Developing Self-Report Instruments to Measure ABET EC 2000 Criterion 3 Professional Outcomes

Jason C. Immekus, Sara Tracy, Jin Eun Yoo, Susan J. Maller, Brian F. French, William C. Oakes

Purdue University

Abstract

The Accreditation Board for Engineering and Technology's *Engineering Criteria 2000* (ABET EC2000)[1] Criterion 3 Programs Outcomes and Assessment specifies outcomes college graduates are expected to know and demonstrate following graduation from accredited engineering programs. The generality of Criterion 3 objectives requires engineering programs to enunciate the desired program outcomes. In recognition of this complex task, this paper presents the process the Engineering Projects in Community Service (EPICS) program is taking to develop precise operational definitions of the outcomes for the design, implementation, and validation of assessment instruments. For didactic purposes, the focus is on outcomes related to professional skills (e.g., communication skills, teamwork) that lend themselves to assessment through self-report instruments (e.g., surveys). This paper is designed to serve as an instructional resource to programs considering the use of self-report instruments to measure professional Criterion 3 outcomes.

1. Criterion 3

ABET EC 2000[1] Criterion 3 Programs Outcomes and Assessment specifies eleven outcomes industry and academia expect college graduates to know and demonstrate following completion of accredited engineering programs. The criteria are intended to enable accredited engineering programs to provide key skills students will need to pursue an engineering career. Specifically, Criterion 3 outcomes include:

(a) an ability to apply knowledge of mathematics, science, and engineering
(b) an ability to design and conduct experiments, as well as to analyze and interpret data
(c) an ability to design a system, component, or process to meet desired needs
(d) an ability to function on multi-disciplinary teams
(e) an ability to identify, formulate, and solve engineering problems
(f) an understanding of professional and ethical responsibility
(g) an ability to communicate effectively
(h) the broad education necessary to understand the impact of engineering solutions in a global and societal context
(i) a recognition of the need for, and an ability to engage in life-long learning
(j) a knowledge of contemporary issues
(k) an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice.

1a. Challenges of Criterion 3 Outcomes

Criterion 3 outcomes identify the key professional skills accredited engineering programs must promote in their graduates by aligning a program's instructional objectives. There are many issues and challenges imposed on engineering programs that desire to measure Criterion 3 outcomes.

Criterion 3 benefits engineering programs in many ways. First, the objectives reflect the understanding of the dynamic nature of engineering practice as it strives to meet the challenges imposed by the demands of an ever-changing society. Not only are engineering students required to demonstrate a strong understanding and knowledge of the application of technical skills, they also are encouraged to show proficiency across a range of professional skills (e.g., communication skills). These skills are increasingly important in today's engineering workplace. Programs that promote the outcomes outlined by Criterion 3 are expected to produce graduates who will be equipped to pursue a life-long engineering career.

There are clear challenges for engineering educators who attempt to meet ABET criteria within their educational programs[28]. Perhaps the most salient challenge to the assessment of Criterion 3 criteria is the lack of operational definitions for the learning outcomes. For example, there are many acceptable perspectives of what constitutes a student's knowledge of contemporary issues (Criterion 3.j), none of which are specified by ABET. This provides an opportunity for disparate measurement of a single outcome, according to the definition adopted by each educational program. Another challenge posed to accredited engineering programs is identifying the sub-domains most indicative of the outcomes. For instance, a student's ability to communicate effectively (Criterion 3.g) can be considered in terms of written or oral ability in formal and informal contexts, communication to a technical or nontechnical audience, and in relation to independent or group work.

In response to these challenges, this paper is designed to provide engineering educators with procedural information to make the design and evaluation of Criterion 3 assessment instruments more consistent. Specifically, outcomes related to professional skills that lend themselves to self-report measures are discussed. Specifically, this paper is intended to provide engineering educators (a) guidelines for implementing a clear and rigorous measurement process to assess professional program outcomes, (b) a survey of available instruments and tools, and (c) procedures for evaluating the quality of the developed assessment instrument(s).

1b. EPICS

EPICS is a service learning program that enables long-term projects in which teams of engineering undergraduates are matched with community service agencies that request technical assistance[21, 23]. EPICS was formally established in 1995[22] and is now a feature at fifteen undergraduate engineering programs across the U.S. More information on EPICS can be found at the website *http://epics.ecn.purdue.edu/*.

Through participation in EPICS students learn many valuable lessons in engineering, including the role of the partner, or "customer," in defining an engineering project; the necessity of teamwork; the difficulty of managing and leading large projects; the need for skills and knowledge from many different disciplines; and the art of solving technical problems. They also learn many valuable lessons in citizenship, including the role of community service in our society; the significant impact that their engineering skills can have on their community; and that assisting others leads to their own substantial growth as individuals, engineers, and citizens. The learning objectives for EPICS are well-matched with Criterion 3 professional outcomes.

2. Methods

Broadly, there are three phases to the process of the development of self-report instruments to measure professional outcomes: (a) define learning outcomes, (b) design assessment instruments, and (c) evaluate reliability and validity evidence of instruments. The subsequent sections of this paper present one approach taken to meet the demands of these processes to provide guidelines and suggestions for other programs.

2a. Define learning outcomes

The objectives engineering programs and Criterion 3 seek to promote and measure are complex abilities that encompass specific, interrelated skills. Figure 1 depicts a visual representation of a hypothetical learning outcome. The apex consists of the general outcome, as per Criterion 3. Typically, learning outcomes can be represented in terms of sub-domains, located at the second-stratum. For instance, attentiveness and perceptiveness are sub-domains that can be used to represent the broad domain of communication competence (3.g)[19]. The third-stratum corresponds to the skills that can be measured directly, operationalized in terms of items. A sound definition and measurement of a particular learning outcome requires careful consideration of the theoretically identified nature of the outcome.

The professional skills of communication competence (3.g), teamwork (3.d), an understanding of a global and societal contexts (3.h), and a knowledge of contemporary issues (3.j.) are outcomes EPICS and Criterion 3 seek to encourage and will be used to illustrate how engineering educators can approach the conceptualization of these skills.

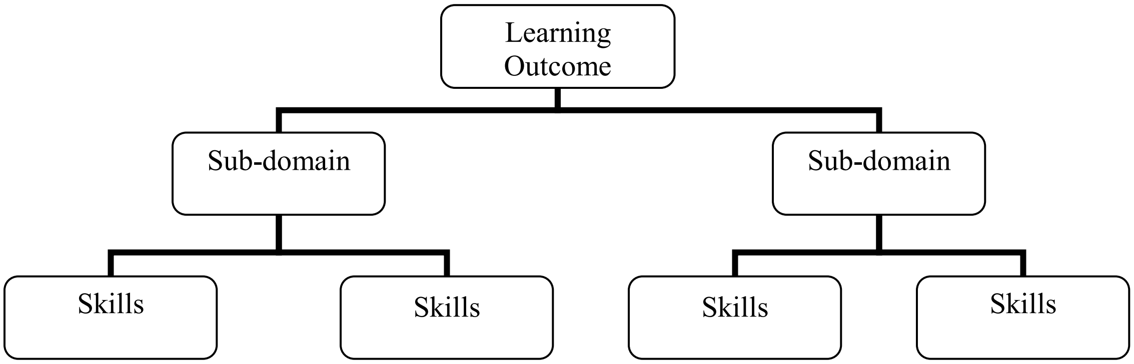

Figure 1. Visual representation of a hypothetical learning outcome

Communication competence can be defined as, "the ability to choose among available communicative behaviors in order that he may successfully accomplish his own interpersonal goals during an encounter while maintaining the face and line of his (or her) fellow interactants within the constraints of the situation"[31]. This definition indicates that communication competence includes any or all of the mechanisms (i.e., written, oral) students use to transmit their intended message to team members, faculty, and/or their community partner. Given that EPICS students are required to communicate with various audiences (e.g., team members, community partners), interpersonal communication is one of the sub-domains we have identified to represent Criterion 3.g. Based on theory, empirical research, and the goals of EPICS, the sub-domains (second-stratum) include: perceptiveness and attentiveness. Scores across these two sub-domains will be used to represent a student's level of communication competence. Separate measures of written and oral communication proficiencies also are being developed.

Other Criterion 3 outcomes are not as readily definable and require conceptualization in terms of the dynamics of a particular engineering program. For instance, the literature on effective teams abounds in representative sub-domains[11]. However, the sub-domains used to represent teamwork are inconsistent across studies. For instance, Hoegl and Gemuenden[11] identified six elements of teamwork, including communication, coordination, effort, and cohesion; other research identified as many as thirteen characteristics of effective teams[6,12]. Engineering educators may take note that the plethora of literature on teamwork is presented in terms of industry-based teams and, in some cases, the content of the measures presented in this research will need to be modified to meet the needs of a specific engineering program. For example, one sub-domain adapted for our teamwork scale is *norms*, or the degree team members adhere to and fulfill the duties required by their assigned role (e.g., team leader, team webmaster, liaison with the community partner). Students' ability to fulfill the obligations of their assigned role is critical in EPICS because teams are composed of students with wide ranging levels of expertise and experience.

An understanding of the impact of engineering solutions in a global and societal context (3.h.) and a knowledge of contemporary issues (3.j.) are professional outcomes that can , in part, be represented by social responsibility, or the degree an individual feels connected to the community[9,25]. EPICS seeks to promote social responsibility through students' involvement with nonprofit community partners. Social responsibility can be conceptualized as a continuous construct that has been represented in terms of five categories (i.e., exploration, clarification, realization, activation, and internalization) [9]. At the highest level, internalization, students are characterized as feeling deep involvement for particular social causes and as a result, often changing their career choices[9]. Currently, our measure of social responsibility includes sub-scales corresponding to *exploration*, *realization*, and *internalization*, as per previous research[25]. After controlling for extraneous factors (e.g., matched sample), scale scores of students participating in EPICS that show growth across the sub-scales, based on pre- and post-test scores, lend support that EPICS is encouraging students to feel a sense of connection to their community.

As shown, several steps are necessary to operationalize Criterion 3 professional outcomes. First, engineering educators should select Criterion 3 outcomes that are linked to one's program. Second, the outcomes should be conceptualized in terms of theory and empirical research. Finally, within each outcome, representative sub-domains should be selected in accordance to one's program. Conceptualizing learning outcomes is a necessary foundation to build representative assessment instruments.

2b. Design assessment instruments

Self-report instruments are valuable tools for evaluation of Criterion 3 outcomes. Specifically, instruments can be designed or adapted to meet the needs of engineering programs, are easy to administer, and the psychometric properties (e.g., reliability, validity) can be evaluated, documented, and used to improve the instrument. Typically, there are three considerations engineering educators need to be aware of when selecting or designing assessment instruments. These include: (a) determining the availability of pre-existing measures, (b) adapting pre-existing measures to meet the needs of a particular engineering program, and (c) constructing new instruments.

Pre-existing instruments, often published in research[6,19,25], provide engineering educators a starting point to measure professional outcomes. When considering pre-existing measures, it is important to examine for what purpose the scale was originally designed to measure (e.g., teamwork) and to whom it was administered (e.g., employees, managers). For example, Campion, Medsker, and Higgs[6] published a scale used to measure team characteristics as they pertained to industry. Several of the instrument's sub-domains (e.g., *social support*, *interdependence*) and items (e.g., everyone on my team does their fair share of work[6]) were applicable to the skills targeted by EPICS. Further, published measures demonstrate how thoroughly the construct (e.g., social responsibility) was operationalized and measured. Such information can be used to evaluate the scale's psychometric properties. Selected scales are listed in Table 1 for further reference.

Adapting pre-existing measures is another way engineering educators can use self-report instruments to assess program outcomes. For instance, the EPICS teamwork scale is partly based on the adaptation of the scale used by Campion et al.[6]. Some of the sub-domains of interest were not measured. The additional teamwork sub-domains we sought to measure were researched in other literature to form a theoretical base for the items developed. By adapting pre-existing instruments and incorporating additional sub-domains that matched our goals, we were able to develop a scale that reflects the aspects of teamwork students are exposed to in EPICS.

The use or adaptation of pre-existing measures may not be an option for the assessment of engineering program outcomes for several reasons. For example, a scale may be published but not available without a fee. In the event that a complete scale needs to be developed, educators will need to identify the outcome and corresponding sub-domains they desire to measure, as previously described. Instructional materials to guide the scale development process are readily available[10].

Table 1. Published Instruments with Measured Domains and Sub-Domains

Scale	Domains	Sub-domains
Communicator Style Questionnaire[18]	Communicator Style	1. Dominant 2. Dramatic 3. Contentious 4. Animated 5. Impression leaving 6. Open 7. Relaxed 8. Attentive 9. Friendly 10. Communicator Image
Interaction Involvement Scale[7]	Communication Involvement	1. Perceptiveness 2. Understanding 3. Attentiveness
Conversation Skills Rating Scale[27]	Conversation Skills	1. Altercentrism 2. Vocal Confidence 3. Posture
Personal Report of Communication Apprehension[16]	Communication Apprehension	Communication apprehension
Center for Research on Learning and Teaching Questionnaire[14]	Social Attitudes/Values	1. Social Attitudes 2. Social Values
Scale of Service Learning Involvement[25]	Social Responsibility	1. Exploration 2. Realization 3. Internalization
Cognitive Flexibility Scale[15]	Life-Long Learning	Cognitive Flexibility
Oddi Continuing Learning Inventory[24]	Life-Long Learning	Self-directed Learning
Self-directed Learning Readiness Scale[13]	Life-Long Learning	1. Openness to learning opportunities 2. Self-concept as an effective self-learner 3. Initiative and independence in learning 4. Acceptance of responsibility 5. Love to Learn 6. Creativity 7. Future orientation 8. Ability to use basic skills and problem-solving skills
Work Group Characteristics Measure[6]	Team work	1. Job Design 2. Interdependence 3. Composition 4. Context 5. Process
Intrinsic Value Scale[26]	Personal Value	Intrinsic Value

2c. Reliability and Validity Evidence of Instruments

The *Standards for Educational and Psychological Measurement*[3] state that evidence of an instrument's psychometric properties (e.g., reliability, validity) must be provided to substantiate the use of obtained scores. This section discusses the rigorous process we are using to examine the psychometric properties of the EPICS scales. Notably, preliminary data collection for EPICS is scheduled for the Spring 2004 semester.

Reliability

Reliability refers to the degree an instrument produces consistent scores[2,8]. Specifically, reliability indicates the degree test scores are not affected by random measurement error[30]. Internal consistency and test-retest reliability estimates will be used to investigate the reliability of EPICS scales. Internal consistency reliability indicates the degree an instrument's items function together to produce a test score[8], and can be estimated based on a single test administration. High internal reliability estimates (e.g., above .90) are desired[20]. Test-retest reliability indicates whether an instrument provides stable scores when administered repeatedly over time[8]. Within this study, the scales will be administered before and after students participate in EPICS. Scale reliability is *a necessary but not sufficient* condition for validity[9].

Validity

Validity is the degree an instrument measures what it was designed to measure[17]. Validated test scores provide engineering educators with meaningful information for evaluating program outcomes. Although there are distinct types of validity (e.g., content, criterion, construct), all validity evidence is ultimately leveled at construct validity[17].

Content validity. Content validity refers to the correspondence between test content and its associated domain. Typically, content validity is established through content review, conducted by content experts. For the EPICS scales, content validity is being established through a panel of students and faculty familiar with the measured outcomes.

Criterion validity. Criterion validity is the relationship, expressed as a correlation coefficient, between an instrument and a criterion, typically another instrument. The relationship between the instrument and criterion can be determined based on the administration of the instruments at the same time (concurrent) or when the instrument is used to predict the criterion (predictive). For this study, predictive validity is of particular importance, because scale scores will be used to predict future outcomes (e.g., academic achievement, retention in major). Although criterion validity provides critical information pertaining to the use of test scores, it cannot be used as sole evidence of an instrument's validity. Specifically, the relationship between and instrument and a criterion may be inflated due to *method bias*. Method bias occurs when factors extraneous to test constructs and specific to the testing method influence test scores. Another potential problem is the possibility that the criterion measure may not be psychometrically sound[17].

Therefore, before the integrity of an instrument's criterion related validity can be concluded, evidence of its construct validity must be provided.

Construct validity. Construct validity deals with whether a scale measures what it was designed to measure. Factor analytic procedures are the most commonly used methods to assess an instrument's factor structure. Factor analysis procedures fall under two categories: exploratory (EFA) and confirmatory (CFA) factor analysis. EFA is data driven; CFA is theory driven[20]. EFA is useful when there is no *a priori* information available to suggest the relationship between observed variables (e.g., items) and latent traits (e.g., communication competence). Because our scales under development are based on theoretical models, CFA will be used to document evidence of construct validity.

CFA is the theoretically preferred method to test whether an instrument represents its measurement model [4,5]. The degree an instrument "fits" its measurement model is evaluated in terms of various "fit" indices (e.g., chi-square, Goodness-of-Fit) [5]. For instance, the EPICS interpersonal teamwork scale, thus far, consists of four sub-scales (e.g., composition, interdependency, norms, and goals), each comprised of a set of items. Once data is collected on the scale, CFA will be used to empirically test whether the items correspond to their respective sub-scales. CFA provides direct evidence regarding whether an instrument's test scores can be used as valid indicators of a measured construct.

Another benefit of CFA is that the factorial similarity of an instrument across groups can be examined. Multisample CFA is used to test the theoretical model of the instrument across groups. If the model "fits" across groups, as determined by a variety of fit statistics, the factor structure is said to be *invariant* (not different), and it is concluded that scores do not measure intended test constructs differently across groups. An advantage of multisample CFA is that specific factor loadings, their associated error variances, and the relationship between factors can be individually tested to determine the specific differences between groups and to better understand how an instrument functions across groups. Particularly, measurement invariance of test scores is critical for between-group analyses (e.g., ANOVA).

3. Conclusion

The goal of this paper was to provide engineering educators with information to proceed with the process of using self-report instruments to measure Criterion 3 professional skills. The demands of ABET EC2000 Criterion 3 outcomes place accredited engineering programs in a situation in which they must demonstrate that students are developing the skills they will need to pursue an engineering career. Self-report instruments can be a valuable tool for evaluating whether an engineering program is reaching its goals of encouraging student proficiencies. Self-report instruments can go beyond simply asking students if they learned a topic or achieved an outcome.

The steps engineering educators should consider through the design process of their self-report instruments to measure their program and Criterion 3 outcomes include:

1. Formal operational definitions of outcomes should be based on theory, empirical research, Criterion 3, and the goals of the engineering program.

2. Self-report instruments to assess specific program outcomes can be based on using pre-existing instruments, adapting pre-existing instruments to meet particular program needs, or constructing new instruments. In the event that a new scale will need developed, it should be designed in accord with theoretical considerations, such as identifying the sub-domains that will be used to comprise the overall measure.

3. Instruments should be selected or designed in accordance to the population the scale will be administered and how obtained scores will be used.

4. Instruments must demonstrate adequate psychometric properties (e.g., reliability, validity) in order for obtained scores to be considered meaningful to the assessment of engineering program and Criterion 3 outcomes.

References

1. ABET (1999). *Criteria for Accrediting Engineering Programs*. The Engineering Accreditation Commision of the Accreditation Board for Engineering and Technology. http://www.abet.org/eac/eac.htm.

2. Allen, Mary, J., & Wendy M. Yen (1979). *Introduction to measurement.* Monterey, CA: Brooks/Cole Publishing.

3. American Educational Researchers Association, American Psychological Association, and the National Council on the Measurement in Education (1999). *Standards for educational and psychological measurement*.

4. Bollen, Kenneth (1989). *Structural equations with latent variables*. New York, NY: Wiley.

5. Byrne, Barbara (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum.

6. Campion, Michael A., Gina J. Medsker, & A. Catherine Higgs (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel Psychology, 46*, 823-850.

7. Cegala, Donald J. (1981). Interaction involvement: A cognitive dimension of communication competence. *Communication Education, 30*, 109-121.

8. Crocker, Linda & James Algina (1984). *Introduction to classical and modern test theory.* Fort Worth, TX: Harcourt Brace.

9. Delve, C.L., S. D. Mintz, & G. M. Stewart (1990). Promoting values development through community service: A design. *New Directions for Student Services, 4*, 7-29.

10. Fowler, Floyd J. (1993). Survey research methods. Newbury Park, CA: Sage.

11. Hoegl, M., & H. G. Gemuenden (2001). Teamwork quality and the success of innovative projects: A theoretical concept and empirical evidence. *Organization Science, 12*, 435-449.

12. Hyatt, D. E., & T. M. Ruddy (1997). An examination of the relationship between work group characteristics and performance: One more into the breech. *Personnel Psychology, 50*, 553-585.

13. Litzinger, Thomas, John Wise, SangHa Lee, & Bjorklund (2003). Assessing readiness for self-directed learning. In *Proceeding of ASEE Annual Conference*. American Society for Eng. Education.

14. Markus, Gregory B., Jeffrey P. Howard, & David C. King (1993). Integrating community service and classroom instruction enhances learning: Results from an experiment. *Educational Evaluation and Policy Analysis, 15*, 410-419.

15. Martin, M. M. & Rubin, R. B. (1995). A new measure of cognitive flexibility. *Psychological Reports, 76*, 623-626.

16. McCroskey, James C., Michael J. Beatty, Patricia Kearney, & Timothy G. Plax (1985). The content validity of the PRCA-24 as a measure of communication apprehension across communication contexts. *Communication Quarterly, 33*, 165-173.

17. Messick, Samual (1995). Validity of psychological assessment. *American Psychologist, 50*, 741-749.

18. Montgomery, Barbara M., & Robert W. Norton (1981). Sex differences and similarities in communicator style. *Communication Monographs, 48*, 121-132.

19. Norton, R. W. (1978). Foundation of a communicator style construct. *Human Communication Research, 4*, 99-112.

20. Nunnally, Jum & Ira Bernstein (1994). *Psychometric Theory* (3rd edition). New York, NY: McGraw-Hill.

21. Oakes, William C., & A. G. Rud Jr. (2001). The EPICS model in engineering education: Perspective on problem-solving abilities needed for success beyond school. In H. Doerr & R. Lesh (eds.). *Beyond Constructivism: A models & modeling perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates.

22. Oakes, William C., Leah H. Jamieson, & Edward J. Coyle (2001). EPICS: Meeting EC 2000 through service-learning, In *Proceeding of ASEE Annual Conference*. American Society for Eng. Education.

23. Oakes, William C., Edward J. Coyle, Richard Fortek, Jeffery Gray, Leah H. Jamieson, Jennifer Watia, & Ronald Wukasch (2000). EPICS: Experiencing engineering design through community service projects. In *Proceeding of ASEE Annual Conference*. American Society for Eng. Education.

24. Oddi, Lorys F. (1986). Development and Validation of an Instrument to Identify Self-Directed Continuing Learners. *Adult Education Quarterly, 36, 97-107.*

25. Olney, Cynthia, & Steve Grande (1995). Validation of a scale to measure development of social responsibility. *Michigan Journal of Community Service Learning*, 43-53.

26. Pintrich, P., & E. W. DeGroot (1990). Motivational and self-regulated learning components of classroom academic. *Journal of Educational Psychology, 82*, 33-40.

27. Spitzberg, Brian H., & H. Thomas Hurt (1994). The measurement of interpersonal skills in instructional contexts. In R. B. Rubin & E. E. Graham (eds.), *Communication research measures: A sourcebook* (pp. 28-45). New York: Guilford Press.

28. Soundarajan, Neelam (2001). Assessing and refining EC 2000, In *Proceeding of ASEE Annual Conference*. American Society for Eng. Education.

29. Sundstrom, Eric, Michael McIntyre, Terry Halfhill, & Heather Richards (2000). Work groups: From the Hawthorne studies to work teams of the 1990s and beyond. *Group Dynamics: Theory, Research, and Practice, 4*, 44-67.

30. Thompson, Bruce (2003).Understanding reliability and coefficient alpha, really. In Bruce Thompson (Ed.), *Score Reliability: Contemporary thinking on reliability issues*. Thousand Oaks, CA: Sage.

31. Wiemann, P. (1977). Explication and test of a model of communicative competence. *Human Communication Research, 3*, 195-213.

Biographical Information

Jason C. Immekus, M.S., is an advanced doctoral student in the Educational Psychology program at Purdue University. His areas of specialization are applied measurement and research design. His research interests include test validity, structural equation modeling, and test development.

Sara Tracy is a master's student in the Educational Psychology program at Purdue University. Her research interests are in parents' influence in young children's learning and development.

Jin Eun Yoo earned her first master's in Gifted Education (May, 2001), earned her second master's in Applied Statistics (May, 2003), and has been working on her doctoral degree in Educational Measurement and Methodology since June of 2001. Her main research interest/specialty lies in large scale assessment, multivariate statistical analysis, item response theory, and missing data imputation.

Susan J. Maller is an associate professor in the Educational Studies at Purdue University. Her areas of specialization include psychometrics and quantitative research methods. Her research interests are in the areas of item response theory, structural equation modeling when used to examine test validity, and item and test bias. She is a Spencer Dissertation Fellow, Associate Editor of *Psychology in the Schools*, and on the Editorial boards of several national refereed journals.

Brian F. French is an Assistant Professor in Educational Studies at Purdue University. His areas of specialization include psychometrics, quantitative methodology, and research design. His research focuses on methods used to obtain test score validity evidence.

William C. Oakes is an Assistant Professor in the Department of Freshman Engineering at Purdue University, where he is a Co-Director of the EPICS Program. He has served on the board of the ASEE Freshman Programs and Educational Research Methods Division and the Steering Committee FIE. He is a former recipient of the ERM Apprentice Faculty Grant and is an Indiana Campus Compact Fellow.