AC 2009-224: ENGINEERING ETHICS CURRICULUM INCORPORATION
METHODS AND RESULTS FROM A NATIONALLY ADMINISTERED
STANDARDIZED EXAMINATION: BACKGROUND, LITERATURE, AND
RESEARCH METHODS

Brock Barry, United States Military Academy
Brock E. Barry is a post-doctoral research assistant in the School of Engineering Education at Purdue University. Dr. Barry received his Ph.D. in Engineering Education from Purdue University and holds a B.S. degree in Civil Engineering Technology from the Rochester Institute of Technology and a M.S. degree in Civil Engineering from the University of Colorado at Boulder. Dr. Barry has accepted a position as an Assistant Professor within the Department of Civil & Mechanical Engineering at the United States Military Academy at West Point. He is a licensed professional engineer in four states and has 10+ years of consulting engineering experience. His research interests include engineering ethics, assessment, motivation theory, and integration of professional skills in the engineering curriculum.

Matthew Ohland, Purdue University
Matthew W. Ohland is an Associate Professor in the School of Engineering Education at Purdue University and is the Past President of Tau Beta Pi, the engineering honor society. He received his Ph.D. in Civil Engineering from the University of Florida in 1996. Previously, he served as Assistant Director of the NSF-sponsored SUCCEED Engineering Education Coalition. He studies longitudinal student records in engineering education, team-member effectiveness, and the implementation of high-engagement teaching methods.
Abstract

The ethics literature within the engineering arena is long on opinions, but short on evidence as to the most effective curriculum models for incorporating an understanding of professional and ethical responsibility. Research related to professional ethics has primarily focused on assessment of student learning, rather than evaluation of curriculum integration methods. A limited number of studies have been published that compare two methods of curriculum integration, yet no rigorous studies that compare multiple methods of curriculum incorporation are known to exist. Without clear evidence of best methods, the debate will continue, and there will be no assurance that the methods currently in use are the most effective.

Within this paper, a recently completed research program is described that evaluated the methods of assimilating ethics into the engineering curriculum to determine if a relationship exists between the curriculum models and the outcome on a nationally administered examination, engineering-specific standardized examination. The study’s population was engineering students during the time period between October 1996 and April 2005 enrolled at nine academic institutions in the southeast United States for which valuable data are available.

A mixed-methods (quantitative and qualitative) research program was designed and executed. The qualitative aspects of the study focused on research questions related to the impetus and considerations given to curriculum changes made by the twenty-three discrete engineering programs that participated in the study. The qualitative research questions were investigated through a process of semi-structured interviews conducted with program representatives and evaluation of an extensive number of ABET Self-Study accreditation documents. Once the curriculum models utilized by the participating programs were identified and defined for the chronological limits of the study, a quantitative process was implemented to compare the curriculum models to performance on the ethics section of the Fundamentals of Engineering Examination. A student-level dataset of subject scores was obtained for all administrations of the Fundamentals of Engineering Examination for each of the participating programs. Statistical techniques were utilized to evaluate the relationship between curriculum methods and examination performance.

This paper will provide a statement of the perceived educational issues and a comprehensive summary of the applicable literature. A detailed discussion of the study’s design and implemented methods will be presented. Subsequent publications will present the findings, discussion, and implications resulting from the completed study. This study was executed to fulfill dissertation research requirements associated with doctoral program in Engineering Education at Purdue University.
Introduction

Objectives
At the micro-ethical level, the ethical practices of an individual define, to a large part, how that individual is perceived by his or her peers. At the macro-ethical level, however, the ethical standards inside a particular organization or profession serve to define the nature of the represented group. Within the various professions, a set of standard ethical practices is expected. Although these standards are typically only loosely defined, those actively working in a particular profession are held accountable to the standards by their peers and by the expectations of society.

Graduates of accredited engineering programs in the United States are expected to leave their colleges and universities with the knowledge, skills and confidence to enter a professional career as an engineer. Engineering is similar to the professions of business, law, and medicine in that individuals practicing in these arenas have the potential to affect society and the communities they serve. Specifically, an engineer creates products and processes that can have considerable impact on human health and life.

Engineers that pursue professional licensure are bound by law to practice in a professional and ethical manner that protects human health and life. However, professional licensure is pursued by a minority of individuals in each of the sub-disciplines of engineering. Accordingly, only a fraction of engineering graduates seek and obtain licensure. With that in mind, one might ask, if practicing, unlicensed engineers are not bound by law to practice professionally and ethically, what is it that compels them to do so? Perhaps the naive answer might be a reliance on the general “goodness” of human beings. Alternatively, one might anticipate that engineering programs would incorporate some form of professional and ethical education into their curricula.

The U.S. Department of Education states that the “goal of accreditation is to ensure that education provided by institutions of higher education meets acceptable levels of quality” (p. 50)\(^1\). Degree granting engineering programs in the United States, seeking to initially obtain or maintain accreditation, are required to meet the curriculum content and learning outcomes established by ABET, Inc. (formerly known as the Accreditation Board for Engineering and Technology, Inc.). ABET traditionally makes relatively minor revisions to the accreditation process and criteria on an annual basis. However, during the mid-1990s a significant paradigm shift was conceived by ABET to substantially revise their curriculum criteria with release of their Accreditation Policy and Procedure Manual\(^2\) for the 2000-2001 accreditation cycle. This revision, which included a move to an outcomes-based program evaluation, is commonly referred to as Engineering Criteria 2000 (EC 2000). Criterion 3 of EC 2000 is a list of eleven technical and professional outcomes that accredited programs are required to demonstrate their students attain. A component of ABET EC 2000, Criterion 3.f, includes knowledge gained in the area of professional and ethical responsibility. Accordingly, all accredited engineering programs must now provide their students with some form of professional and ethical instruction to meet ABET standards. Prior to the implementation of EC 2000, specific curriculum content in the area of professional and ethical responsibility was not required.
Multiple methods for incorporating professionalism-based and ethics-based teaching and learning into the engineering curriculum are presently being used by universities and colleges. This study investigated what methods were being utilized within the engineering programs at a particular group of academic institutions and during a particular time frame. Representatives of those engineering programs were interviewed and accreditation documents were reviewed as part of the investigation. Further, an examination designed to measure engineering competency was utilized as the assessment tool to evaluate the various professionalism and ethics incorporation methods. The study population was engineering students from nine academic institutions in the southeast United States. The academic institutions were members of Multiple-Institution Database for Investigating Engineering Longitudinal Development (MIDFIELD). Individual academic institutions and MIDFIELD are described in greater detail subsequently.

The assessment tool utilized in the executed study was a nationally administered, engineering-specific examination generated and governed by the National Council of Examiners for Engineering and Surveying (NCEES). NCEES is an organization composed of licensing boards representing all U.S. states and territories. Engineering graduates who aspire to professional licensure must obtain an ABET-accredited degree, pass the NCEES examination known as the Fundamentals of Engineering (FE) Examination, complete a required number of years of practical experience, and pass the NCEES examination known as the Principles and Practice Examination. The FE Examination is the only examination designed to align with the intended knowledge gained as part of an ABET-accredited program.

The policy that prompts engineering programs to incorporate ethics within their curricula is generated by ABET and enforced through an accreditation review process. The study detailed herein has the potential to influence the means by which the engineering community assesses professional competency and/or the policy that the engineering community uses to hold engineering programs accountable for instilling an understanding of professional and ethical responsibility in their graduates.

The research discussed herein was designed in a manner that permitted an effective and efficient means of addressing the stated research questions, yet also permitted the collection of data for follow-on research.

Framework
The study presented within this document utilizes a mixed-methods approach grounded in the principles of curriculum evaluation theory. The mixed-methods approach uses both quantitative and qualitative research techniques. Miles & Huberman\(^3\) note that mixed-methods studies “can help sequentially (results of the first method inform the second’s sampling, instrumentation, etc.) and can expand the scope and breadth of a study by using different methods in different components” (p. 41).

In addition to utilizing mixed methods, this document represents the efforts associated with the first two steps in what could be a more extensive research program. The steps of this research program are sequential in that the second step could not be initiated without completion of the first. Accordingly, the research questions in the subsequent section of this document (Research Questions and Hypothesis) are presented in a manner that relates specific inquiry to each step in
the research program. The research questions associated with the first-step in the research program were best suited for qualitative methods of discovery. In turn, the findings of the first step of this research program helped to inform and shape the research question related to the second step. The second-step research question was best suited for quantitative methods of discovery.

No testable theories related to ethics curriculum incorporation methods are known to presently exist. Thus, this research was conducted in an emergent manner that allows for the possible development of a curriculum-related theory as a result of the study. In addition, the investigation has been designed with a pragmatic approach to addressing the research questions. As a result, the intent has been to produce a study whose quality is judged by “…its intended purposes, available resources, procedures followed, and results obtained…” (p. 71)⁴.

Qualitative Framework
The qualitative framework for this study is based on the concept of applied reality testing. This study attempted to understand the nature of conditions that are deemed important to society. This framework is commonly employed in evaluation-based research,⁴ such as the study presented within this document.

The nature of the research questions associated with reality testing typically attempt to determine what the reality of a particular situation is. For the purposes of this study, the “reality” that is being investigated in the first-step of the research, is to determine what curriculum methods are being used among particular engineering programs to incorporate accreditation requirements related to professional and ethical responsibility. In addition to determining what methods are in current use, the reality testing framework approach also supports a desire to determine if changes in curriculum methods have occurred, when they occurred, and why they occurred. To address those questions, two traditional qualitative methods of investigation were implemented and are discussed in greater detail subsequently.

Quantitative Framework
The quantitative framework for this study was shaped based on what Green & Stone⁵ describe as the Outcome Model of curriculum evaluation research. In support for the Outcome Model, Green & Stone offer the following statement:

Too few systematically evaluate the impact of their curriculums [sic] in terms of the satisfactions and dissatisfactions of their graduates and their levels of professional or vocational competence. How successful and how competent are the graduates of the universities’ vocational or professional curriculums? (p. 63)

The Outcome Model has been implemented in this study using a summative approach to appraise and compare various curriculum methods. A graphical reproduction of the Outcome Model, as presented by Green & Stone is illustrated in Figure 1. This study begins in the box labeled the “The Graduate,” and specifically relates to the professional estimate of competence. The instrument used in this study to compare between curriculum methods is the FE Examination, an engineering-specific examination designed as a measure of minimal competency to practice as a professional engineer.
Within the model, the next set of concurrent components is “Overall Program Goals” and “Specific Curriculum Objectives.” A statement of program goals is a required section of ABET Self-Study documents prepared by individual engineering programs as part of an accreditation review. Such program goals are then mapped to specific curriculum objectives. In the case of ABET-accredited engineering programs, those curriculum objectives are defined by the ABET EC 2000 documents and include a criterion directly related to professional and ethical responsibility.

The overall program goals and specific curriculum objectives feed into the “Decisions to be Made” component of the Outcome Model. In this case, the decision to be made is the curriculum method adopted at a program level to satisfy the professional and ethical responsibility aspects of ABET accreditation. As noted, those decisions are influenced by both “Input from the Profession” and “Input from Society.” The input received from the profession can come in many forms; including program advisory councils, alumni surveys, and educational policies developed by technical and professional organizations. Input from society to the program-level tends to be less direct, but in the case of professional and ethical responsibility, society does tend to define normative standards.

The decisions made within the Outcome Model then feed into several aspects of the program. Of particular interest to this study is the potential for “Revised Curriculum Plans and New Approaches to Implementation.”

Finally, the Outcome Model is a continuous loop in the recognition that changes made to a program will influence the student at various levels: “The Beginning Student,” “The Senior Student,” and the “The Graduate.” Again, for the purpose of this study, the primary interest has been in the graduates’ competency related to professional and ethical responsibility.
Research Questions and Hypothesis
To address the problem statement made previously and in accordance with the mixed-methods approach, the following two distinct sets of research questions were generated:

Research Step 1 – Qualitative Investigation
- How have the incorporation methods utilized by the MIDFIELD institutions for accomplishing ABET Criterion 3.f changed during the time period between 1995 and 2005?
- If changes in the incorporation methods have occurred, what general trends can be recognized?
- If changes in the incorporation methods have occurred, what change agents prompted the shift in incorporation method?
- What are the primary considerations that MIDFIELD program representatives cite relative to selecting a particular incorporation method?
- What type of classification scheme would be most appropriate for describing the ABET Criterion 3.f curriculum incorporation methods utilized by the MIDFIELD institutions?

Research Step 2 – Quantitative Investigation
- Having knowledge of the methods utilized by the MIDFIELD programs to accomplish ABET Criterion 3.f, does a statistically significant relationship exist between the amount of required applicable content and student performance on the professional ethics section of a nationally administered, engineering-specific standardized examination?

A valid, testable, and relevant hypothesis has been generated to give direction to the quantitative research question. The null hypothesis is stated as follows:
- No statistically significant relationship exists between the required applicable content utilized by the MIDFIELD academic institutions to accomplish ABET Criterion 3.f and student performance on the professional ethics section of a nationally administered, engineering-specific standardized examination.

Delimitation and Limitations of Study
Performance of this study involved a review of student-generated scores on the FE Examination administered by the National Council of Examiners for Engineering and Surveying (NCEES). There are only two nationally administered, engineering-specific examinations, the FE Examination and the Principles and Practice Examination. The FE Examination is the only examination with content which is focused on the intended knowledge gained in ABET-accredited engineering programs. The Principles and Practice Examination is traditionally administered to graduates of ABET-accredited engineering programs after a requisite number of years in professional practice. Scores for the FE Examination were obtained for the period between and inclusive of October 1996 and April 2005. During this period, the FE Examination used a consistent format and allocation of content. 6, 7

The academic institutions used in this study are members of the Multiple-Institution Database for Investigating Engineering Longitudinal Development (MIDFIELD) program. The nine MIDFIELD member institutions are: Clemson University, Florida A&M University, Florida State University, Georgia Institute of Technology, North Carolina A&T State University, North Carolina State University, University of Florida, University of North Carolina at Charlotte, and Virginia Polytechnic Institute and State University. It should be noted that engineering students
from Florida A&M University and Florida State University attend a joint College of Engineering for their engineering coursework.

This study was limited to three engineering sub-disciplines common to the nine MIDFIELD institutions; civil engineering, electrical engineering, and mechanical engineering. From a pragmatic standpoint, limiting to these three engineering sub-disciplines ensures representation of the three most common engineering sub-disciplines, while maintaining a reasonable population size. Among all engineering sub-disciplines, civil engineering, electrical engineering, and mechanical engineering have the highest percentages of licensed engineering graduates at 44%, 9%, and 23%, respectively. This study did not include engineering technology programs. While most of the MIDFIELD academic institutions do not have engineering technology programs, at the few MIDFIELD academic institutions where engineering technology programs are offered, those programs are housed in departments separate from the engineering programs and they utilize disparate curricula.

The chronological limits of this study have been established as 1995 to 2005. Within that time period most of the participating academic programs have gone through a minimum of one ABET accreditation review prior to implementation of EC 2000 and a minimum of one ABET accreditation review subsequent to implementation of EC 2000. The anticipated exceptions to this condition were programs that voluntarily adopted the EC 2000 criteria before the year 2000 implementation. Programs that voluntarily adopted the EC 2000 criteria early may have undergone two accreditation reviews using the EC 2000 criteria during the chronological limits of this study.

This study intentionally omitted particular subject matter. The population was not analyzed relative to gender, race, or socio-economic status. This study was not concerned with ranking MIDFIELD institutions or programs by their performance on the FE Examination. Accordingly, results of this study are reported without institution identifiers. This study also did not attempt to make a connection between professional and ethical understanding and practice. As Pfatteicher notes, ABET Criterion 3.f does not require that engineering programs demonstrate that graduates are ethical. Rather Criterion 3.f requires that graduates understand professional and ethical responsibility. However, it should be noted that some academic institutions aspire to outcomes that specifically include a demonstration of ethical responsibility.

This study has inherent limitations as the result of its design. It is often difficult to make particularly strong conclusions from a study that uses a single quantitative tool as the means of assessment. However, this limitation of the study is needed to address the research questions in a reasonable period of time. Alternatively, a long-term study could have been designed that includes yearly assessment of students using a second tool in addition to the FE Examination.

Curriculum Incorporation Methods

The operational definitions adopted for the curriculum incorporation methods initially used in the design of this study were based on the writing of Joseph Herkert. In particular, Herkert’s article titled Engineering Ethics Education in the USA: Content, Pedagogy and Curriculum provides the basic classification scheme for defining several of the curriculum incorporation methods.
• Required course in engineering ethics within the program: Commonly conducted as a full-semester, multiple-credit class, which all students within a given discipline are required to complete. This model allows the class content to focus on discipline-specific issues.

• Required course in ethics external to the program: This method relies on course offerings from philosophy or religion departments and exposes students to a more general ethics background, while sacrificing the discipline context captured in the “within the program” method.

• Across-the-curriculum: This approach exposes students to ethical considerations repetitively, in multiple courses, during a progression towards their degree. This method requires commitment among all department faculty to capitalize on ethics discussions within traditional non-ethics focused courses.\(^\text{11}\) This research study assumes that via this method ethics is discussed in a minimum of four separate courses.

• Linking ethics with society: This approach uses a curriculum model with a series of required core courses that strongly emphasizes professional ethics and the role of the discipline within society. This model is found in programs that design their curriculum around the principles of professionalism and society. Typically, these programs incorporate service-learning activities within the curriculum.

Prior to the adoption of ABET EC 2000, it is possible that accredited engineering programs simply did not incorporate professional ethics within their curriculum. Thus, a program which made no concerted effort to incorporate professional ethics would fall into the “none” category. Although Herkert did not explicitly identify this category, it was used in the research design.

**Literature Review**

**Introduction**

There are several methods of incorporating an understanding of professional and ethical responsibility within curricula currently used by engineering programs. Much consideration has been given to those methods as a result of ABET’s introduction of EC 2000, which for the first time required that accredited engineering programs infuse their curricula with professional and ethical responsibility to achieve a particular aspect of the outcomes-based EC 2000 criteria.

Accordingly, leading up to and subsequent to the EC 2000 implementation, the engineering education literature has generated a sizeable volume of material on the subject of curriculum models. Interestingly, the education literature associated with several other professions, including medicine, law, and business, also includes an extensive amount of publication on the subject of professional ethics curriculum incorporation methods. The interested reader is encouraged to consult the Journal of Engineering Education article by Barry & Ohland\(^\text{12}\) as well as the full dissertation by Barry\(^\text{13}\) for further details.

This literature review contained herein focuses on the engineering education literature relative to curriculum models and assessment using a nationally administered examination, as well as a review of prior research efforts associated with each of those concepts.
Research-Specific Literature

ABET EC 2000
In 1991, John Prados, then President of ABET, delivered a report to the ABET Board of Directors. The focus of that report was the perceived challenges facing ABET and engineering education. The take-away message from Prados’ report, supported by many engineering educators and members of industry, was that “engineering education must change significantly to support this new environment, and that ABET’s rigid application of the accreditation criteria created a significant barrier to the needed innovation.”

In response to this call for change, ABET established the Accreditation Process Review Committee (APRC) in 1992 and charged the committee with the task of advising ABET how to “increase flexibility in the engineering accreditation criteria, while maintaining a strong emphasis on educational quality.” In 1995 the APRC issued a report titled The Vision for Change on behalf of ABET, The National Science Foundation, and the ABET Industry Advisory Council. The APRC report summarized a series of consensus building workshops conducted between May and August 1994. Out of those workshops, and the collective subsequent efforts of ABET and its affiliates, a significant revision to the accreditation review process developed. For the first time, the ABET accreditation process was structured around outcomes, assessment, and continuous improvement rather than prescriptive and detailed curricular specifications.

What is now commonly referred to as Engineering Criteria 2000 (EC 2000) was initially approved for publication in October 1995 by the ABET Board of Directors. Given the sweeping changes proposed in EC 2000, ABET wisely released the ground-breaking criteria for public review and comment well in advance of the proposed implementation. Final authorization for EC 2000 was granted on November 2, 1996.

As noted previously, Criterion 3.f of EC 2000 states that students must be able to demonstrate “an understanding of professional and ethical responsibility” (p. 32). The Criterion 3.f wording has remained unchanged between the original implementation in 2000 to the approved 2008-2009 criteria. Prior to the approval and implementation of EC 2000, the ABET accreditation criteria used less definitive and enforceable language in Criterion IV.C.3.j which stated:

An understanding of ethical, social, economic, and safety considerations in engineering practice is essential for a successful engineering career. Course work may be provided for this purpose, but as a minimum it should be the responsibility of the engineering faculty to infuse professional concepts into all engineering course work.

Five engineering programs voluntarily went through accreditation review using the EC 2000 criteria during the 1996-97 and 1997-98 academic years. Another 104 programs were evaluated using the EC 2000 criteria during the transition years between the 1998-1999 and the 2000-2001 academic years. At present, all ABET-accredited engineering programs have undergone at least one review using the EC 2000 criteria.

Incorporating ABET Criterion 3.f in the Engineering Curriculum
While engineering educators, in general, supported the premise of an outcomes-based accreditation system, the less prescriptive process generated by EC 2000 raised many questions and concerns relative to methods for teaching and incorporating the professional skills in the
engineering curriculum. The landmark article by Shuman, Besterfield-Sacre, and McGourty\textsuperscript{16} provides an exceptional overview of teaching and assessment methods for all six EC 2000 professional skills. Multiple articles, released both in advance and subsequent to the introduction of EC 2000, focused on the teaching, assessment, and curriculum incorporation challenges associated with Criterion 3.f.\textsuperscript{9-11, 16, 25-34}

Whereas the methods of ethics instruction and assessment are often left to the discretion of the instructor/professor, methods of curriculum incorporation are typically established at the department and/or institutional level. The predominant methods of integrating ethics within curricula include: required courses within the discipline, elective courses outside the discipline, across-the-curriculum, and the linking of ethics with society.\textsuperscript{10, 29}

The discussion of ethics curriculum incorporation within the engineering literature is marked by a wide variation in opinions relative to best methods. Herkert\textsuperscript{10} notes his preference for the linking of ethics with society approach. Drake, Griffin, Kirkman, & Swann\textsuperscript{35} show support for a required course in the engineering discipline. Three Hastings Center commissioned reports, discuss a variety of curriculum incorporation procedures used in engineering.\textsuperscript{36-38} Ultimately, the Hastings Center reports advocate for the use of a stand-alone course in ethics. However, the reports then discuss the merits of, but fall short of supporting, the across-the-curriculum approach. It is worth noting that the literature, at times, blurs the lines between instructional methodology and curriculum incorporation. For example, Haws\textsuperscript{39} claims to be reviewing pedagogical approaches in his meta-analysis, but one of the methods he discusses, service learning, is viewed by others\textsuperscript{10, 29} as a component of a curriculum incorporation approach.

There appears to be a significant amount of debate related to methods of incorporation, but very little rigorous investigation to determine which methods of curriculum incorporation are most common or most effective. Drake et al.\textsuperscript{35} performed a comparison of two methods of integrating ethics in the engineering curriculum, a stand-alone course and several modules within general engineering courses (quasi-across-the-curriculum). This rare comparison of two curriculum incorporation methods was unsuccessful in generating conclusive statistics in support of either method. Discussion and debate over curriculum incorporation methods is not a recent phenomenon. A 1955 \textit{Journal of Engineering Education} article evaluated various aspects of a stand-alone ethics course within the engineering curriculum versus a quasi-across-the-curriculum method.\textsuperscript{40}

\textit{Assessment of EC 2000 Criterion 3}

The technical and professional skills encompassed within ABET’s EC 2000 Criterion 3 are presented in a manner that allows for flexibility in implementation. Besterfield-Sacre et al.\textsuperscript{41} note that by design, the Outcomes are vaguely constructed to “encourage each engineering program’s faculty to add its own, hopefully unique specificity” (p. 100) and further, that the flexibility “reflects a sensitivity on ABET’s part to the importance of differing institutional missions and programs” (p. 100). The downside to the flexibility of interpretation is: “ambiguities in defining and measuring the skills that students must demonstrate if a program is to meet the intent of the criteria.”(p. 2).\textsuperscript{42}
Strauss & Terenzini\textsuperscript{42} note that in response to the implementation of EC 2000, a large number of associated assessment instruments have emerged, but most fail to meet accepted test development standards. Further, the vast majority of the existing EC 2000-related assessments measure an individual a-k criterion. Strauss & Terenzini were unable to identify an existing “psychometrically sound, comprehensive, and practical measure of all of the learning outcomes” (p. 4) listed in EC 2000. Accordingly, in 2002 ABET commissioned a study through the Center for the Study of Higher Education at Penn State University, in which Strauss & Terenzini set out to develop such an instrument with the ultimate intent of utilization by accredited engineering programs. They generated a 36-item measure of student performance on all 11 EC 2000 outcomes.\textsuperscript{42, 43} However, despite the practical nature and good psychometric properties, this instrument has not found wide-spread use among engineering programs preparing for ABET accreditation.

**Impact of ABET EC 2000**

Lattuca, Terenzini, and Volkewin\textsuperscript{44} conducted a three-year investigation titled *Engineering Change: A Study of the Impact of EC2000 (Engineering Change)* to determine if post-EC 2000 engineering graduates are “better prepared to enter the profession than were their pre-EC 2000 counterparts” (p. 1). The ABET funded study looked at the impact of EC 2000 relative to student learning outcomes and organizational and educational policies and practices. The conceptual framework for this study suggested that program change, including curriculum change, would ultimately lead to changes in student learning. Thus, the *Engineering Change* study attempted to investigate similar concepts as the study detailed herein, although via a different framework and methods. The *Engineering Change* study generated an extensive number of interesting findings, including many related directly to the study detailed herein, and it concluded that the implementation of EC 2000 has had a positive impact on student learning.

**Fundamentals of Engineering Examination**

While the ABET APRC tackled the issue of revising the accreditation criteria, another group of stakeholders was considering methods of assessing the knowledge of engineering students. The Joint Task Force on Engineering Education Assessment (Task Force) was formed in 1994 with the mission of examining the value of using the FE Examination in undergraduate engineering education assessment.\textsuperscript{45} The Task Force consisted of representatives from NCEES, ABET, the American Society for Engineering Education (ASEE), the ASEE Engineering Deans Council (EDC), and the National Society of Professional Engineers (NSPE). The Task Force determined that the FE Examination should be “restructured so that it more broadly measures outcomes of the total engineering education experience.”\textsuperscript{45} Ultimately, this group was suggesting a closer alignment between the, then pending, EC 2000 outcomes and the content of the FE Examination.

While NCEES presently advocates for the use of the FE Examination as a general research and assessment tool, they acknowledge that the examination’s primary function is to measure minimum technical competency required for professional licensure of engineers.\textsuperscript{7, 46, 47} The primary construct of the examination appropriately remains at the forefront among those that suggest using it for other purposes.\textsuperscript{8} The American Society of Civil Engineering (ASCE)\textsuperscript{48} issued a statement advocating that the FE Examination should not be modified to meet additional objectives, such as program assessment and research. However, ASCE also notes that “…FE exam results could provide valuable information to educators.”\textsuperscript{48}
In addition to their role in the Task Force, ABET appears to be supportive of the use of the FE Examination as an outcomes assessment tool. The *Criteria for Accrediting Engineering Programs* 2000-2001 through 2003-2004 each carried the following language:

The assessment process must demonstrate that the outcomes important to the mission of the institution and the objectives of the program, including those listed above, are being measured. Evidence that may be used includes, but is not limited to the following: student portfolios, including design projects; **nationally normed subject content examinations** [emphasis added]; alumni surveys that document professional accomplishments and career development activities; employer surveys; and placement data of graduates.  

The FE Examination is the only nationally normed, engineering-specific examination that could satisfy that statement. Notably, the entire section quoted above has been struck from the ABET Criteria beginning with the 2004-2005 version. In addition, starting with the 2004-2005 ABET Criteria, all such lists of suggested accreditation methodologies were removed to alleviate the perception that ABET was being prescriptive. Colby & Sullivan note, the ABET guidelines do not specify particular “strategies to meet the requirements” (p. 327) of ABET Criterion 3.f. It should also be noted that the FE Examination is “normed” against peer institutions during program evaluation. As discussed subsequently, the examination is a criterion-referenced, not a norm-referenced, assessment.

The FE Examination has been offered since 1956 and has undergone a variety of changes. Upon the recommendation of the Task Force, the FE Examination was changed in 1996 from a 8-hour general engineering content examination to a 4-hour morning session that covered general engineering content and a 4-hour afternoon session that provided the test-taker the option of choosing content from among several engineering sub-disciplines. Those changes were implemented during the October 1996 test administration. The structure and content distribution of the examination remained unchanged until a new set of examination specifics were implemented in October 2005.

The pass-fail examination is administered twice a year (spring and fall) to an approximate total of 40,000 individuals. The examination consists of 120 multiple-choice questions in the morning sessions and 60 multiple-choice questions during the afternoon session. Throughout the October 1996 to April 2005 administrations, 4% of the morning session questions related to professional and ethical responsibility.

The FE Examination is a criterion-referenced assessment. A criterion-referenced test uses a specified content domain as its interpretive reference. Relative to grading, NCEES notes that scoring for each exam is “based on its own merits with no regard for a predetermined percentage of examinees that should pass or fail.” In contrast, norm-referenced testing interprets an individual’s test score by comparing it to the results obtained by other test takers. A passing score for the FE Examination is calculated as the number of correct answers required to indicate a minimum level of competency. The passing score is determined by a panel of subject-matter experts.

Although the FE Examination was designed to measure competency to practice as an engineer, the test has been successfully applied to other forms of assessment. LeFevre, Steadman, Tietjen,
& White\textsuperscript{7} authored a report on behalf of NCEES that discusses the use of the FE Examination in program-level assessment. Steadman\textsuperscript{52} also discusses the possible use of the FE Examination for program assessment, but stresses that assessment is a process and cannot rely completely on the results of a single examination. In fact, the Task Force, while advocating for the application of the examination as an assessment tool, noted that it should be used as one part of a larger process.\textsuperscript{45} LeFevre\textsuperscript{6} demonstrates how the FE Examination was successfully implemented as part of a program review process as the University of Arkansas.

Nirmalakhandan, Daniel, & White\textsuperscript{53} note that the FE Examination could be used for not only program-level assessment, but also for intra-institution comparisons. They state, “One of the uses of the results would be in comparing the performance of students in an institution against the peer performance on a subject basis to gauge the effectiveness of individual or series of courses”(p. 73).\textsuperscript{53} They go on to discuss the methods behind the successful application of an intra-institution study using the FE Examination. Wicker, Quintana, & Tarquin\textsuperscript{54} discussed a similar study performed using FE Examination data for more than 50 universities.

In the only known publication of a program-level assessment study of engineering ethics utilizing the FE Examination, Davis & Butkus\textsuperscript{55} describe an attempt to measure changes in student performance over a four-year period. Terenzini, Lattuca, Ohland, & Long\textsuperscript{56} have presented a design to combine a database generated during the previously referenced ABET commissioned study of the impact of EC 2000 and a database of FE Examination results. Similar to the study detailed within this document, the study proposed by Terenzini et al. is designed to evaluate possible correlations between student learning outcomes and program-level examination performance.

While some researchers and academicians advocate for the use of the FE Examination in program assessment, the community does not speak with one voice on this issue. For example, Drnevich & Tener\textsuperscript{57} expressed opposition to using the examination as a mandatory program assessment tool and Schimmel, King, & Ilias\textsuperscript{58} recommend the use of department-developed assessment tools rather than the FE Examination. Some of the most common knocks against using the examination for program assessment and research include: the examination measures minimum competencies, the examination scores are influenced by variable student motivation, the small number of individuals that take the examination is not representative of the full population, the examination is administered before graduation (prior to coverage of certain content), and the examination constricts programs from developing according to unique mission statements.\textsuperscript{8, 53, 54, 57, 58} Each of these issues ultimately influences the reliability and validity of the FE Examination as an assessment tool.

An article authored by Lawson\textsuperscript{8} provides an extensive analysis of the examination score’s reliability and validity. Specifically, Lawson scrutinizes the examination relative to its application for assessing three discrete constructs: individual competency, program accreditation, and college performance. Specific to reliability, Lawson notes that the examination is machine graded, thereby eliminating the possibility of grader bias and reducing the possibility of grading error. Further, the examination consists of a relatively large number of questions, each of which have undergone statistical performance analysis to reduce the potential for external error.\textsuperscript{47} Lawson states that the use of an external, absolute standard for passing (criterion-reference)
creates a strong argument for face validity. Given that the examination was initially developed to assess individual professional competence to practice as a licensed engineer, it is not surprising that Lawson concludes that examination scores are adequately reliable and valid.

Lawson then moves on to the second application of the FE Examination, program-level assessment. Lawson notes that the relatively low number of students taking the examination makes it non-representative of the engineering population and thus, unreliable for program assessment. Lawson also discusses a weakness of the examination’s validity in that it measures only minimal technical competency, which he states is “hardly the goal of an engineering program strategically focused on quality and excellence…” (p. 323). Ultimately, Lawson notes that the FE Examination is most reliable and valid for program-level assessment when a large percentage of graduates take the examination and when a discipline-specific examination exists to match the mission of the program.

The analysis is completed with a discussion of college-level assessment using the FE Examination. Lawson identifies the not-too-uncommon desire of the key constituencies of higher education to seek “clear, simple, and accurate indicators that demonstrate the value and quality of higher education” (p. 324). Performance funding is a tempting issue to connect with examination results. In his discussion, Lawson addresses reliability and validity by drawing parallels between the engineering profession and the professions of medicine and law. He notes that a significant difference between these professions is that physicians and attorneys must pass an examination equivalent to the FE Examination to practice in their field, whereas engineers may practice (with limitations) without professional licensure. Finally, Lawson identifies that the multiple-discipline format of the FE Examination makes it inherently less reliable than the more uniform medical and law examinations.

In summary, Lawson methodically addresses the reliability and validity of FE Examination scores with respect to their application of three levels of assessment. Assessment of individual competency is shown to be an appropriate application of the examination. Program-level assessment is suggested to have the potential for weaknesses in reliability and validity. Lawson makes a strong case against using the FE Examination for college-level assessment.

Lawson’s article has significant bearing on this research. The performed curriculum comparison is similar to what is defined as a multiple-program assessment. The academic institutions included in this study were not expected to have examination participation rates sufficiently high enough to overcome the reliability issues associated with a non-representative examinee population. The argument can be made that no other means of engineering-specific assessment, with or without higher participation rates, currently exists. Thus, in spite of the perceived weakness of the FE Examination, it remains the soundest instrument for this type of research.

When considering a detailed analysis of a single aspect of the FE Examination in relation to a single ABET outcome, particular emphasis on the examination’s construct validity is important. That is, is the FE Examination an appropriate measure of ABET Criterion 3.f; a student’s ability to “understand professional and ethical responsibility?” Due to the security required to protect the database of FE Examination items, a direct inspection of professional ethics questions was
not possible. In lieu of a direct inspection of actual examination items, the content of four of the most widely used FE Examination review manuals were evaluated\textsuperscript{59-62}. The discussion sections of each of these examination review texts focused on the NCEES Model Rules of Professional Conduct (code of practice); the obligation of the professional engineer to society, employers and clients, and other engineers; and ethical and legal standards. These four references collectively offer 123 practice questions for the professional ethics section of the FE Examination. An item-by-item inspection of these practice questions indicates that they align well with the investigator’s interpretation of the EC 2000 Criterion 3.f, as well as the interpretation offered by Shuman et al.\textsuperscript{16} If the content of these review texts are representative of the actual examination content, as they claim to be, then it can be comfortably stated that the ethics section of the FE Examination is an appropriate assessment of Criterion 3.f.

A copy of the Exam Development Procedures Manual: Exam Development, Scoring, and General Procedures\textsuperscript{47} and the Security and Administrative Procedures Manual\textsuperscript{49} were provided by NCEES for review as part of this study. Downing\textsuperscript{63} notes that security of tests and test items is a constant concern during all steps of development. Accordingly, NCEES takes extensive steps to maintain the security of the FE Examination during all stages of item development, test preparation, administration, and scoring. Included in those security measures is the treatment of the aforementioned manuals as confidential documents. Thus, the contents of these manuals cannot be fully detailed as part of this research. However, at a minimum it can be revealed that the manuals contain extensive information on exam development, equating, administration, security and scoring, as well as evidence of appropriate psychometric characteristics.

Studies across Disciplines and across Institutions
The research detailed within this document includes an assessment of students’ understanding of professional and ethical responsibility across disciplines and across institutions. Several studies have previously been conducted to determine if academic major (discipline) and/or institution environment affect the development of moral reasoning.

In the landmark study by Pascarella & Terenzini,\textsuperscript{64} the authors make the case that the general experience of attending college is linked with statistically significant increases in the use of principled reasoning to judge moral issues. They state that “…college attendance is associated with a humanizing of values and attitudes concerning the rights and welfare of others” (p. 64).\textsuperscript{64} Classical moral development literature also supports the concept that moral development can be stimulated through educational intervention.\textsuperscript{65-67} Pascarella & Terenzini go on to say that “…the college experience itself has a unique positive influence on increases in principled moral reasoning” (p. 347).

Pascarella & Terenzini identify several studies that conclude significant differences in moral reasoning can result from a disparity in an academic institution’s mission. Good & Cartwright\textsuperscript{68} conducted a study with a comparison between three academic institutions: a state university, a Christian liberal arts university, and a Bible university (an academic institution that emphasizes Bible courses, communication courses, and ministry vocation preparation courses). The results of the Good & Cartwright study indicate no initial statistical difference in moral reasoning among freshmen between the three universities. However, when comparisons in freshman-to-senior gains are made, the state university and Christian liberal arts university produced greater gains.
than the Bible university. McNeel\textsuperscript{69} also performed a study that looked at an institution’s effect on moral development and identified variability in moral development based on mission of academic institution.

Some evidence has been found for the connection between academic major and moral development. Jeffrey\textsuperscript{70} performed a study that compared the ethical development of accounting students, non-accounting business students, and liberal arts students. Within that study, accounting students generated the highest moral reasoning scores on a standardized assessment. Cummings, Dyas, Maddux & Kochman\textsuperscript{71} compared the moral reasoning ability of education majors to that of several other majors. The results of that study indicated that education majors had lower moral reasoning skills. McNeel’s\textsuperscript{69} study across disciplines showed that psychology, nursing, and English majors obtained higher moral reasoning scores than education and business majors. From that study McNeel concluded “it has become increasingly clear that moral issues are integrally bound up in the content of the various disciplines…” (p. 28). McNeel also noted that the “highest growth took place in majors that focus on understanding humans and/or majors that include a central integration of ethical considerations within the content of a professional course of study” (p. 34). Surprisingly, Jeffrey does not believe that score variation across disciplines is the result of curriculum content, whereas, both Cummings, et al. and McNeel contend there is a connection between ethics curriculum incorporation and moral development. Pascarella & Terenzini\textsuperscript{64} believe that the body of evidence is currently too small to conclude that academic major affects moral development. Pascarella & Terenzini’s contentions appear to be supported by the meta-analysis by Borkowski and Ugras.\textsuperscript{72} Their meta-analysis included 30 studies of undergraduate major and ethical reasoning (number of subjects = 11,818) and generated mixed and inconclusive results.

Literature Limitations
As revealed in this literature review, multiple methods of incorporating ethics within the engineering curriculum are currently in use. What is absent from the literature is any indication that these curriculum incorporation methods have been thoroughly studied for effectiveness in developing students with an understanding of professional and ethical responsibility. Thus, it is possible that the curriculum techniques currently in use within various academic programs are not producing the anticipated outcome. The completed research detailed herein begins to address the gap in the engineering literature.

Research Procedures

Introduction
The research questions, identified previously, have been addressed in accordance with the pragmatically designed procedures detailed herein. In general, information relative to current and historic methods of ethics curriculum incorporation, within each of the study’s participating programs, was gathered through interviews with key personnel and by inspection of ABET accreditation documents. A database of student-level performance on the FE Examination was obtained from NCEES. The database permitted student-level investigation of performance of individual examination topics between curriculum incorporation methods.
The methods employed during this research program consisted of a mixture of evaluation and correlation analysis. A summative evaluation was performed to identify the methods of ABET Criterion 3.f curriculum incorporation in use by the MIDFIELD academic institutions. The correlation analysis was performed to determine if a relationship exists between the amount of content in the required curriculum used to accomplish ABET Criterion 3.f and performance on a standardized examination. The research was exploratory in that it is investigating relatively unknown territory – the applicable literature indicates that research of this type has not previously been executed. Successful performance of exploratory research can lead to better understanding of phenomena, such as variations in the level of understanding of professional and ethical responsibility.

Specific Procedures
The following subsections provide a detailed discussion of the procedures employed in this research. Each subsection is grouped by major content area.

Institutional Review Board
In accordance with Federal, State, and Purdue University regulations for research involving human participants, approval for performance of this study was obtained from the Purdue University Institutional Review Board (IRB) prior to substantial commencement of the study. The research was “conducted in established or commonly accepted educational settings, and involves normal educational practices.” Accordingly, an “exempt” status was obtained. The Purdue University IRB website states that “exempt” status “generally applies to the study of teaching methods, strategies, and curricula in the process of education.”

Confidentiality
In accordance with the agreement established with the MIDFIELD institutions and the requirements set forth in the IRB protocol, this study was conducted in a manner that protects the confidentiality of the participants, at both student-level and institution-level. It is not the intent of this study to highlight any given program as either exemplary or lacking in any manner. The evidence recorded within this document has been presented in a manner that does not readily permit identification of individual students or participating academic programs at the MIDFIELD institutions. Randomly generated codes have been assigned to the participating academic programs. Accordingly, those utilizing this document are discouraged from attempting to make such connections between the evidence and academic programs.

Interviews
Semi-structured interviews were performed with representatives of the civil, electrical, and mechanical engineering programs that agreed to participate in the interview phase of the study. The intent of the interviews was to establish the current and historic method(s) each program utilizes to address ABET Criterion 3.f within their curriculum, as well as discuss reasons why the program opted to utilize a particular method and/or make changes.

A record of contacts for the MIDFIELD academic institutions was assembled as part of this study. The record included the name, title, and contact information for individuals within each program who were likely to be able to assist with the study. These individuals served as the
initial point of contact. The record was continuously updated with information for secondary contacts within the participating programs.

A letter of introduction was sent to each of the initial department contacts. The letter identified the contact’s academic institution as a member of MIDFIELD and that the institution has provided authorization for release of information to this specific study. Further, the letter briefly described the purpose and goals of the research. It clearly stated a request for the opportunity to conduct a brief interview and to subsequently inspect the program’s ABET records. An attempt was also made in the letter to describe how the program’s participation in the study could be useful in the preparation of future ABET accreditation review documentation. The letter provided contact information for Mr. Brock E. Barry; including mailing address, email address, fax number, and phone number.

One week after the introduction letters were mailed out, an individually addressed email was sent as a follow-up to non-responsive contacts. The introductory letter was attached to the follow-up email as a PDF document. Follow-on phone calls and additional emails were generated by Mr. Brock E. Barry approximately one week after the initial follow-up email. In some cases, contacts were prompted for a response by Dr. Matthew W. Ohland. In several instances, the primary program contact delegated the interview and ABET file release process to another program representative. When that occurred, an email was sent to the alternative contact to provide introductory information and to schedule the interview. The introductory letters and follow-up correspondence also stressed the request for ABET documentation. Programs were encouraged to send documentation via email in PDF format, although hard copy was also accepted.

All interviews were conducted over the phone and at the relative convenience of the department contact. In general, phone interviews lasted between 30 and 60-minutes, although time was allotted in the event that they ran longer. Interviews were recorded, with authorization, to permit subsequent transcription. For consistency, all interviews were conducted by Mr. Brock E. Barry.

Program representatives were provided with a copy of the interview script and were encouraged to review the document in advance of the interview. The script begins with a general overview of the research program and nature of the questions. Then an attempt was made to clarify that the interviewee was indeed the appropriate person to speak with relative to historic program knowledge and ability to address the interview questions. The methods of incorporating ABET Criterion 3.f in the curriculum was the primary focus of the interview. However, to provide context for the primary topic of interest, as well as provide content for future research activity, the interview questions also broadly probed the interviewees’ thoughts on professional and ethical responsibility in engineering, the FE Examination, and the ABET accreditation process. The script included multiple open-ended questions to encourage the interviewee to provide depth of content in their response. Each program representative was asked to describe the method or methods of curriculum incorporation that their program used during the study’s chronological limits. Based on the description offered by the representative, Mr. Brock E. Barry then read back the operational definition for the curriculum incorporation method that most closely matches the method described by the program representative. The representative was asked to confirm if the operational definition was an appropriate description of their methods.
The interview process was beta-tested during mock interviews with undergraduate representatives of two separate engineering programs within the College of Engineering at Purdue University. During the beta-testing process, notes were collected and used for making protocol revisions. Beta-testing was completed prior to performance of interviews with the MIDFIELD programs.

Transcription
Recorded interviews were transcribed by Mr. Richard G. Martin of Golden, Colorado under contract to provide such services to Purdue University. Mr. Brock E. Barry proofed all transcripts against the original recordings. As only a portion of each interview directly relates to the current study, the full transcripts have not been included in this research document. Refer to Barry\textsuperscript{13} for a copy of the Interview Summary Sheets.

ABET Self-Study Documents
ABET accreditation reviews are performed at the program-level. However, typically those reviews are executed on a college-wide basis during a coordinated single campus visit. All of the MIDFIELD institutions have maintained accreditation throughout the chronological limits of the study. A direct request was made to ABET to obtain the dates of prior reviews for each of the study’s academic programs, but that information is not publicly available. Accordingly, that data was gathered directly from the participating programs as part of this study. The literature indicates that at least one of the study’s participating programs, Georgia Institute of Technology, voluntarily underwent an institution-wide review, in accordance with the EC 2000 guidelines, prior to the actual implementation.\textsuperscript{15} During performance of the study, efforts were made to identify all participating programs that went through such voluntary reviews during the time period between the initial introduction of EC 2000 in the 1996-1997 academic year and the required implementation during the 2000-2001 academic year.

As part of the ABET accreditation process, programs under review are required to generate a Self-Study report. The Self-Study reports include statements that describe the methods used by the program in achieving ABET Criterion 3.f.\textsuperscript{75} For each of the participating programs, documents associated with the prior three accreditation reviews were requested. Release of most of the ABET documents reviewed during this study was negotiated with representatives of the academic institution. In a few instances self-study reports for ABET accreditation reviews were publicly available and downloaded from academic websites.

The majority of the programs that agreed to release Self-Study documents provided the entire document. However, in some cases, programs extracted portions of the document prior to release. Each of the obtained Self-Study reports was reviewed for evidence of how a program addressed Criterion 3.f requirements. Pre-EC 2000 self-study reports were also evaluated based on the EC 2000 requirement of “understanding of professional and ethical responsibility.” Individual summary sheets were generated for each ABET document reviewed. See Barry\textsuperscript{13} for a copy of the ABET Document Review Summary Sheets.

MIDFIELD
The MIDFIELD database and related studies are supported by a grant from the National Science Foundation. The database contains records for all degree seeking, undergraduate students for
each of the member institutions. The longitudinal database includes records from fall 1987 through the summer 2005. The MIDFIELD database grew out of and has expanded upon the work of the Southeastern University and College Coalition for Engineering Education (SUCCEED). Collectively, the MIDFIELD institutions award 1/12 of all engineering degrees in the United States.

The study’s participating academic institutions are members of MIDFIELD. Data associated with scores on the FE Examination were not previously a part of the MIDFIELD database. However, the existing MIDFIELD agreement was augmented by obtaining written consent from each of the participating institutions to obtain that information from NCEES.

NCEES
NCEES authorized MIDFIELD to obtain FE Examination data for purposes of research in a letter dated 19 November 2004. A copy of that letter is on file in the MIDFIELD repository. That authorization letter was subsequently endorsed by NCEES in an email correspondence between the NCEES Director of Professional Services, Mr. E. Davis McDowell, P.E. and the MIDFIELD Project Director, Dr. Mathew W. Ohland, dated 29 January 2007. NCEES required that MIDFIELD obtain signed letters of authorization from each participating institution. Letters generated by each participating institution, granting access to the NCEES data are on file within the MIDFIELD repository.

An NCEES data file of student-level FE Examination scores for the participating academic institutions was initially released to MIDFIELD on 27 June 2008 and additional data were released on 7 August 2008. The data represents student-level scoring of individual examination topic areas for all administrations of the examination between October 1996 and April 2008 for the MIDFIELD institutions. That data was then evaluated relative to the research questions, as discussed in greater detail subsequently. This study represents the first published attempt to utilize NCEES data for the purpose of rigorous educational research.

Research Population
The research population consists of nine public academic institutions in the southeast United States, all of which are members of MIDFIELD. The population represents a purposeful sample of convenience. The established relationship and agreement in place between the academic institutions and MIDFIELD made for convenient and expedient access to an information-rich dataset. Further, the purposive aspect of the sample is associated with the intentional selection of particular engineering sub-disciplines. The participating institutions range in total enrollment (2004-05 academic year) from slightly more than 10,000 to nearly 48,000 students. Table 1 presents general information about each academic institution.

Several systems have been developed to classify, rank, and/or group academic institutions. The most widely used system is the *Carnegie Classification of Institutions of Higher Education* (Carnegie Classification). The Carnegie Foundation for the Advancement of Teaching initially developed their classification system in 1970 as a means of supporting its program of research and policy analysis. The Carnegie Classification has seen multiple updates during the past three decades. Perhaps the mostly widely used and recognized version of the Carnegie Classification is the 2000 update. That particular update groups academic institutions based predominately on
When NCEES generates Subject Matter Reports (formerly known as Report 5), they identify comparator groupings that utilize the 2000 version of the Carnegie Classification. Table 1 includes the MIDFIELD institutions’ Carnegie Classification utilizing both the 2000 and the 2005 versions of the classification system. Given that this study focuses on undergraduate engineering curricula, only the Undergraduate Instructional Program is reported from the 2005 multi-level classification system in Table 1.

Table 1. General Information for MIDFIELD Academic Institutions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Clemson University</td>
<td>Clemson, South Carolina</td>
<td>17,110</td>
<td>Carnegie 1 Research Extensive</td>
<td>Prof+A&amp;S/HGC</td>
</tr>
<tr>
<td>Florida A&amp;M University</td>
<td>Tallahassee, Florida</td>
<td>13,067</td>
<td>Carnegie 3 Master’s Colleges and Universities I</td>
<td>Prof+A&amp;S/SGC</td>
</tr>
<tr>
<td>Florida State University</td>
<td>Tallahassee, Florida</td>
<td>38,431</td>
<td>Carnegie 1 Research Extensive</td>
<td>Bal/HGC</td>
</tr>
<tr>
<td>Georgia Institute of Technology</td>
<td>Atlanta, Georgia</td>
<td>16,841</td>
<td>Carnegie 1 Research Extensive</td>
<td>Prof+A&amp;S/HGC</td>
</tr>
<tr>
<td>North Carolina A&amp;T University</td>
<td>Greensboro, North Carolina</td>
<td>10,383</td>
<td>Carnegie 3 Master’s Colleges and Universities I</td>
<td>Prof+A&amp;S/SGC</td>
</tr>
<tr>
<td>North Carolina State University</td>
<td>Raleigh, North Carolina</td>
<td>29,957</td>
<td>Carnegie 1 Research Extensive</td>
<td>Bal/HGC</td>
</tr>
<tr>
<td>University of Florida</td>
<td>Gainesville, Florida</td>
<td>47,993</td>
<td>Carnegie 1 Research Extensive</td>
<td>Prof+A&amp;S/HGC</td>
</tr>
<tr>
<td>University of North Carolina - Charlotte</td>
<td>Charlotte, North Carolina</td>
<td>19,846</td>
<td>Carnegie 3 Master’s Colleges and Universities I</td>
<td>Bal/HGC</td>
</tr>
<tr>
<td>Virginia Polytechnic Institute and State University</td>
<td>Blacksburg, Virginia</td>
<td>27,619</td>
<td>Carnegie 1 Research Extensive</td>
<td>Prof+A&amp;S/HGC</td>
</tr>
</tbody>
</table>

From Table 1, it can be observed that six of the MIDFIELD institutions are classified as Research Extensive (Carnegie 1) and the remaining three institutions are classified as Master’s Colleges and Universities I (Carnegie 3) using the 2000 system. Research Extensive institutions offer a wide range of baccalaureate programs and award a minimum of 50 doctoral degrees across 15 different disciplines, whereas, Master’s Colleges and Universities I offer a range of...
degree programs up to the master’s degree level. All of the MIDFIELD academic institutions offer baccalaureate degrees in civil, mechanical, and electrical engineering.

A unique aspect of two MIDFIELD institutions, not stated in Table 1, should be presented for clarification. Florida A&M University and Florida State University separately grant degrees in engineering, but they have a common college of engineering. The joint institution shares resources including, funding, faculty, and academic space. Thus, for the purpose of this study Florida A&M University and Florida State University were considered as a single academic institution.

**Instrumentation**

The instrument that was used in this study is the FE Examination as developed and administered by NCEES. There are only two nationally administered, engineering-specific examinations, the FE Examination and the Principles and Practice Examination. The FE Examination is the only examination whose content is focused on the intended knowledge gained in ABET-accredited engineering programs. The proposed construct of the FE Examination is competency to practice as an engineer.

The FE Examination consists of 180 multiple-choice questions and is scored using optical scanners. The examination is administered over a period of eight hours with one four-hour session in the morning and a second four-hour session in the afternoon. Content in the morning session comprises concepts common among all engineering disciplines. During the afternoon session, examinees have the option of taking an examination with general engineering content or an examination with content in a specific engineering discipline. A single reference handbook is supplied to each examinee, but no other reference material is permitted in the examination site.

Eligibility for taking the examination is determined by individual State Boards of Professional Registration (or equivalent organization). In the states were the MIDFIELD institutions reside, each state licensing board currently mandates that the students must be in their final year of studies in an ABET approved engineering program to register for the examination. Similar requirements are currently in place in other States.

The FE Examination is a criterion-referenced test. When a form of the examination is offered for the first time, or when significant changes in examination content are made, the passing score for the first administration is determined by a group of professionally licensed engineers acting as content experts. A statistical technique known as equating is used to preserve the passing standard on all subsequent administrations of the examination. In the event that a particular examination administration is more difficult than prior administrations, fewer questions must be answered correctly to achieve a passing score. NCEES clearly states that each examination is scored based on its own merits and without a predetermined target pass rate.

As noted in the work of Lawson, NCEES employs very rigorous administration procedures as a means of strengthening the examinations reliability. These practices include advanced assignment of examination entrance documents, photo identification for entry, limitations on examination aids such as calculators, strict prohibition of two-way communication devices (cell phones, pagers, etc.), consistent nationwide administration dates and times, distribution of
writing instruments, pronouncement of pre-test and post-test scripts, and test collection procedures.  

As noted previously, a copy of the *Exam Development Procedures Manual: Exam Development, Scoring, and General Procedures (Exam Manual)* and the *Security and Administration Procedures Manual* were provided by NCEES for review as part of this study. Due to the confidential nature of those documents, discussion of the specific details of examination development and administration is not permitted. However, it is appropriate to indicate here that NCEES adheres to testing preparation, administration, and security procedures such as the *Standards for Educational Psychological Testing*. 

Data Collection  
The primary dataset for this study was student-level, topic specific FE Examination scores that were obtained directly from NCEES. After each administration of the examination, NCEES generates topic-by-topic results that can be sorted by engineering discipline and by institution, and then can be compared to both nationwide averages and peer institution averages. A student-level dataset of FE Examination scores was obtained from NCEES. The database included a record of academic institution identifier, examination administration identifier, undergraduate major, and afternoon section of the examination selected, as well as topic-level scores for each student. The dataset also included an identifier for student name that consisted of the first four letters of the last name followed the first letter of the first name. Student names were used for the purpose of identifying potential repeat takers of the examination and the student identifiers were removed from the data set subsequently. 

Additional data necessary for performance of this study was related to documentation of professionalism and ethics curriculum incorporation methods used by the civil, electrical, and mechanical engineering programs at the MIDFIELD institutions during the chronological limits of the study. As noted previously, this data was collected via two separate methods; direct interviews with program representatives and review of ABET accreditation review documents. The intent of obtaining curriculum data via two disparate methods, when possible, was to triangulate wherein the multiple methods provide relatively independent verification of the findings. Reviewing the ABET documents was particularly useful in supplementing the memories of the interviewed program representatives. 

Interviews with program representatives were conducted between 6 March and 10 July 2008. Transcripts were prepared and reviewed for accuracy within two weeks of performance of individual interviews. ABET accreditation documents were collected and reviewed between 5 March and 20 August 2008. Interviews were consistently performed prior to review of associated ABET documents. The only exception to this condition was in the case of programs for which ABET documents were obtained, but no interview was conducted. 

Treatment of the Data  
This section of the research document provides an overview of the applied statistical analysis performed to address the research questions. This section has been organized to speak to the treatment of data related to each research question, separately.
Qualitative Data Treatment
The completed research study was designed in a manner that required minimal treatment of the collected data in support of the qualitative-oriented research questions. The interview protocol allowed for direct interpretation and confirmation of responses provided by program representatives. Likewise, the process of reviewing the ABET Self-Study documents permitted investigation of evidence in support of particular curriculum incorporation methods. A system of open coding and selective coding was used to extract theme-based evidence from the interview transcripts and ABET documents. An iterative process of comparison of findings was used during the course of qualitative data collection and evaluation. Extensive consideration was given to methods of presenting the qualitative data in a manner that addresses the associated research questions and preserves the integrity of data.

Quantitative Data Treatment
The database of FE Examination records required manipulation upon receipt to meet the needs of the current study. Data treatment included sorting and removal of data associated with examination administrations beyond the chronological limits of this study, as well as removal of student-level data associated with engineering disciplines other than the disciplines associated with this study (civil, electrical, and mechanical engineering). The manner by which these special conditions were addressed is detailed in the subsequent section of this document. After the database was filtered to meet the needs of this study, a total of 9,772 individual student-level records were used in this study.

The reported topic scores associated with each student were independently normalized relative to performance on the examination as a whole and peer statistics as described subsequently. The measure employed to minimize bias is based on the work of Wicker, Quintana, & Tarquin.\cite{Wicker99} Wicker et al. proposed an adjustment of program-level FE Examination scores based on national averages. The initial design of this study anticipated adjustments based on peer institutions as suggested by Nirmalakhandan, Daniel, & White.\cite{Nirmalakhandan97} It was expected that adjustments generated relative to a peer institution would result in a more accurate correction than would have been generated using national averages. Further, it was anticipated that peer institution subject matter results could be obtained directly from the Subject Matter Reports for each administration of the examination. However, it was discovered that both the national statistics and the peer institution statistics are generated for each combination of afternoon examination (e.g. general, civil, mechanical, etc.) and major (civil, environmental, construction, etc.). Therefore, the comparator statistics could not be readily obtained. Accordingly, the implemented method of adjusting for overall student score included the generation of topic-level peer statistics directly from the student-level database provided by NCEES. As all individuals take a common AM examination, peer-level data was generated for each administration of the examination for the full population of students, regardless of afternoon examination or undergraduate major. The application of the bias adjustment measure to student-level data, rather than program-level, represents another deviation from the work proposed by Wicker et al.\cite{Wicker99}
The process of score correction began with calculation of a relative score using the following equation:

\[ RS_{ijlms} = APC_{ijlms} - PAPC_{ij} \]  

(1)

Where:
- \( RS_{ijlms} \) = student-level relative score for a specific examination topic
- \( i \) = FE Examination topic (professional ethics)
- \( j \) = FE Examination administration (year and month)
- \( l \) = academic institution
- \( m \) = engineering discipline (civil, electrical, or mechanical)
- \( s \) = student identification value
- \( APC_{ijlms} \) = average percent correct for individual completing the examination (reported by NCEES)
- \( PAPC_{ij} \) = peer average percent correct (calculated from the NCEES data set)

Student bias can further be removed through calculation of a number that “provides a basis for comparison of performance in individual subjects relative to the other subjects” (p. 50). In essence, this adjustment accounts for general performance on the full examination. Wicker et al. referred to such a number as the “Tarquin” number. As this study builds upon the work of Wicker et al. by adapting their equation to utilize student-level data and peer comparisons, the used terminology within this study is Tarquin’ and is calculated as follows:

\[
TARQUIN'_{ijlms} = \left( APC_{ijlms} - PAPC_{ij} \right) - \frac{\sum_{k=1}^{n} (APC_{kjlms} - PAPC_{kj})}{n}
\]

(2)

Where:
- TARQUIN’\(_{ijlms}\) = student-level relative score adjusted for student bias based on peer institution results
- \( n \) = total number of FE Examination topic areas used in the model
- \( k \) = individual FE Examination topic other than \( i \)

This formula allows for correction of ethics scores relative to overall performance on the examination. The power of this formula is in the use of multiple additional examination topics (\( k \)) in the correction.

Tarquin’ data generated in this study aligns with discrete independent variables (e.g. 1 course, 2 courses, 3 courses, etc.). Therefore, using Microsoft Excel, a code was developed that randomized values associated with each independent variable. The result is a “jittering” effect (p. 121), wherein the individual data points are randomly plotted within an interval that is centralized around their associated independent variable, but the position of the dependent variable remains unaltered. However, the jittering effect impacts the subsequent statistical calculations. Therefore, calculations were performed on original, non-randomized data, and the jittering was utilized only for the purposes of data presentation.

As Wicker et al. note, an important assumption in the calculation of the Tarquin’ number is that performance on the individual topic areas of the FE Examination are normally distributed. During this study, the generated Tarquin’ values were checked for normality utilizing the
Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling tests. In addition, a histogram and q-q plot were generated to evaluate the normality graphically.

An analysis of variance (ANOVA) was performed to determine if a statistically significant difference existed between the mean value of the dependent variable (Tarquin') and several other variables. The variables under consideration were the academic institution, undergraduate major, administration time (pre- or post-curriculum changes), and the amount of required applicable content. The amount of required content was evaluated in terms of the number of required courses and the weighted percent of required course credits. The pre- and post- administration time was evaluated related to changes in the number of required courses and the weighted percent of required course credits, independently. With significant values generated by the ANOVA procedure, a Tukey pairwise comparison was performed to determine which sets of population means were different from one another. A Tukey test makes a comparison between all possible pairs of means and determines if the difference between any two means is greater than the expected standard error.\(^{57}\) Specific results of the ANOVA and Tukey procedures are reported in the Findings section of this document. All statistical analysis was performed using SAS 9.1.3.

It should be noted that the quantitative research question associated with this study seeks the relationship between the amount of applicable content within the engineering curriculum and student performance on the professional ethics section of the FE Examination. Accordingly, student performance was evaluated through the use of the Tarquin' values, which is a measure of individual student performance on the ethics section of the FE Examination adjusted for performance on the Examination as a whole, as well as peer Examination performance. Thus, the quantitative research question is addressed through consideration of relative performance. The assumption is made that the Tarquin' values are representative of actual student performance. The study by Wicker et al.\(^ {54}\) provides evidence in support of that assumption.

**Special Conditions**
Several special conditions were addressed during performance of the study. The manner by which each of these conditions was addressed is discussed in the subsequent sub-sections of this document.

**Multiple Examinations**
During the period studied, several states permitted students to take the examination earlier in their formal education and to take the exam multiple times. Within this time period, the licensing boards that determine examination registration eligibility for each state enacted consistent regulation as reflected today. As discussed previously, to be eligible to register for the examination, students in each state associated with this study, must currently be in the final semester of their accredited degree. The exact time when the shift in registration policy occurred was not verified. Rather, the data was filtered to remove any repeat takers of the examination. The procedure used for identifying and removing duplicates is a multi-step process. First the data set was sorted by “academic institution,” “administration,” and “name.” In this case, “name” is the 5-letter code described above. A formula was entered into the database that checks for and flags consecutive cells with duplicate “name.” Each flagged cell is manually inspected. If the apparent duplicate lines of data both reflect a “pass-pass” combination, the data is not duplicate.
An assumption has been made that individuals will not take the examination a second time if they pass on the first attempt. This assumption is considered reasonable when considering the time and expense associated with taking the examination multiple times. Further, individuals that complete the examination receive notification of pass or fail status without a reported numerical score. Thus, in the absence of numerical score reports, the final grade on the examination is not considered a matter of prestige within the engineering community. If the apparent duplicate lines of data reflect a “fail-pass” or “fail-fail” combination, the data is considered to be representative of an individual that took that examination more than once (duplicate). Consideration is also given to the level of commonality of the name (e.g. browj) and the size of the represented academic institution. When repeat examinees (duplicates) are identified, only the chronologically first line of data is retained.

Engineering Major
The pre-examination data collection process used by NCEES includes recording of the student’s self-identified academic major. Within the dataset, there were instances of students that self-identified with an engineering discipline or major that did not exist at a particular academic institution. For example, North Carolina State University’s and Virginia Polytechnic and State University’s civil engineering programs allows students to take a construction-oriented educational track. However, an ABET-accredited construction engineering degree is not offered at those academic institutions. This condition was also noted for environmental engineering at Clemson University, Florida A&M University-Florida State University, University of North Carolina – Charlotte, North Carolina A&T University, and North Carolina State University. For individual cases where this was identified the student-level data records were manual adjusted to reflect civil engineering.

Engineering Technology
In several states, engineering technology students are also eligible to register for the FE Examination. Many of the state-level regulations regarding engineering technology programs have changed during the chronological limits of the study. However, it was determined that the only MIDFIELD academic institution with engineering technology programs that directly parallel the engineering programs (e.g., civil engineering technology and civil engineering), that also resides in a state that does and/or did permit engineering technology students to sit for the examination, was University of North Carolina – Charlotte. Unfortunately, the NCEES pre-examination data record sheet does not include documentation of student classification as an engineering or engineering technology student separately. In this case, the student-level dataset was compared with records housed within the Dean’s office at University of North Carolina – Charlotte, to identify and filter engineering technology students from the database. Ultimately, the dataset, as provided by NCEES, contained relatively few engineering technology students for the University of North Carolina – Charlotte.

School Codes
A part of the pre-examination data collection process requires that students record a school code on the examination answer sheet. Each academic institution that graduates students eligible to take the examination has a unique school code. During several administrations of the FE Examination at Georgia Institute of Technology and one administration at Virginia Polytechnic and State University, the examination proctors failed to clearly identify the school code for
students to record. Accordingly, data for students from those particular examination administrations was not directly attributed to that academic institution. Thus, data for nine administrations at Georgia Institute of Technology and one administration at Virginia Polytechnic and State University were excluded from this study. All eighteen administrations of the examination, within the chronological limits of this study, were used for the remaining participating academic institutions.

**Effect Delay**

When the MIDFIELD programs made curriculum changes, those changes did not result in immediately measurable effects. The curriculum changes documented during this study include the addition and/or modification of courses in each year of undergraduate study. However, the effect was measured through administration of the FE Examination in the final year of study. Thus, curriculum changes made in the junior-year would require a minimum of one-year before the effect was measured, changes made in the sophomore year would require a minimum of two-years before the effect was measured, and finally, changes made in the freshman-year would require a minimum of three-years before the effect was measured.

For purposes of this study, the year of implementation noted in the supporting documents and illustrations are reported for the specific year that curriculum changes occurred. However, prior to performance of the statistical analysis, the dataset was manually edited to reflect the specific year that the number of required courses and/or the weighted percent of required course credits would have had an effect on the FE Examination.

**Other Special Conditions**

There are other conditions that could influence student performance on the FE Examination. For example, it is always possible that students may take an ethics-based free elective outside of the required engineering curriculum. The available dataset did not permit identification and adjustment of student-level scores for individuals that participated in such coursework. However, it is anticipated that the overall number of individuals that pursue such coursework is small.

Extracurricular activities, events, and experiences are also expected to influence the ethical and professional development of students. Sutkus, Carpenter, Finelli, and Harding\(^8^8\) detail an ongoing study to evaluate the influence of such activities, events, and experiences. The data collected during the study detailed in this document did not permit consideration of such influences.

Students enrolled in final semester of an ABET accredited graduate program are also eligible to take the FE Examination. Within the defined chronological limits of this study, the master of science in environmental engineering at Virginia Polytechnic and State University (initially accredited in 1984) was the only such program. Thus, it is possible that students enrolled in this program may be represented in the data set. The American Society for Engineering Education’s Engineering Data Management System indicates that the number of graduates from that program was 4 for 1998 (oldest available data).\(^8^9\) Accepting this value as the upper possible limit of representation in the study’s data set and recognizing that it is unlikely that full participation
would be achieved, it is comfortably acknowledged that a small number of such students could be included in the analyzed data set.

Finally, it is also recognized that discussions of ethics and professionalism may occur within required engineering courses that are not documented as part of either the interviews or ABET document reviews. Again, such an occurrence could not be accounted for, but is also not anticipated to affect this study.

Reliability and Validity

Appropriately, educational research, like most forms of research, is judged on its reliability and validity. For the purposes of this study, the interpretation of reliability and validity offered by Kirk & Miller\textsuperscript{90} has been adopted.

Reliability is the “degree to which the finding is independent of accidental circumstance of the research”(p. 20).\textsuperscript{90} Patton\textsuperscript{4} states that within a reality-oriented approach to inquiry the researcher will:

\begin{quote}
…establish an ‘audit trial’ to verify the rigor of your fieldwork and confirmability of the data collected because you want to minimize bias, maximize accuracy, and report impartially believing that ‘inaccuracy and bias are unacceptable in any case study. (p. 93)
\end{quote}

Miles & Huberman\textsuperscript{3} state that qualitative investigation and reporting should be neutral and free from bias. Reporting within this document has focused on firm description and analysis, rather than personal voice and viewpoint. The investigator’s personal motivation for conducting this study was based on discovery alone and was not influenced by preconceived notions related to curriculum methods. An iterative process of comparison of findings was used during the course of qualitative data collection to ensure consistency of methods. During the quantitative analysis statistical results were evaluated for reasonableness.

Within this document efforts have been made to be transparent in the methods employed and the evidence collected. An “audit trail” has been created that permits ease of replication by others (within the limits of required confidentiality). Extensive consideration was given to the design of the study and performance in adherence to strict levels of consistency. A process of revisiting interviews and ABET document summaries was instituted to ensure consistency throughout the data collection process.

Triangulation of data sources has been employed to strengthen the accuracy and credibility of the findings. Triangulation is accomplished within this study by looking for evidence within the ABET Self-Study documents that supports the statements made by department representatives during the semi-structured interviews. The ABET documentation and the statements made by departmental representatives are expected to be factual in nature. Others utilizing ABET Self-Study documents for education-based research have made similar assumptions.\textsuperscript{34, 91} Thus, there is little potential for accidental circumstances related to the findings of this study, which would impact the study’s reliability.

Validity is the “degree to which the finding is interpreted in a correct way”(p. 20).\textsuperscript{90} Within the limits of the required confidentiality, the \textit{Findings} section of this document presents the collected
data. The reader is free to interpret that data as one sees appropriate. However, the data is fairly unambiguous and is not likely to generate alternative interpretations.

A detailed description of the population has been provided. Accordingly, this permits the reader to make comparison to other populations. Limiting effects associated with the population have been identified, which controls the level of generalization. However, the conclusions reached from the findings have been presented in a sufficiently generic format to allow replication. This document also includes identification of settings where the study could be tested further.

The findings and conclusions are presented in a manner that makes them readily understood and accepted as plausible. Where appropriate, areas of uncertain data and related conclusions are clearly identified. In addition, rival explanations for the findings are discussed when suitable.

Patton\textsuperscript{4} says that the established criteria for quality of reality testing, a form of evaluation research, is “…plausibility of findings; credibility, impartiality, and independence of judgment; confirmability, consistency and dependability of data; and explainable inconsistencies…” (p. 93).

Kirk & Miller\textsuperscript{90} make the following statement when discussing reliability and validity: “No experiment can be perfectly controlled, and no measuring instrument can be perfectly calibrated” (p. 21). While their statement is inarguable, as detailed herein, this study has taken appropriate and pragmatic steps to ensure that the design, execution, and interpretation have been performed in a manner to maximize the reliability and validity.

**Generalization**

Even though a large dataset was used during this study, there are limitations to the extent that the findings and conclusions can be generalized. The MIDFIELD data represents a collection of academic institutions that is limited regionally to the southeastern United States. One out of every twelve recipients of accredited engineering degrees in the United States graduates from a MIDFIELD program. The study’s population included 23 programs at nine MIDFIELD academic institutions. ABET indicates that as of 2007 there were 1,798 accredited engineering programs at 368 colleges and universities.\textsuperscript{92} Thus, the population used in this study is not representative of all engineering students in the United States. Findings and conclusions drawn as part of this study should not be generalized beyond the MIDFIELD programs that participated in the study.

**Findings, Discussion, and Implications**

The research program detailed within this document has recently been completed. The findings have proven to be very rich with information which in turn has generated significant discussion. The implications of those findings and related discussion have particular bearing and implications for specific groups of engineering education stake-holders; including engineer administrators, ABET, and NCEES. Subsequent publication will detail the findings, conclusions, and implications of this study.
References


13 Barry, B.E., Engineering ethics curriculum incorporation methods and results from a nationally administered standardized examination, in School of Engineering Education. 2009, Purdue University: West Lafayette. p. 294.


85 American Educational Research Association., American Psychological Association., National Council on Measurement in Education., and Joint Committee on Standards for Educational and...


