

Validity of the Contextual Competence Scale for Engineering Students

Hyun Kyoung Ro,^a Dan Merson,^b
Lisa R. Lattuca,^c and Patrick T. Terenzini^b

^aCarnegie Mellon University, ^bPennsylvania State University,
^cUniversity of Michigan

Abstract

Background Engineering educators and practitioners increasingly value contextual competence. A psychometrically sound, practical, and useful scale for assessing undergraduate engineering students' contextual competence is needed.

Purpose/Hypothesis This article provides comprehensive evidence of the content, structural, discriminant, and criterion-related validity of the contextual competence scale.

Design/Method This study used student, alumni, and faculty survey data from a nationally representative sample of 120 U.S. engineering programs from 31 four-year institutions. Validity evidence was obtained by expert review of questions, a pilot test, factor analyses, and several analyses utilizing *t*-tests, correlations, and regression.

Results Experts constructed the questions used in the scale (content validity). Those questions combined to measure a single concept (structural validity), the scale reliably measures (Cronbach's alpha = .91) that concept, and it focuses on contextual competence instead of other student outcomes (discriminant validity). The contextual competence scale varies according to students' characteristics and curricular experiences as well as similarities and differences in student and alumni levels of contextual competence in the same programs and institutions (criterion-related validity). Finally, the scale may be a more accurate measure of contextual competence than faculty members' perceptions of students' ability.

Conclusions The contextual competence scale allows engineering programs to meet ABET and other self-study requirements or do large-scale research with relative ease and little expense. The process described in this article can be used by other researchers in engineering education for their scale development efforts.

Keywords contextual competence; measuring ABET outcomes; validity tests

Introduction

The importance of contextual competence in engineering design work is well understood and established within the profession (ABET Engineering Accreditation Commission, 2008; National Academy of Engineering [NAE], 2004, 2006). But a valid, parsimonious measure has not yet been developed to assess and enhance students' contextual competence in local self-assessment and larger-scale studies of undergraduate engineering education. This study applies

rigorous tests of validity to such a measure of contextual competence and describes a method researchers can use to assess the validity of other scales intended to measure student outcomes.

We developed a unidimensional, four-item scale that measures contextual competence as part of a national study of 120 engineering programs and 31 four-year institution (Ro, Lattuca, Merson, & Terenzini, 2012). Contextual competence is an engineer's ability to anticipate and understand the constraints and impacts of social, cultural, environmental, political, and other contexts on engineering solutions (Ro et al., 2012). Our definition of contextual competence was based on research in undergraduate professional fields, including engineering (Stark & Lowther, 1988). Ro et al. (2012) reported preliminary evidence of the validity of the scale using statistical procedures and professional judgment. The scale is conceptually clean, psychometrically sound, and internally consistent (Cronbach's $\alpha = .91$).

This article describes the development and the validation of the contextual competence (CC) scale and provides guidance for engineering educators and researchers on how to conduct validity tests on measures of engineering students' learning outcomes. Extending Ro et al.'s (2012) preliminary analyses, this article uses a broader literature review to further validate tests. Following the latest standards developed by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (1999), we focus on the scale's test content, internal structure, and relations to other variables. Drawing on the resulting validity evidence, we discuss the implications of the scale's use for practice in local self-assessment and large-scale studies of undergraduate engineering education.

Literature Review

Engineering educators and practitioners have increasingly come to value contextual competence. The ABET program accreditation criteria 3.c, 3.f, 3.h, and 3.j promote contextualization of engineering practice (ABET Engineering Accreditation Commission, 2008). The National Academy of Engineering also envisioned the workplace of the near future as one of dynamic technological change that requires engineers to understand complex societal, global, and professional contexts (NAE, 2004, 2006). The growing body of research literature also explores students' contextual understanding and ways to incorporate contextual competence into the engineering curriculum (Dym, Wesner, & Winner, 2003; Lau, 2004; National Research Council, 1995). Unfortunately, technical engineering courses seldom help students to develop non-technical knowledge, such as about contemporary regional, national, and global societal issues or ethics (Kastenberg, Hauser-Kastenberg, & Norris, 2006; Loui, 2006). Our study focuses on the measurement of contextual competence, which is widely acknowledged as critical but is underemphasized in undergraduate engineering programs.

Research Limitations

Despite widespread interest in the importance of engineering students' contextual competence, little research examines how to assess it using valid measures from literature-based definitions and validation studies. We summarize previous studies in terms of three limitations: lack of a comprehensive definition, methodological limitations, and lack of rigorous evidence of validity.

Because engineering design is central to engineering practice, a good deal of research focuses on how engineering students are taught to incorporate contextual understanding into their design practices. Kilgore, Atman, Yasuhara, Barker, and Morozov (2007) refer to the "breadth of problem-scoping" and the "context" of engineering as important to the consideration of

global and societal implications of engineering design (p. 321), which is a critical part of the design process. Their definition of contextual competence focuses on the breadth of problem-scoping behaviors when a student solves an engineering design task. Many of the judgments practicing engineers make extend beyond design, and thus the need for contextual competence extends beyond the design task itself (Sheppard, Macatangay, Colby, & Sullivan, 2008). Working with educators in eight undergraduate professional fields, including engineering, Stark and Lowther (1988) defined contextual competence as the “understanding of the societal context (environment) in which the profession is practiced” in order to prepare a professional “to make judgments in light of historical, social, economic, scientific, and political realities” (p. 23).

While contextual competence has many aspects (Stark & Lowther, 1988), most studies focus on only one of these dimensions and typically define contextual competence in terms of cultural competence, which includes the cultural and social aspects of engineering practice (Besterfield-Sacre, Atman, & Shuman, 1998; Mikic & Grasso, 2002). Common emphases are global competence, including cross-cultural sensitivity, adaptability, and communication skills (Shen, Jesiek, & Chang, 2011), as well as global teamwork skills and working with people who define problems differently from oneself (Bielefeldt, 2008; Downey et al., 2006). Bordogna, Fromm, and Ernst (1993), for example, defined contextual understanding capability as the appreciation of the economic, industrial, and international environment in which engineering is practiced and the ability to provide societal leadership effectively. These studies address aspects of contextual competence, but they rarely include other dimensions of contexts.

In addition to lacking a broad definition of contextual competence, previous studies had several methodological limitations. Most of the authors developed their instruments using a single institution or a small number of colleges, so the instruments might be useful only in similar types of institutions with similar kinds of students. Also, these studies did not use multiple programs to validate the instrument. Different engineering disciplines have different curricular emphases, policies, practices, and faculty cultures that influence their approach to increasing their students’ competencies (Lattuca, Terenzini, Harper, & Yin, 2010), and thus single-discipline analyses might not be generalizable to other disciplines. Further, these studies focus on either course-level assessment, a single-institution program evaluation, or self-study for accreditation rather than evaluating students across many institutions.

Few studies that have developed contextual competence measures provide comprehensive evidence on the validity of the measures. For example, Strauss and Terenzini (2005) provided a psychometric study of how to develop nine factors representing learning outcomes identified by ABET, including a factor called Societal and Global Issues. Although they provided the results of a factor analysis and the scale’s internal consistency reliability (Cronbach’s alpha), they completed no further research specifically providing other validity evidence of that measure.

Measure Validity

According to the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (1999), “Validity refers to the degree to which evidence and theory support the interpretation of test scores entailed by proposed uses of tests” (p. 9). Their joint standards state that the examination of a measure’s validity relies on the collection of evidence that is based on test content, response processes, consequences of testing, internal structure, and relations to other variables.

Evidence based on test content refers to the extent to which a scale’s items, in the aggregate, constitute a representative sample of the topic’s content domain. Do the items reflect

what has been defined as “contextual competence” (Suen, 2008; Trochim, 2006)? To answer the question, content experts are consulted and their professional judgment is taken to reflect the degree of what was traditionally called “content validity.” Evidence of response processes pertains to the method respondents use to complete the measurement. The process used should authentically simulate the processes required by the theoretical construct being measured (Suen, 2008). Evidence about the consequences of testing relates to the potential for intended or, more often, unintended consequences of the use of test scores to disadvantage one group compared with another if the consequences are a result of the instrument’s inability to accurately measure the construct in question for the intended purpose.

Evidence based on the internal structure indicates the extent to which a set of items or a measure reflects the underlying construct (Suen, 2008). Evaluation of the internal structure typically entails using factor analytic techniques to determine whether the items, in the aggregate, constitute a common, underlying factor (the construct). This approach has traditionally been called “structural validity” because it illustrates the statistical structure of the instrument’s items and constructs. In addition, factor analysis can provide evidence on whether a factor represents the construct it is intended to measure and does not represent others that are theoretically different (Leighton, 2008; Trochim, 2006). This concept was previously called “discriminant validity” and refers to the measure’s ability to discriminate between similar constructs.

Evidence based on relations to other variables pertains to whether the measure’s score correlates with scores on other variables; such a correlation would be expected based on predictions of the theory underlying the construct (Suen, 2008; Trochim, 2006). The goal is to have a scale that can identify differences between individuals or groups that might be reasonably expected to differ on the trait reflected in the measure. This concept was previously referred to as either “concurrent” or “predictive validity,” depending on whether the different variables were measured concurrently or if the construct in question was believed to have a causal effect on the other variable(s). In order to conduct analyses that will provide evidence pertaining to a construct’s relation to other variables, one must first determine the nature of the expected relationships.

Relationships between contextual competence and other variables We reviewed considerable literature and determined that our contextual competence (CC) scale relates to students’ demographics, curricular experiences (academic major and design course enrollment), and other measures reported by alumni and faculty members. We examine these expected relationships in the following sections of this literature review.

Relationship with students’ demographics Engineering students’ contextual competence is related to demographics, such as gender and class year. Women place a larger importance on conceptual relationships and connections than do men and focus on larger, broader pictures rather than individual details (Häussler & Hoffmann, 2002; Kilgore et al., 2007). Like the gender gap in general self-confidence (Brainard & Carlin, 1998), however, women’s self-rated ability to consider contexts may be lower than men’s. Women in science, technology, engineering, and mathematics (STEM) fields tend to lose confidence upon matriculation, despite entering college with similar achievement and confidence levels (Seymour, 1995; Whitt, Pascarella, Elkins Neisheim, Marth, & Pierson, 2003). First-year students and seniors have different understandings of design context because the seniors typically have more design and work experience through capstone design courses and a variety of co-curricular activities (Atman, Cardella, Turns, & Adams, 2005; Bogusch, Turns, & Atman, 2000; Kilgore et al., 2007).

Relationship with students' curricular experiences Engineering students' contextual competence also varies according to their academic major and curriculum. The literature suggests that academic programs' curricular emphases and faculty members' engagement in teaching, which influence students' experiences and learning (Pascarella & Terenzini, 2005), vary across engineering disciplines. These differences are consistent with the patterns suggested by Holland's typology-based disciplinary environments (Holland, 1973; Lattuca et al., 2010). Faculty members in what Holland termed "enterprising" (industrial engineering) and "investigative" (chemical and civil engineering) programs reported greater emphasis than those in "realistic" programs (electrical and mechanical engineering) on professional and ethical responsibilities and on the societal and global implications of engineering solutions in their courses.

Furthermore, working through design tasks is one avenue to enhancing a student's contextual competence. In a case study that analyzed six institutions with exemplary engineering programs, Lattuca et al. (2010) found that first-year design and capstone design courses were the most likely curricular venues for developing students' contextual competence. In a companion study that surveyed more than 5,000 engineering undergraduate students, Palmer et al. (2011) examined the influence on contextual competence of four types of curricular emphases: professional skills, such as leadership, teamwork, communication, and project management; professional values, such as ethics, lifelong learning, and diversity; core engineering thinking, such as problem-solving skills and creativity; and broad and systems perspectives (i.e., systems thinking, such as understanding how nonengineering fields can help solve an engineering problem). They found that curricula that emphasized broad and systems perspectives increased contextual competence the most. In addition to the inclusion of broad societal context components into engineering curricula, Kilgore et al. (2007) found that first-year engineering students were receptive to this approach to learning; many of the students in their study were sophisticated enough to incorporate both technical problems and solutions that were based on societal issues.

Relationship with other measures of contextual competence Despite the engineering education profession's increasing attention to contextual competence, the evidence suggests that engineering students generally lack key aspects of this competence (Kranov, Hauser, Olsen, & Girardeau, 2008). Lattuca, Terenzini, and Volkwein (2006) found that, among a variety of learning outcomes ABET emphasizes, engineering seniors were least confident about their ability to understand the broader societal contexts that influenced engineering practice. Employers corroborated these findings. Forty-eight percent of the 1,622 employers surveyed as part of their study judged new hires as inadequately prepared to understand how a variety of contexts shape, and might be shaped by, engineering solutions (Lattuca et al., 2006). Professional engineers expect graduating students to possess adequate levels of contextual competence. We predict that current students and recent graduates from the same program should report similar levels of senior-year competence when compared to other programs or institutions. Alumni experience with solving real-world problems, however, should lead them to score higher on contextual competence than current students.

Faculty members' evaluations are a common source of realistic and objective rating of their students' abilities. Despite faculty's expertise, the reliability and validity of their assessments are open to question, just as are students' self-ratings (Ward, Gruppen, & Regehr, 2002). Previous comprehensive summaries of findings in the self-assessment literature in a variety of disciplines, including engineering, suggested that expert ratings and student self-assessments are not highly correlated (Falchikov & Boud, 1989).

Methods

The contextual competency (CC) scale was developed as part of a study, Prototype to Production: Conditions and Processes for Educating the Engineer of 2020 (P2P), funded by the National Science Foundation (NSF Grant EEC-0550608). The overall goal of the P2P study was to examine the curricular, instructional, cultural, and organizational features that support high-quality learning in engineering programs. To provide a comprehensive perspective of undergraduate engineering education, the study design consolidated data from several constituent groups: engineering faculty members, program chairs, and associate deans for undergraduate education, as well as engineering undergraduates and alumni. The CC scale was developed for use in the engineering student and alumni surveys. Faculty members were also asked a single question about their senior students' contextual competence.

Survey Design

A team of education and engineering researchers collaborated on the development of five surveys for engineering students, alumni, faculty, program chairs, and administrators during a rigorous, two-year process. Figure 1 shows how the research team developed the survey instruments. The team first conducted an extensive literature review on key topics related to contextual competence in engineering, but also in fields outside engineering. In addition to reviewing this literature and developing a bank of relevant items from existing instruments and studies, the team spent a year conducting interviews and focus groups with engineering administrators, faculty members, students, and alumni on five campuses (two universities, a branch of one university, and two community colleges) to understand how engineering programs sought to develop students' contextual competence through the curriculum and co-curriculum.

Interview and focus group transcripts were distributed to the research team. The research team included faculty members who had extensive experience with quantitative social science studies of college student experiences and outcomes; an engineering faculty member; and doctoral students who had worked in colleges of engineering, had previous engineering experience in both college and industry, or were graduate engineering students. The team met weekly for almost a year to discuss the findings and to refine and draft items for the five

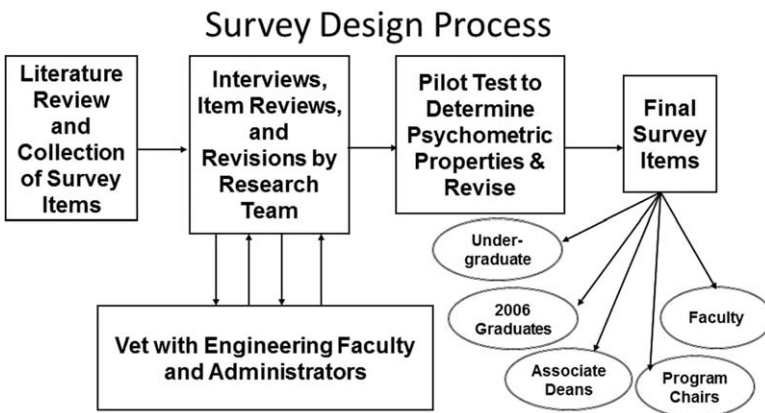


Figure 1 The survey instruments design process.

surveys. As part of this process, the team developed sets of items for use in examining specific student outcomes of the study, one of which was contextual competence. Other outcomes included fundamental skills, design skills, interdisciplinary skills, communication skills, teamwork skills, and leadership skills.

Once drafted, the surveys were reviewed by engineering faculty and administrators at the Pennsylvania State University, who met in focus groups with the members of the team to revise and refine the individual items. The four-year student, two-year student, and faculty instruments were then pilot tested as described in the Internal Structure section of the Findings section below. After the pilot test, the research team again met with focus groups of engineering faculty members and administrators from Penn State to review the final student survey and to incorporate feedback from the project's National Advisory Board, before administering the surveys to the full Prototype to Production sample.

Sample and Data Treatment

The Prototype to Production (P2P) study utilized a cross-sectional survey design to assess engineering education at 31 four-year U.S. institutions. A disproportional, stratified random sampling plan was used to produce a nationally representative sample of four-year baccalaureate engineering programs that offer two or more ABET-accredited programs in six engineering disciplines: bio/biomedical, chemical, civil, electrical, industrial, and mechanical engineering. Because the P2P study was also designed to inform analyses of a closely related set of six case studies, one of which offered only a baccalaureate-level general engineering program, the sample was refined to include three institutions that offered only a general engineering program or a general program in addition to their discipline-based programs. All faculty members, program chairs, and sophomore, junior, and senior (second-, third-, and fourth-year) students at the 31 institutions were invited to participate in web-based surveys. The student survey included students' self-assessments of selected learning outcomes, including contextual competence.

Table 1 summarizes the demographics of the respondents. The student survey was administered to 32,737 engineering students, 5,249 (16%) of whom responded. As well, 1,403 alumni (19% of the 7,307 invited) participated. Such relatively low response rates for these surveys are not uncommon by current standards. Survey response rates have been declining for several decades (Baruch, 1999; Dey, 1997; Smith, 1995; Steeh, 1981), and web-based surveys often have low response rates (Sax, Gilmartin, & Bryant, 2003; Zhang, 2000). After cleaning the data to eliminate invalid entries and prepare variables for analysis, the final sample consisted of 5,115 students and 1,283 alumni. Also included in the analysis is a single measure of faculty members' perceptions of their senior students' contextual competence and a question from the program-chair survey asking whether their program had a first-year design course requirement.

To adjust for possible response bias due to self-selection, responses were weighted at the campus level according to the following characteristics: student data with respect to gender, race/ethnicity, class year, and engineering discipline; and alumni data with respect to gender, race/ethnicity, and engineering discipline. However, weights were not applied for the factor analysis because doing so would have created scale variables that would become double weighted when used in subsequent analyses where weights were applied. Data that were missing from student, alumni, faculty, and program chairs due to individual question nonresponse were imputed using the expectation-maximization method recommended by Dempster, Laird, and Rubin (1977) and Graham (2009).

Table 1 Student and Alumni Demographics

Characteristic	2008 Students ^a		2006 Alumni ^a	
	N (%)	Respondents (%)	N (%)	Respondents (%)
Discipline				
Bio/Biomedical	6.5	8.7	5.7	6.3
Chemical	10.4	14.4	8.5	9.1
Civil	19.5	17.3	17.1	14.8
Electrical	21.8	17.5	28.0	32.1
Industrial	6.1	4.3	7.2	8.1
Mechanical	32.1	29.1	31.2	24.3
General	3.6	8.1	2.3	5.3
Gender				
Men	81.5	71.9	79.9	79.3
Women	18.5	28.1	20.1	20.7
Race/Ethnicity				
African American	5.2	2.8	4.7	5.3
Asian or Pacific Islander	12.1	8.1	12.7	15.6
Hispanic	6.5	5.8	6.7	7.4
American Indian	0.6	0.2	0.5	0.1
Other	6.1	5.9	7.1	3.8
Foreign national	5.9	12.5	6.9	6.6
Caucasian	63.5	64.8	61.3	61.2
Class standing^b				
Sophomore	6.1	17.9		
Junior	39.0	33.9		
Senior	54.9	48.2		

^aWeighted and imputed. ^bSelf-reported class year, not number of earned credits or hours (e.g., sophomores consider themselves to be second-year students).

Validity Testing

As described in the literature review, five types of evidence can be used to validate a scale. For this article we focused on test content, internal structure, and relations to other variables. We did not address evidence on response process or consequences of testing for several reasons. First, the P2P research project did not intend to develop a measure of contextual competence that could be used for purposes of placement or individual performance assessment. Development of such a measure would have been prohibitively expensive and time consuming. Our goal, rather, was to develop a scale of contextual competence that could measure group performance for use in local and scholarly research and outcomes assessment (e.g., for ABET accreditation review). Second, for such research and assessment purposes, individual students were not evaluated on their contextual competence for either formative or summative purposes, so there are no individual-student consequences of the measurement to consider. Finally, since our study was cross-sectional, it did not measure preconditions to address selection bias or to follow up with the individual students. We address this and other limitations in the Limitations section below.

The methods employed in this study to examine the validity of the CC scale are organized based on the type of validity evidence collected as outlined in the literature review: test

content, internal structure, and relations to other variables. All of the evidence presented in this article relates to the final datasets unless the pilot tests are mentioned.

Test content As described in the Survey Design section above, content experts were consulted and their professional judgment was taken to reflect the degree to which contextual competence was adequately operationalized in the scale.

Internal structure To examine the internal structure of the CC scale, the research team applied principal components analysis to the pilot-test data and factor analytic techniques to the full four-year student dataset. These procedures are detailed in Ro et al. (2012) and summarized below.

In particular, the factor analysis of the full dataset included all of the student outcome items on the four-year student survey. We intended to measure multiple skill sets by means of the student outcome scales, not just contextual competence. A factor that includes strongly loaded items and does not highly correlate with other factors can clearly measure its intended construct while distinguishing it from the other constructs being measured in the survey.

Relations to other variables First, we conducted a linear, ordinary least-squares regression analysis of contextual competence, based on gender, class standing, citizenship, and academic major. We also added a curricular variable: whether the students were required to take a first-year design course. Drawing on the literature, we hypothesized that students who take a first-year design course are more likely to develop contextual competence. We tested that hypothesis with a *t*-test instead of regression because we had only 2,009 student cases from the matched student and program data (the program-chair data contained 1,197 students who are in programs that require first-year design courses and 812 who were not). Including this variable in the regression analysis would have greatly reduced the number of cases and power of the analysis, so we first examined the demographic variables to understand their relationship with contextual competence and then examined the influence of the first-year design course separately.

We also examined the correlations between students' contextual competence and alumni respondents' measures of their contextual competence when they were college seniors. We calculated student and alumni mean scores by program within the same institution (e.g., a mean score of mechanical engineering students and a mean score of mechanical engineering alumni from Penn State). We then compared student and alumni respondents in the same program on each campus. We also used paired *t*-tests to determine if group mean scores differed between student and alumni respondents' current levels of contextual competence to test the hypothesis that professional engineers would score higher because they would have more experience solving real-world problems requiring the consideration of relevant context. Both analyses are organized by engineering program within specific institutions.

Finally, we examined students' self-ratings from the four-item CC scale compared with their faculty members' assessment of them based on a single question on the faculty survey. We also included single-question faculty assessments of other student outcomes. Responses were again clustered within engineering program and institution.

Limitations

Although the scale items covered multiple aspects of contextual competence, each subdimension was not measured by multiple items. Parsimony, balanced with measurement detail, is a common goal of scale creation (DeVellis, 2003). The parsimoniousness of the CC scale is advantageous, but also precludes its use for purposes that require detailed measurement results, such as academic placement or individual performance assessment. The scale was

developed for research and assessment purposes; because of the complexity of the larger study of which it was but a part, the need to optimize response rates imposed strict limitations on the number of items that measured any given student outcome. Consequently, the complexity at which any outcome could be explored was limited. Thus, the CC scale must be understood as a global measure of what is probably a substantially more complex engineering skill. The scale is probably best used in situations that require parsimony, whether for space requirements, as in a study of multiple outcomes, or time, as in a quick assessment of a particular class unit, course, or program intended to promote contextual competence.

The CC scale in this study was examined through respondents' self-reports. Higher education researchers and administrators have frequently used self-reported assessment as indicators of student learning or ability, but the literature disagrees on their accuracy. Bowman (2010) reported that some researchers found a strong correlation between subjective and objective assessments, while others reported a strong divergence. Although a direct measure of contextual competence might be preferable (e.g., performance on some task, product, or other factual demonstration of their contextual competence), the logistics, costs, and time associated with gathering such evidence from the many respondents at 31 institutions precluded including such a step in the research design. Moreover, there is no widely used standardized test of contextual competence that can be applied in any type of engineering practice or design setting. It is also worth noting that standardized measures have their own limitations for survey research or assessments (e.g., their greater length, cost, and administration requirements).

The format of the survey itself may have been psychometrically problematic. The survey was administered online and was presented as a series of webpages. When respondents felt they had completed a page of questions, they were asked to click "Next" to move on. Question items were organized by topic or learning outcome with a common question stem and response scale. Thus, the contextual competence questions appeared together on the same page. The factor analysis was possibly influenced by this visual and organizational structure. However, the results of the full factor analysis of all of the learning outcome questions from the four-year student survey shows that some questions, contrary to expectations, loaded on other learning outcome factors, even on outcomes that were on other pages of the survey. Administering the survey with a different or randomized organization may isolate any formatting effects to provide a more precise psychometric analysis, but it was impractical to do so for such a large, comprehensive study.

This study used only one method to examine the scale's reliability, calculating Cronbach's alpha (Cronbach, 1951; de Gruijter & van der Kamp, 2008). Analysis based on a single administration freezes certain types of error at the time of administration and thus does not include them in the estimate of instrument error variance (Haertel, 2006). Parallel forms of reliability evidence can be calculated by administering separate versions that are psychometrically equivalent in key aspects (de Gruijter & van der Kamp, 2008; Haertel, 2006). We are unable to provide this type of evidence because the research team developed and tested a single version of the scale. Test-retest reliability is also frequently calculated to assess the extent to which the trait being measured is stable over situation and time (de Gruijter & van der Kamp, 2008; Haertel, 2006). The cross-sectional nature of the study's design precluded any assessment of the scale's test-retest reliability. We recommend that future research utilizing this scale should provide additional information on the psychometric properties of the scale.

Findings

Test Content

Professionals with expertise in social science research, higher education, engineering education, and engineering practice crafted and selected survey items that, in their judgment, reflected the meaning and content of “contextual competence” as defined for this scale. Other engineering faculty members and members of the project’s National Advisory Board agreed with the project team members and focus group participants that the measure’s content validity was adequate for its intended purposes.

Internal Structure

The Prototype to Production research team used principal components analysis with the pilot-study sample and principal axis factoring with the full national sample to identify the conceptual structure of each outcome scale, including contextual competence. A confirmatory factor analysis using traditional factor analytic methods (Garson, 2010) showed that all four items loaded onto a single factor, the resulting scale had high internal consistency reliability (Cronbach’s alpha = .91), and the reliability of the scale would decrease if any item was removed (Table 2).

The factor analysis of the full dataset included all of the student outcome items on the four-year student survey. The resulting factor structure discriminated between contextual competence and other areas of engineering student competence that factored into eight discrete scales that may be related to, but are distinctly different from, contextual competence; these scales are fundamental skills, design skills, leadership skills, teamwork skills, communication skills, interdisciplinary skills, recognizing disciplinary perspectives, and reflective behavior practice.

The P2P team recognized that the outcomes may be theoretically, as well as empirically, related to each other, so they chose an oblique rotation in the factor analysis, which allowed for factors that correlate. The highest correlation associated with contextual competence was with design skills (.65); all other correlations were below .60 (Table 3). The moderately strong correlation between contextual competence and design is consistent with the view that contextual competence is an important element of many successful design solutions. Social science researchers

Table 2 Contextual Competence Scale

Please rate your	Factor loading	<i>M</i>	<i>SD</i>	Alpha if deleted
Knowledge of contexts (social, political, economic, cultural, environmental, ethical, etc.) that might affect the solution to an engineering problem	.83	3.35	1.03	.88
Knowledge of the connections between technological solutions and their implications for the society or groups they are intended to benefit	.85	3.35	1.02	.87
Ability to use what you know about different cultures, social values, or political systems in developing engineering solutions	.86	3.24	1.13	.88
Ability to recognize how different contexts can change a solution	.78	3.54	0.99	.89

Note. Response scale: 1 = weak/none, 2 = fair, 3 = good, 4 = very good, 5 = excellent. Alpha = .91.

typically accept even small-to-moderate correlations as evidence of relationships between constructs, particularly when the underlying constructs of the factors can be theoretically expected to correlate, so our results are not entirely obvious. One would expect several of the student outcomes to be correlated with each other because they represent a set of skills that are theoretically and functionally similar, such as leadership, teamwork, and communication skills. We did find medium-to-large correlations between the CC scale and other outcomes, and the level of correlation one considers to be small or large depends on one's behavioral science discipline (Cohen, 1988). There is no generally accepted cut-off value for the purposes of determining discriminant validity, although .85 has been given by Garson (2010). The result that neither the contextual competence factor nor its individual items correlated very highly with any of the other constructs provides some evidence of its ability to discriminate between contextual competence and other, related engineering-skill constructs.

Relations to Other Variables

To develop a scale that can identify differences among groups that might differ for that trait, we conducted a series of analyses to evaluate the CC scale's ability to make such differentiations. We examined student demographics, first-year design, and alumni contextual competence.

Student demographics To assess the relations identified in our literature review, we conducted a linear regression analysis of contextual competence with gender, class-year standing, and academic major. This technique allowed us to determine the relationship between level of contextual competence and each of the selected demographic variables while simultaneously controlling for the influence of the other variables, such as race/ethnicity and achievement (measured by combined SAT scores; Table 4). The full linear regression model was significant ($F(11, 5069) = 17.577, p < .001$). Graphs of the residuals indicate that the model is appropriately specified and that residuals are not related to the other variables in the model.

The literature review indicated that women tend to rate their own abilities lower than do men, regardless of their actual ability (Seymour, 1995; Whitt et al., 2003). As shown in Table 4, men in our study reported higher levels of contextual competence than did women. The finding is consistent with other analyses conducted by our research team (McKenna, Plumb, Kremer, Yin, & Ro, 2011). As might be expected, junior- and senior-year students also scored higher in contextual competence than did sophomores.

Finally, we found no significant differences in contextual competence among most of the engineering programs except for general engineering and industrial engineering, whose students reported higher levels of contextual competence than did mechanical engineering students. These findings are consistent with previous research that indicated that programs (such as industrial engineering) classified in Holland's model of occupational themes (Holland, 1973) as "enterprising" emphasize the societal and global implications of engineering solutions more than do "realistic" programs (Lattuca, et al., 2010). Our findings related to gender, class year, and academic majors are consistent with both theory and previous research and provide evidence that supports the scale's validity.

Table 3 Contextual Competence Scale and Student Outcomes

Outcome variables	Correlation
Design skills	.65
Leadership skills	.58
Communication skills	.55
Teamwork skills	.54
Interdisciplinary skills	.47
Recognizing disciplinary perspectives	.43
Fundamental skills	.35
Reflective behavior practice	.21

First-year design course requirement By conducting a *t*-test, we also examined differences in contextual competence between students who took required first-year design courses and those who did not. Students who were required to take a first-year design course had significantly higher contextual competence scores (mean = 3.37) than those who were not (mean = 3.23, $t = -3.32$, $p < .001$; pooled standard deviation = .92; significant Levene's test indicates unequal group variances). The Cohen's *d* effect size of .15 (reflecting the magnitude of the group difference after adjusting for different group sizes and standard deviations), however, indicates the first-year design course may have little or no importance in practice (Cohen, 1988).

Student and alumni contextual competence

We explored the correlation of senior students' assessment of current ability and the alumni's responses to questions that asked them to report their level of contextual competence when they were seniors in college (Table 5). We also compared the overall factor scale scores. Table 5 indicates the statistically significant correlations between the two groups on both the individual items and the overall scale score; these correlations provide additional evidence of the scale's validity.

We also hypothesized that alumni's item scores would be higher than those of students because professional engineers might have more experience solving real-world problems that require consideration of relevant context. After applying a Bonferroni-Holm correction to adjust for the accumulation of Type I error, the results for the overall scale scores and one of the four individual items support this hypothesis (Table 6).

Student and faculty measures of contextual competence Finally, we compared senior students' estimates of their contextual competence using the four-item CC scale to their faculty members' assessment of their students' contextual competence, which is derived from a single-item question. The two measures of contextual competence are not correlated, although faculty members' assessment of their students' learning outcomes are somewhat to highly correlated with each other (Table 7). At least three reasons may explain these results. First, the faculty and student scores were matched by program, but faculty were asked about their students in general, so there was no way to match individual students and their faculty members' assessment of their learning outcomes. Second, the student's CC scale consists of multiple items, but the faculty measure is a single question: "Now, think about current seniors in your program. Rate their abilities in the following areas." It is generally accepted that well-constructed scales are superior measures of a latent construct than single items (de Gruijter & van der Kamp, 2008; DeVellis, 2003). These findings may support that view. Finally, all of the faculty measures of their students' learning outcomes are correlated with each other, ranging from .29 to .77 (Table 7). This result may suggest a "halo effect" (Thorndike, 1920): faculty who think their students perform well (or poorly) in one area may think they perform well (or poorly) in others areas.

Table 4 Contextual Competence and Student Characteristics

Variable	β
Woman	-0.067*
Juniors ^a	0.121***
Seniors ^a	0.302***
Biomedical/bioengineering ^b	0.048
Chemical engineering ^b	-0.048
Civil engineering ^b	0.066
Electrical engineering ^b	-0.045
General engineering ^b	0.133*
Industrial engineering ^b	0.308***
Black, Latino, Native American ^c	0.038
Asian American ^c	-0.157***
Race other ^c	0.048
SAT composite score ^d	0.004***
Adjusted R^2	0.042***

Note. $n = 5,115$. β = the standardized regression coefficient. ^aReference group is sophomore students. ^bReference group is mechanical engineering. ^cReference group is Caucasian/White. ^dSAT composite score adjusted from one- to ten-point intervals. * $p < .05$; *** $p < .001$.

Table 5 Seniors' Contextual Competence and Alumni's Ability When Seniors

Current seniors' responses of their ability	Alumni's responses of their ability senior year				
	1	2	3	4	Scale
1. Knowledge of contexts (social, political, economic, cultural, environmental, ethical, etc.) that might affect the solution to an engineering problem	.18*				
2. Knowledge of the connections between technological solutions and their implications for the society or groups they are intended to benefit		.23*			
3. Ability to use what you know about different cultures, social values, or political systems in developing engineering solutions			.27**		
4. Ability to recognize how different contexts can change a solution				.24**	
Contextual competence scale					.28**

* $p < .05$; ** $p < .01$.

Table 6 *T*-tests of Student and Alumni Levels of Contextual Competence

Contextual competence items	Senior	Alumni	Mean difference	Effect size <i>d</i>
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)		
Knowledge of contexts (social, political, economic, cultural, environmental, ethical, etc.) that might affect the solution to an engineering problem	3.46 (0.37)	3.73 (0.47)	0.27***	0.57
Knowledge of the connections between technological solutions and their implications for the society or groups they are intended to benefit	3.46 (0.34)	3.57 (0.48)	0.11	0.23
Ability to use what you know about different cultures, social values, or political systems in developing engineering solutions	3.36 (0.45)	3.41 (0.49)	0.05	0.10
Ability to recognize how different contexts can change a solution	3.63 (0.40)	3.73 (0.46)	0.10	0.22
Contextual competence scale	3.48 (0.34)	3.61 (0.39)	0.13**	0.33

** $p < .01$; *** $p < .001$.

Discussion and Implications

Stark, Lowther, and Hagerty (1987) were among the first to propose contextual competence as an important learning outcome for professional education, including engineering. A decade later, ABET reinforced the importance of contextual competence by including it as one of 11 outcomes criteria to be reviewed for reaccreditation (ABET, 1997). More recently, the need to produce engineering graduates with contextual competence was stated in national

Table 7 Correlation of Student Learning Outcomes as Assessed by Faculty

Student learning outcomes	1	2	3	4	5	6	7
1. Contextual competence	–	.60	.67	.62	.55	.55	.52
2. Math, science, & eng. fundamentals		–	.58	.29	.33	.39	.32
3. Design skills			–	.49	.53	.48	.47
4. Interdisciplinary competence				–	.55	.48	.55
5. Communication skills					–	.71	.71
6. Teamwork skills						–	.77
7. Leadership skills							–

Note. All faculty assessments are correlated with each other at $p < .001$.

reports to the engineering community. With a few exceptions (Atman, 2009; Atman, Kilgore, Yasuhara, & Morozov, 2008), however, research on students' development of contextual competence is scarce, and sound, parsimonious instruments that assess students' contextual competence beyond the design context are rare.

We have shown the CC scale to be a useful, practical measure of contextual competence that is psychometrically rigorous and practical (Muchinsky, 2004). The scale allows engineering programs to meet ABET and other self-study requirements with relative ease and little expense. The items can be completed by students in small and large groups easily and quickly. Because the scale permits both sound and parsimonious assessment of students' overall contextual competence, it leaves room in survey instruments to assess other areas of interest to engineering programs without overburdening students and reducing survey response rates.

The practical advantages of adopting the scale are wide and varied. Engineering colleges and individual academic programs can now assess their students' ability more frequently and consistently. The scale allows a college to produce regularly updated evidence for summative accreditation reviews and to improve the programs and other interventions used to improve students' contextual competence.

The scale distinguishes levels of contextual competence according to gender, which the literature indicates might affect students' perception of their abilities. Knowing such differences will help researchers to more accurately measure inequalities in women's and men's perceptions of their abilities and to suggest changes to policy and practice to address them. Previous research indicated there are differences between self-report and objective measures of ability. Our analyses confirmed what the literature and theory suggest of contextual competence: students' self-rated ability rises as they move from sophomore to junior and senior status; the level of competence is higher if students are in a major that focuses on the societal and global implications of engineering solutions; and alumni, who have more experience in the field, believe they are more knowledgeable about the effects of context than are senior students. Researchers and educators can confidently use the scale to measure student and alumni perceptions of their contextual competence, and how that ability may actually change according to student demographics or instructor interventions.

The findings of this study also have implications for engineering curricula and teaching practice. While most engineering students conduct design tasks by the junior or senior level

(e.g., capstone courses), some engineering programs include design components in lower-division courses or the first-year curriculum. We found that participating in required first-year design courses may slightly improve students' contextual competence; this finding suggests the benefits of requiring first-year design courses.

This study used self-ratings of contextual competence. The validity analyses could be further strengthened by comparing students' self-reported assessments to performance evaluations by faculty members or engineering educators. For example, some engineering faculty members test their students' contextual competence level on midterm or final exams in capstone design courses. If the grade, students' self-assessment of contextual competence, and faculty members' rating of their students were concurrently collected, the three components could be used to examine the validity of students' self-rating of contextual competence.

Although we used test content, internal structure, and relations to other variables to validate the contextual context scale, we could not investigate validity evidence based on response processes or consequences of the scale due to limitations stemming from the study's design. Further validity tests using this scale should be conducted in the future. For example, evidence of the appropriateness of the response process would be gained by comparing students' self-reports to design course grades and faculty assessments. The long-term consequences of using the scale are important because curricular reforms based on the results of using the scale can lead to more effective development of the contextual competence of future engineers.

A comprehensive and iterative survey-development process that involves content experts is an important first step for providing evidence of validity. Conducting a pilot study and factor analysis of the results is another important but often overlooked step in examining the internal structure of a scale. Finally, appropriate statistical procedures should be used to provide evidence of how the scale may relate to other variables to ensure that the resulting scale adequately measures important dimensions of the core concept as is predicted by the underlying theory. The process used by this study can be adopted (or adapted) by other researchers to develop measures of learning outcomes and evaluate the validity of their use for both local and large-scale assessments.

Acknowledgments

This study was supported by the National Science Foundation Grant EEC-0550608. Opinions expressed herein are those of the authors, and no endorsement by the National Science Foundation should be inferred.

This article substantially extends H. K. Ro, L. R. Lattuca, D. Merson, and P. T. Terenzini (2012).

References

- ABET. (1997). *Engineering criteria 2000: Criteria for accrediting programs in engineering in the United States*. Baltimore, MD: Author.
- ABET Engineering Accreditation Commission. (2008). Criteria for accrediting engineering programs: Effective for evaluations during the 2009–2010 accreditation cycle. Retrieved from <http://www.abet.org/Linked%20Documents-UPDATE/Criteria%20and%20PP/E001%2009-10%20EAC%20Criteria%2012-01-08.pdf>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.

- Atman, C. J. (2009). *Surprise happens to the prepared mind: Linking context and creativity in engineering design*. Paper presented at the ACM Creativity and Cognition in Engineering Design Conference, Berkeley, CA.
- Atman, C. J., Cardella, M. E., Turns, J., & Adams, R. (2005). Comparing freshman and senior engineering design processes: An in-depth follow-up study. *Design Studies*, 26(4), 325–357. doi: 10.1016/j.destud.2004.09.005
- Atman, C. J., Kilgore, D., Yasuhara, K., & Morozov, A. (2008). *Considering context over time: Emerging findings from a longitudinal study of engineering students*. Paper presented at the Research on Engineering Education Symposium, Davos, Switzerland.
- Baruch, Y. (1999). Response rate in academic studies: A comparative analysis. *Human Relations*, 52(4), 421–438. doi: 10.1177/001872679905200401
- Besterfield-Sacre, M., Atman, C. J., & Shuman, L. J. (1998). Engineering student attitudes assessment. *Journal of Engineering Education*, 87(2), 133–141. doi: 10.1002/j.2168-9830.1998.tb00333.x
- Bielefeldt, A. (2008). *Cultural competency assessment*. Paper presented at the ASEE Annual Conference and Exposition, Pittsburgh, PA.
- Bogusch, L. L., Turns, J., & Atman, C. J. (2000). *Engineering design factors: How broadly do students define problems?* Paper presented at the Frontiers in Education Conference, Kansas City, MO. Retrieved from <http://fie-conference.org/fie2000/papers/1182.pdf>
- Bordogna, J., Fromm, E., & Ernst, E. W. (1993). Engineering education: Innovation through integration. *Journal of Engineering Education*, 82(1), 3–8. doi: 10.1002/j.2168-9830.1993.tb00065.x
- Bowman, N. A. (2010). Can 1st-year college students accurately report their learning and development. *American Educational Research Journal*, 47, 466–496.
- Brainard, S. G., & Carlin, L. (1998). A six-year longitudinal study of undergraduate women in engineering and science. *Journal of Engineering Education*, 87(4), 369–375. doi: 10.1002/j.2168-9830.1998.tb00367.x
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. doi: 10.1007/BF02310555
- de Gruijter, D. N. M., & van der Kamp, L. J. T. (2008). *Statistical test theory for the behavioral sciences*. Boca Raton, FL: Chapman & Hall/CRC.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1), 1–38.
- DeVellis, R. F. (2003). *Scale development: Theory and applications*. Thousand Oaks, CA: Sage.
- Dey, E. L. (1997). Working with low survey response rates: The efficacy of weighting adjustments. *Research in Higher Education*, 38(2), 215–227. doi: 10.1023/a:1024985704202
- Downey, G. L., Lucena, J. C., Moskal, B. M., Parkhurst, R., Bigley, T., Hays, C., . . . Nichols-Belo, A. (2006). The globally competent engineer: Working effectively with people who define problems differently. *Journal of Engineering Education*, 95(2), 107–122. doi: 10.1002/j.2168-9830.2006.tb00883.x
- Dym, C. L., Wesner, J. W., & Winner, L. (2003). Social dimensions of engineering design: Observations from Mudd Design Workshop III. *Journal of Engineering Education*, 92(1), 105–107. doi: 10.1002/j.2168-9830.2003.tb00746.x

- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59, 395–430. doi: 10.3102/00346543059004395
- Garson, G. D. (2010). Factor analysis. *Statnotes: Topics in Multivariate Analysis*. Retrieved from <http://faculty.chass.ncsu.edu/garson/PA765/factor.htm>
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60(1), 549–576. doi: 10.1146/annurev.psych.58.110405.085530
- Haertel, E. H. (2006). Reliability. In R. Brennan (Ed.), *Educational measurement* (pp. 65–110). Westport, CT: Praeger.
- Häussler, P., & Hoffmann, L. (2002). An intervention study to enhance girls' interest, self-concept, and achievement in physics classes. *Journal of Research in Science Teaching*, 39(9), 870–888. doi: 10.1002/tea.10048
- Holland, J. L. (1973). *Making vocational choices: A theory of careers*. Englewood Cliffs, NJ: Prentice Hall.
- Kastenber, W. E., Hauser-Kastenber, G., & Norris, D. (2006). *An approach to undergraduate engineering education for the 21st century*. Paper presented at the 36th Annual Frontiers in Education Conference, San Diego, CA. Retrieved from <http://fie-conference.org/fie2006/index.html>
- Kilgore, D., Atman, C. J., Yasuhara, K., Barker, T. J., & Morozov, A. (2007). Considering context: A study of first-year engineering students. *Journal of Engineering Education*, 96(4), 321–334. doi: 10.1002/j.2168-9830.2007.tb00942.x
- Kranov, A. A., Hauser, C., Olsen, R., & Girardeau, L. (2008). *A direct method for teaching and assessing professional skills in engineering programs*. Paper presented at the ASEE Annual Conference and Exhibition, Pittsburgh, PA.
- Lattuca, L. R., Terenzini, P. T., Harper, B. J., & Yin, A. C. (2010). Academic environments in detail: Holland's theory at the subdiscipline level. *Research in Higher Education*, 51(1), 21–39. doi: 10.1007/s11162-009-9144-9
- Lattuca, L. R., Terenzini, P. T., & Volkwein, J. F. (2006). *Engineering change: A study of the impact of EC2000*. Baltimore, MD: ABET.
- Lau, A. S. (2004). *Life-centered design – A paradigm for engineering in the 21st century*. Paper presented at the ASEE Annual Conference and Exposition, Salt Lake City, UT.
- Leighton, J. P. (2008). Validity. In N. J. Salkind & K. Rasmussen (Eds.), *Encyclopedia of educational psychology* (Vol. 2, pp. 995–1000). Thousand Oaks, CA: Sage.
- Loui, M. C. (2006). Ethics and the development of professional identities of engineering students. *Journal of Engineering Education*, 94(4), 383–390. doi: 10.1002/j.2168-9830.2005.tb00866.x
- McKenna, A., Plumb, C., Kremer, G. E., Yin, A. C., & Ro, H. K. (2011). *Approaches to engaging students in engineering design and problem solving*. Paper presented at the ASEE Annual Conference and Exposition, Vancouver, Canada. Retrieved from <http://www.asee.org/public/conferences/1/papers/599/view>
- Mikic, B., & Grasso, D. (2002). Socially-relevant design: The Toytech project at Smith College. *Journal of Engineering Education*, 91(3), 319–326. doi: 10.1002/j.2168-9830.2002.tb00709.x
- Muchinsky, P. M. (2004). When the psychometrics of test development meets organizational realities: A conceptual framework for organizational change, examples, and recommendations. *Personnel Psychology*, 57(1), 175–209. doi: 10.1111/j.1744-6570.2004.tb02488.x

- National Academy of Engineering (NAE). (2004). *The engineer of 2020: Visions of engineering in the new century*. Washington, DC: The National Academies Press.
- National Academy of Engineering (NAE). (2006). *Educating the engineer of 2020: Adapting engineering education in the new century*. Washington, DC: The National Academies Press.
- National Research Council (NRC). (1995). *Engineering education: Designing an adaptive system*. Washington, DC: National Academies Press.
- Palmer, B., Terenzini, P. T., McKenna, A., Harper, B., & Merson, D. (2011). *Design in context: Where do the engineers of 2020 learn this skill?* Paper presented at the ASEE Annual Conference and Exposition, Vancouver, BC. Retrieved from <http://www.asee.org/public/conferences/1/papers/599/view>
- Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research*. San Francisco, CA: Jossey-Bass.
- Ro, H. K., Lattuca, L. R., Merson, D., & Terenzini, P. T. (2012). *Measuring engineering students' contextual competence*. Paper presented at the ASEE Annual Conference and Exposition, San Antonio, TX. Retrieved from <http://www.asee.org/public/conferences/8/papers/4880/view>
- Sax, L. J., Gilmartin, S. K., & Bryant, A. N. (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education*, 44(4), 409–432. doi: 10.1023/A:1024232915870
- Seymour, E. (1995). The loss of women from science, mathematics, and engineering undergraduate majors: An explanatory account. *Science Education*, 79, 437–473. doi: 10.1002/sce.3730790406
- Shen, Y., Jesiek, B. K., & Chang, Y. (2011). *Cultural orientation and global competency: A comparative assessment of engineering students*. Paper presented at the ASEE Annual Conference & Exposition, Vancouver, BC. Retrieved from <http://www.asee.org/public/conferences/1/papers/1992/view>
- Sheppard, S. D., Macatangay, K., Colby, A., & Sullivan, W. M. (2008). *Educating engineers: Designing for the future of the field*. San Francisco, CA: Jossey-Bass.
- Smith, T. W. (1995). Trends in non-response rates. *International Journal of Public Opinion Research*, 7(2), 157–171. doi: 10.1093/ijpor/7.2.157
- Stark, J. S., & Lowther, M. A. (1988). *Strengthening the ties that bind: Integrating undergraduate liberal and professional study*. Report of the Professional Preparation Network. Ann Arbor, MI: Professional Preparation Project.
- Stark, J. S., Lowther, M. A., & Hagerty, B. M. K. (1987). Faculty perceptions of professional preparation environments: Testing a conceptual framework. *Journal of Higher Education*, 58(5), 530–561. doi: 10.2307/1981786
- Steeh, C. G. (1981). Trends in nonresponse rates, 1952–1979. *Public Opinion Quarterly*, 45(1), 40–57. doi: 10.1086/268633
- Strauss, L. C., & Terenzini, P. T. (2005). *Measuring student performance on EC2000 criterion 3.a–k*. Paper presented at the ASEE Conference and Exposition, Portland, OR.
- Suen, H. K. (2008). Measurement. In N. J. Salkind & K. Rasmussen (Eds.), *Encyclopedia of educational psychology* (Vol. 2, pp. 641–647). Thousand Oaks, CA: Sage.
- Thorndike, E. L. (1920). A constant error on psychological rating. *Journal of Applied Psychology*, 9, 25–29. doi: 10.1037/h0071663
- Trochim, W. M. K. (2006). Research methods knowledge base. Retrieved from <http://www.socialresearchmethods.net/kb>

- Ward, M., Gruppen, L., & Regehr, G. (2002). Measuring self-assessment: Current state of the art. *Advances in Health Sciences Education*, 7(1), 63–80. doi: 10.1023/a:1014585522084
- Whitt, E. J., Pascarella, E. T., Elkins Neisheim, B. S., Marth, B. P., & Pierson, C. T. (2003). Differences between women and men in objectively measured outcomes, and the factors that influence those outcomes, in the first three years of college. *Journal of College Student Development*, 44(5), 587–610. doi: 10.1353/csd.2003.0060
- Zhang, Y. (2000). Using the Internet for survey research: A case study. *Journal of the American Society for Information Science*, 51(1), 57–68. doi: 10.1002/(sici)1097-4571(2000)51:1<57::aid-asi9>3.0.co;2-w

Authors

Hyun Kyoung Ro is a research analyst in Institutional Research and Analysis at Carnegie Mellon University, 5000 Forbes Avenue, UTDC Building, Pittsburgh, PA 15213; hro@cmu.edu.

Dan Merson is an independent educational research and statistics consultant as well as a research associate for the Leonhard Center for Enhancement of Engineering Education, The Pennsylvania State University, 201 Hammond Building, University Park, PA 16802; DanMerson@gmail.com.

Lisa R. Lattuca is a professor in the Higher Education Program and the Center for the Study of Higher and Postsecondary Education, the University of Michigan, 610 E. University Avenue, Room 2117, Ann Arbor, MI 48109-1259; LLatt@umich.edu.

Patrick T. Terenzini is a distinguished professor and senior scientist emeritus in the Higher Education Program and the Center for the Study of Higher Education, The Pennsylvania State University, 400 Rackley Building, University Park, PA 16802; Terenzini@psu.edu.