AC 2012-4038: APPRAISAL SYSTEM FOR SUPERIOR ENGINEERING EDUCATION EVALUATION - INSTRUMENT SHARING AND SCHOLARSHIP (ASSESS)

Dr. Denny C. Davis P.E., Washington State University

Denny Davis is professor of chemical engineering and bioengineering at Washington State University. He launched and directed the Engineering Education Research Center between 2005 and 2011. His scholarly work addresses engineering design learning and assessment. He is a Fellow of the American Society for Engineering Education.

Prof. Michael S. Trevisan, Washington State University

Mike Trevisan is a professor of educational psychology at Washington State University and the Associate Dean for Research in the College of Education. For more than 17 years, he has worked with engineering educators across the country to develop engineering design curriculum and assessments for a variety of engineering disciplines. His key collaborator is Dr. Denny Davis, Washington State University.

Dr. Howard P. Davis, Washington State University

Howard Davis received degrees from the Evergreen State College (B.A. 1976), WSU (B.S. 1981, M.S. 1988), and the University of Oregon (Ph.D. 1993). He is currently a Clinical Assistant Professor in the Gene and Linda Voiland School of Chemical Engineering and Bioengineering. He has been the President and CEO of IPM, a medical device company and Total Dynamics, LLC, a software company. He is also on the board of directors of Developing World Technologies, a company started by former students of the capstone class that he teaches. His interests include engineering and entrepreneurship pedagogy and assessment, technology development, and clinical applications of biomedical instrumentation.

Dr. Shane A. Brown P.E., Washington State University

Shane Brown conducts research in conceptual and epistemological change, social capital, and diffusion of innovations. In 2011, he received the NSF CAREER Award to investigate how engineers think about and use concepts that academics consider to be important.

Dr. Brian F. French, Washington State University

Brian F. French is an Associate Professor of educational leadership and counseling psychology and Co-Director of the Learning and Performance Research Center at Washington State University. His area of research focuses on applied and methodological issues in educational and psychological measurement.

Appraisal System for Superior Engineering Education Evaluation-Instrument Sharing and Scholarship (ASSESS)

Abstract

Although engineering educators implement many educational innovations to improve student achievement, few evaluate the impacts of their innovations sufficiently to support confident adoption of their innovations by others. A national panel of 30 engineering education and evaluation professionals has called for a national resource to enable effective evaluation of engineering education projects. This paper reports on the process and framework for creating a library of superior evaluation instruments, the ASSESS system, that supports scholarly innovation in engineering education. The overarching goal of the ASSESS project is to create and test the system, and engage the user community to position ASSESS resources for successful adoption and implementation. The project seeks to disseminate evaluation instruments and to build the evaluation capacity of the engineering education community for more effective evaluation of engineering education development projects.

A team of engineering education and evaluation professionals and project consultants have established a framework for characterizing evaluation instruments, structuring a database, and creating a web interface that supports desired user functionality. Each instrument is characterized by a descriptive summary, applications to engineering education, underlying theoretical foundations, technical specifications (e.g., reliability, validity data), implementation features, and directions for accessing the instrument. Published information on the instrument and its testing are referenced. Instruments are described in ways sensitive to the needs of novice users as well as needs of evaluation experts.

The evaluation instrument inventory is contained in a web-based database that supports information storage, retrieval, and input. The database contains prose textual fields for descriptive information, as well as quantitative and selected item fields for searchable data on the instruments. Users are able to search the database for instruments of interest based on keyword searches and selected item characteristics that fit their evaluation needs. The interface will also support expert review and user feedback on instruments to facilitate instrument selection and to guide instrument improvement.

Adoption research is being conducted to determine factors that influence the use of the database. User needs and adoption preferences will be determined and used to guide the final deployment of the ASSESS system. Experts from the engineering education evaluation community will continue to review, guide, and contribute superior instruments that are important to the engineering education community.

Introduction

The world faces many challenges that will require innovative, responsible contributions by engineers and other professionals[1-6]. At the same time, national leaders are expressing concerns about the educational approaches used to prepare the next generation of engineers[7-13]. Engineering education scholars are responding by calling for research-driven changes to engineering education that can attract, retain, and properly prepare diverse students with thinking skills required for the challenging and rapidly changing engineering profession[7, 8, 14-21]. On a broader scale, the public is asking whether the millions of taxpayer dollars invested annually in engineering education research are changing the practice of educating engineers to meet present and future challenges[22, 23].

Questions about the effectiveness of engineering education research investments can take different forms. To what extent are engineering education innovations effective in improving the recruitment, retention, and success of engineering students? How effectively are they improving the preparation of engineering and engineering technology graduates for future challenges? How well are investments building the capacity of engineering and technology faculty to facilitate and assess student learning and performance? Are engineering educators and educational researchers finding and properly using evaluation instruments to effectively improve and properly judge the impacts of educational innovations? What voids in evaluation instruments are preventing proper evaluation of innovations and merit further research and development investments? These and related questions drive the work reported in this paper.

In 2009, a team of engineering education and evaluation scholars began addressing questions of engineering educator evaluation capacity under an exploratory grant from the National Science Foundation's Course and Curriculum Development (CCLI) program[24]. The team conducted preliminary searches for evaluation instruments for engineering education, identified tentative groups of stakeholders for these instruments, and drafted criteria for cataloguing instruments for a future database of engineering education evaluation instruments. They convened a national panel of thirty engineering educators, evaluation experts, and developers of evaluation instruments in Arlington, Virginia to further their definition of the needs for an engineering education evaluation instrument database. This group of stakeholders helped the research team revise persona for stakeholder groups, clarify instrument characteristics important to different groups, identify additional sources of instruments, and reinforce the need for a widely accessible database of evaluation instruments that can support educators, evaluators, instrument developers, and funding program officers. This preliminary work and its results are reported in a recent conference paper[25].

To continue the earlier work, an expansion grant, "Appraisal System for Superior Engineering Education Evaluation-Instrument Sharing and Scholarship" (ASSESS), was funded in 2011 under the National Science Foundation's Transforming Undergraduate Education in Science, Technology, Engineering and Mathematics (TUES) program. This paper reports on the process and framework for creating ASSESS, a library of superior evaluation instruments that supports scholarly innovation in engineering education.

Vision, Goal and Objectives

The long-term vision for ASSESS is a proven web-based system that builds the evaluation capacity of the engineering education community. The short-term project vision is an effective, attractive ASSESS system ready for scholarly utilization by the engineering education community.

The overarching goal of the project is to build a robust, user-driven, web-based database of evaluation instrument information that collects information on high quality instruments and helps users select and properly use evaluation instruments to support scholarly engineering education project evaluation and implementation. This project builds on previous proof-of-concept work to create, thoroughly test, and strategically communicate the ASSESS system, positioning it for successful adoption and implementation by the engineering education community. Two sub-goals are:

1. *Instrument utilization*. Enable engineering educators to locate and effectively deploy superior evaluation instruments to enhance engineering education project discoveries and successes, and

2. *Instrument development*. Support the engineering education evaluation community in the identification and refinement of evaluation instruments with potential to become major assets to the profession.

Four project objectives define components toward achieving the project goals: a software support system, an instrument appraisal process, rigorous testing of the system, and targeted dissemination and adoption research. Objectives and observable outcomes are:

1. **Software Support System**. Develop an evaluation instrument database and user interface that support users in accessing desired instrument information and in providing feedback on instrument and system features.

   Achievement of this objective requires delivery of a prototype system that (a) has a web interface with appearance and functionality expected by the community, (b) enables users to easily locate desired resources, (c) instructs users in proper use and interpretation of instrument results, and (d) enables users to provide useful feedback on the instrument and system.

2. **Instrument Selection and Appraisal Process.** Establish an appraisal process that identifies, reviews, certifies, and supports refinement of superior evaluation instruments for engineering education.

   Achievement of this objective requires (a) a search strategy that effectively identifies potential evaluation tools for the database, (b) instrument appraisal properly overseen by engineering education evaluation experts, (c) characterization of evaluation instruments following appropriate evaluation standards, (d) selection of superior instruments through peer review by engineering education and evaluation professionals, and (e) identification of additional evaluation instruments with high potential for the community through unbiased vetting with the community.

3. **System and Process Testing.** Document the performance of the integrated software system and instrument appraisal process to ensure desired value added to the engineering education and evaluation community.

   Achievement of this objective requires that: (a) users report satisfaction with the overall web interface and system functionality, (b) characterizations of instruments in ASSESS are judged reliable by authors and users, (c) users report satisfaction with the feedback processes enabled by ASSESS, and (d) users report satisfaction with the process for nominating and critiquing prospective instruments for ASSESS.

4. **Dissemination and Adoption.** Conduct communication, vetting, and adoption research of the ASSESS system needed to establish its value and to define steps essential for broad adoption by the engineering education community.

   Achievement of this objective will be demonstrated by (a) a poster session at the ASEE annual conference drawing support from engineering educators, (b) an ASEE workshop providing evidence that new users of ASSESS can rapidly utilize and gain value from the system, (c) a review article in the *Journal of Engineering Education* drawing responses evidencing perceived value of ASSESS, and (d) adoption research findings aligning ASSESS characteristics with factors needed for adoption.

Methods

The ASSESS project, an NSF TUES type 2 project, builds upon the foundation laid by the precursor NSF CCLI type 1 project, (DUE 0839898) "Inventory of Evaluation Tools for Engineering Education Projects."[24] The precursor project identified a pool of individuals with varied backgrounds and project evaluation experiences as a resource for continued probing of interests and expectations of the engineering education evaluation community. In addition, the Inventory project (a) established the need for a readily available repository of high quality evaluation tools to support evaluation of engineering education projects, (b) identified some of the tools and characteristics that may be valuable for an evaluation tools database, and (c) identified potential user groups who could benefit from an engineering education evaluation tools database.

The core ASSESS project activities began with a meeting of five project consultants on January 5, 2011 on the campus of the project investigators. The purpose of the meeting was to review and revise the project team's understandings of the requirements for an evaluation tools database that can best support the engineering education community in building its evaluation capacity. The consultant group was purposefully comprised of individuals representing diverse background experiences in engineering education and evaluation. Consultants in attendance, three of whom participated in the precursor project, were:

- Gregory Chung – Senior Research Associate, National Center for Research on Evaluation, Standards and Student Testing (CRESST), University of California Los Angeles

- Edward Gehringer – Associate Professor, Computer Science, North Carolina State University
- Arlen Gullickson – Emeritus Researcher, former Director, The Evaluation Center, Western Michigan University; Chair of the Joint Committee on Standards for Educational Evaluation (JCSEE).
- PK Imbrie – Associate Professor, Engineering Education, Purdue University
- Thomas Litzinger – Professor, Mechanical Engineering, Pennsylvania State University; Director of the Leonhard Center for Enhancement of Engineering Education

Questions posed to the project consultants to guide their input to the project team addressed audience, instrument characteristics, search process, and feedback. Questions included:

1. Who is the audience we should be trying to attract to ASSESS?
2. What kind of instruments does the ASSESS infrastructure need to support?
3. How should the 'quality' of an instrument be defined?
4. What set of instruments (or instrument types) should initially be deployed in ASSESS?
5. What strategy should be deployed for identifying relevant measurement and evaluation instruments for Engineering Education?
6. What tagging scheme should be used for the instruments to aid in users' searches?
7. What interface requirements (or models) should be used to guide development of the ASSESS database interface?
8. What instrument user feedback should be supported in the system?

The consultants provided input to the project team in the form of recommendations resulting from their on-campus meeting and multiple follow-up communications that provided feedback on work of the project team. They reviewed plans for ASSESS database development, reviewed prototype website configurations, and provided reviews of evaluation instruments suitable for the ASSESS database.

The project team allocates project work to three work groups, as shown in Table 1, addressing the principal components of the ASSESS system. The engineering work group plans, designs, prototypes, and tests the database and the user interface for the database. They also develop algorithms for selecting desirable instruments based on their characteristics designated in the ASSESS database. The measurements work group identifies the characteristics of high quality evaluation instruments and ways of cataloguing and communicating these to users. They also identify evaluation instruments of interest to the community and prepare summaries of the instruments useful to user communities. The adoption work group conducts research to identify from potential users what they need, currently use, and would like to use to locate appropriate evaluation instruments for their applications. This group also directs feedback to the other work groups to guide their definitions and development of the ASSESS system to make it more adoptable. Activities of each work group are described below.

The project team used regular whole-team meetings to coordinate, plan, develop consensus, and obtain feedback on work of the separate groups. When developments reached a stage of relative stability, the consultants were contacted to obtain their review and suggestions.

Table 1. Project Work Groups and their Responsibilities

Group	**Responsibilities**
Engineering Work Group	Design and implementation of database, web interface, instrument search tools; testing and refinement of system
Measurements Work Group	Definition of instrument characterization structure, identification and characterization of candidate instruments
Adoption Work Group	Definition of ASSESS stakeholders and their needs; gathering feedback on ASSESS and informing refinements of the system

The engineering work group led development of the software support system for the IDEALS database. This group built upon the strawman evaluation tools database created in the precursor Inventory project and drew from other software tools with desired functionality to avoid unnecessary development work. The software development process followed the Rapid Application Development (RAD) methodology[26], which included three overall phases:

(a) Needs analysis produced software specifications to guide programming. The work group reviewed the Inventory project report, probed perceived needs of the entire project team, and gleaned suggestions from the consultants to compile anticipated needs of different stakeholders. The consultants reviewed the final list, added a few refinements, and encouraged the project team to conduct a broader user needs analysis. This will be done as part of the adoption research (see below).

(b) Integration and coding of database and user interface to define and demonstrate desired functionality. A database was prototyped based upon characterizations of instruments from the Inventory project report, and three example instruments were used to populate the database. This led to prototyping of a user interface that enabled users to review the contents of the database for individual instruments. Example reports pages were prototyped to show possible ways for instrument information to be presented to users.

(c) User testing (by the project team) to test the software and associated evaluation instrument data. The project team reviewed the prototyped user interface and the instrument information contained in the database. Feedback was used to make revisions to the website layout and contents of reports.

The measurements work group defined processes for identifying and rating the quality of evaluation instruments identified. The initial identification strategy incorporated three components: (1) search of peer-reviewed journal articles, (2) use of branching techniques from reference sections of papers and reports, and (3) identification of innovative projects that may have developed and employed evaluation instruments. Consultants added other important sources of instruments for the group to consider. Evaluation instruments in the ASSESS database will need to be of high quality, based on criteria set by the evaluation community for providing adequate information about the instruments and their proper use, as well as information about the reliability and validity of results expected from use of the instrument.

The strategy for users to search the ASSESS database to identify instruments of interest was co-led by the engineering work group and the measurements work group. The measurements group identified characteristics of greatest interest to users and the engineering work group identified algorithms for searching based on different types of data. Consensus definitions of a search strategy were prototyped on the ASSESS development website for the project team to try. After the project team was generally satisfied with the search strategy's implementation, the consultants were invited to review the search strategy implementation and to provide feedback. This led to additional refinements, as well as encouragement regarding the value of the ASSESS system.

Based on consultant feedback, the adoption work group will conduct a formal needs analysis for prospective ASSESS users. Surveys of potential users will address questions of two types to probe how users might use the system and what information they would seek from the system. Sample questions are listed below.

A. "Pull" Questions: What users would like in an evaluation instrument database

1. What instruments do you use in your classroom that might be broadly accessible for others wanting to conduct classroom assessments? Would you be willing to share these instruments as part of the tool database?
2. Where do you currently go to access these instruments?
3. What instruments would you like to see us include in the database?
4. What kind of information would you like to know about these instruments?

B. "Push" Questions: What users think about the ASSESS database after reviewing it

1. What information do you find useful?
2. What additional information would you like to see?
3. What suggestions do you have for the layout of the screen?
4. Would you use this tool? Why or why not?
5. Who would you envision using this tool?

The adoption work group planned adoption research to access perspectives and opinions of users and potential users of the ASSESS system that can be used to make the system more adoptable and broadly used. The combined "push" and "pull" approach described above was selected to obtain a broader perspective and richer data set to guide development efforts.

The "push" effort is aimed at existing and new users and their interpretations of the usability and adoptability of the ASSESS system. Adoptability will be investigated using the Diffusion of Innovations framework, specifically the five characteristics of innovations that have been shown to affect adoption: relative advantage, compatibility, complexity, trialability, and observability[27, 28]. A single group of individuals will be interviewed four times over the course of the project. Each time they will be presented with the current state of the ASSESS system and asked questions about how useable and adoptable the system is. This information will be summarized and provided to the development team after each set of interviews for improvement to the ASSESS system. Additionally, focus groups will be held at the 2012 ASEE and FIE national conferences to gather more perspectives.

The “pull” effort is aimed at how non-users of ASSESS access and assess the quality of assessment instruments. We will interview one repeat group of individuals (Group 1) and one changing group of individuals (Group 2: groups A, B, C, D) for this effort. The repeat group (Group 1) will provide data on the adoption impacts of ASSESS refinements, while the changing group will provide a means of validating and extending perspectives to additional prospective users. Group 1 will be interviewed each semester and findings will lead to a report. Group 1 and Group 2 will be interviewed in the same time periods, and these results will lead to additional reports. Table 2 shows the anticipated schedule for reports from the different interviews.

Table 2. Schedule for Reports from Adoption Research Interviews

Time of Report	Group 1 (Constant)	Group 2 (Groups A, B, C, D)
Fall 2011	“Push” interviews by phone	“Pull” interviews (A) by phone
Spring 2012	“Push” interviews by phone	“Pull” interviews (B) by phone
Summer 2012	“Push” interviews by phone	“Push” interviews (C) at FIE conference
Fall 2012	“Push” interviews by phone	“Push” interviews (D) at ASEE conference
Spring 2013	Summary report	Summary report

Results

Design of the ASSESS database will be guided by criteria established from data gained from project consultants, the precursor Inventory project panel, and input from ASSESS project leaders. The audience for the ASSESS website and database will be comprised of five stakeholder groups as defined in Table 3. These include project leaders, engineering educators, faculty developers, instrument developers, and program officers who will identify high quality evaluation instruments, deploy them in projects, or recommend them to others. These groups have overlapping but varied needs that will be addressed by the database and user interface.

Table 3. Stakeholders for ASSESS Website and Database

Stakeholder Group	Description
Project leaders	Leaders of educational innovation projects (e.g., NSF CCLI/TUES projects) who need to evaluate impacts of their projects to report their successes and to guide improvements by themselves and others
Engineering educators	Educators implementing innovations in their classes or programs (perhaps novices to experts in evaluation) who will benefit from using evaluation tools to determine impacts of their innovations
Faculty developers	Faculty development professionals (e.g., Teaching & Learning Center, Office of Faculty Development) who attempt to help their faculty learn how to evaluate learning and other impacts of what they do
Instrument developers	Engineering education or evaluation professionals developing new evaluation instruments (e.g., funded educational researchers) who seek to understand instrument quality and community needs to guide their development efforts
Program officers	Grant program officers (e.g., NSF TUES program managers) who desire high quality evaluation instruments that can be recommended to project directors for evaluating their innovations

The ASSESS database should include instrument types that are suitable for engineering education applications and that are easily characterized according to features or qualities of importance to stakeholders. Types of instruments identified for inclusion, summarized in Table 4, include tests of learning, self reports, and project leader data instruments. These address direct measures of project impacts as well as indirect measures, and cognitive measures as well as affective measures. Project leader data provides additional information useful to individuals who manage awards for engineering education innovation and research.

Table 4. Instrument Types for Inclusion in ASSESS

Instrument Type	Description
Tests of Learning	Achievement tests, concept inventories, process evaluations (e.g., teamwork), self-regulation, metacognition
Self Reports	Interest, affective, vocation, background, demographics, attitude inventories
Project Leader Data	Surveys for gathering data from project leaders

Additional specifications for the database have been identified from practical constraints and goals of the project. For example, the database should not (at least initially) seek to store assessment data, but only store evaluation instrument information. Only well-developed, high quality instruments should be included in the database. The database should be built to accept data from users, even though this capability may not be implemented initially; this capability will enable instrument developers to extend features of instruments in the future. User access should be role based: Users with different roles (e.g., occasional user of search feature, user conducting searches on a regular basis, instrument developer) should be given different privileges and have different responsibilities to support ASSESS. The system should have built-in data collection features (appropriate to user role and patterns of use) that identify who downloads information, at what campus, and to what effect. Further, the system should receive reviews by users, with reviews being rated by an algorithm (several are available) designed for this purpose. The evaluation instruments included in ASSESS should be rated for quality, based on evaluation protocols (e.g., NASA, USAID) made known to users.

Instruments to populate ASSESS should be identified following strategies that will locate high quality measurement and evaluation instruments suitable for use in Engineering Education. Archival journals and grant project databases should be searched for periods reaching from the present to more than ten years into the past. For collecting demographic types of data, national and international high stakes tests (e.g., NAEP, NSSE, HERI, PISA, TIMMS, SAT/GRE, ABET) are a good resource for these types of questions. Searching should include a complete search of archival journals and possibly dissertations and technical reports. Search terms should be broadened to include “validity” and “validation” in addition to “evaluation instruments” within the engineering education field.

Suitable evaluation instruments may be found in existing repositories and organizations involved in educational development. Examples include the FLAG website (http://www.flaguide.org/), the National Science Digital Library (http://nsdl.org), and the Engineering Pathways website (http://engineeringpathway.com). Evaluation study reports for grant programs (e.g., NSF CCLI) are another source that may identify instruments used in these studies. Program managers for

NSF and other grant programs are a resource for identifying instruments developed with funding from their programs or project PIs who are engaged in developing relevant instruments. Other possible sources for instrument information include: the Defense Technical Information Center (a repository for final reports of projects funded by Department of Defense), the Mental Measurements Yearbook (cataloguing 3500 commercially available tests), and regional educational laboratories (e.g., Northwest Regional Education Lab in Portland, OR).

Within the ASSESS database, instruments must be characterized to aid in identifying those meeting criteria set by the user – quality criteria, ranges of suitable application, and instrument features. A tagging scheme is preferred to multiple levels of characterization because tagging is more flexible and less constrained by needing to order the dimensions of the characterizations. Using multiple tags will enable labeling an instrument by type of instrument (e.g., affective, attitudinal, achievement), types of question (Likert, multiple choice, matching), compliance with standards (e.g., APA standards for Educational and Psychological testing), and many other dimensions independent of a hierarchy of these dimensions.

During the first year of the ASSESS project, the Engineering Work Group (EWG) developed a series of five prototype templates for characterizing evaluation instruments. Each template was designed to present important instrument information in a concise and useful form, prototyped on an ASSESS development website, and reviewed by others on the project team. With each new prototype, information was presented in a format that allowed the user to view additional information on demand. Each revision also responded to feedback from the project team and/or project consultants.

At the end of the first project year, the template for presenting assessment instrument information gives the user five tabs for viewing information in different ways: SUMMARY, DESCRIPTION, SPECS, REVIEWS, and LINKS. Figure 1 shows these tabs.

SUMMARY Tab. As shown in Figure 1, the Summary tab contains information presented under subheadings. The first, under the Summary subheading, provides an overview of the selected instrument, as interpreted by the ASSESS project team.

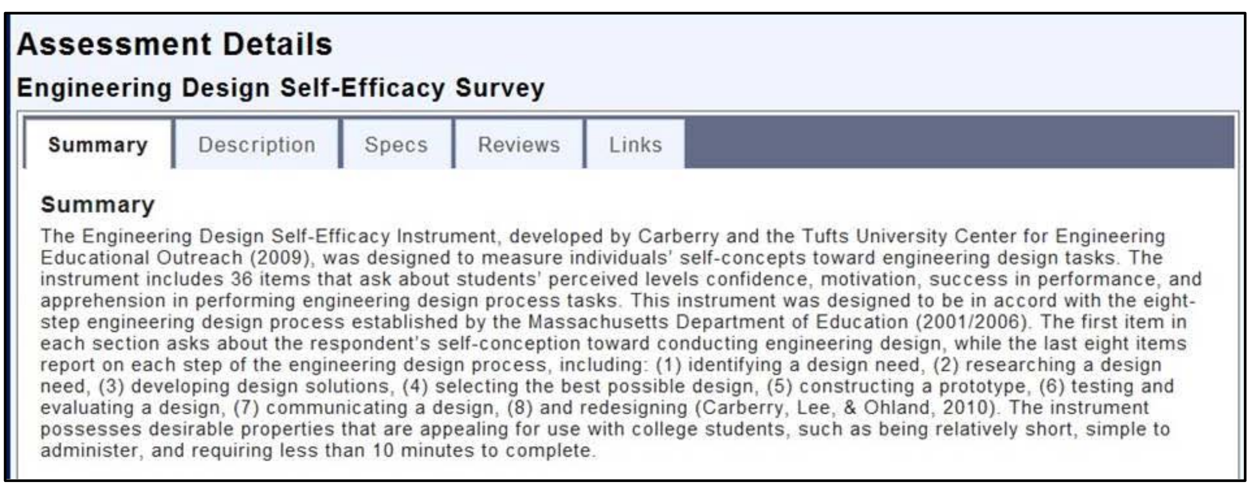

Figure 1. ASSESS Website Tabs and Summary (part 1) for a Sample Instrument

In the first summary paragraph (Summary subheading), the user finds information in narrative form: the instrument name and developers, purpose of the instrument, description of instrument items, and salient information about its usability. This summary provides the potential user a quick overview of the instrument that supports a decision to explore the instrument further.

Under other subheadings, the user finds additional information about the instrument's usefulness in engineering education, recommendations for its use, and reliability and validity summary statements, as shown in Figure 2. The Uses in Engineering Education subsection describes reported uses of the instrument in engineering education contexts, and cites references for additional information. From this, the reader can see how widespread is the instrument's use and where this may have occurred. The Recommendations subsection interprets this use and proposes more broadly the ways in which this instrument might be used appropriately in engineering education contexts.

Uses in Engineering Education

Because the Engineering Design Self-Efficacy Instrument was recently established, few studies have reported using the instrument to measure self-efficacy. The reference below is an ASEE conference proceeding describing the pilot study for the instrument.

Carberry, A., Ohland, M., & Lee, H-S. (2009). Developing an instrument to measure engineering design self-efficacy: A pilot study, American Society for Engineering Education, AC2009-206.

The pilot study is an investigation of how to develop an instrument that measures students' self-efficacy regarding engineering design. 36 items were developed and tested using three types of validity evidence. First, the content of the instrument was tested to ensure that the full domain (each subdimension) of the engineering design process was represented. Second, the instrument was tested for whether responses to the instrument could identify groups with various levels of engineering design experience. Finally, theoretical connections between motivation, expectancy for success, and anxiety were tested to determine their appropriateness in the measurement of self-efficacy. Results confirmed an accurate reading of engineering design self-efficacy for 82 volunteer respondents with diverse engineering expertise.

Continuation of Summary

Recommendations

The use of the instrument seems appropriate for examining students' beliefs about their engineering design skills and knowledge of the process. Scores on the Engineering Design Self-Efficacy Instrument could be used as a predictor variable or as an outcome variable depending on the research questions being asked. This scale could be used in local evaluations of programs as well as wide scale research efforts with engineering students.

Reliability

The reliability information for the scores from the Engineering Design Self-Efficacy Instrument has been in the form of internal consistency. The following Cronbach's alpha values were found for the task-specific self-concepts: self-efficacy (0.967), motivation (0.955), outcome expectancy (0.967), and anxiety (0.940) (Carberry, Lee, & Ohland, 2010).

Validity

Content validity, criterion-related validity, and construct validity were established for the Engineering Design Self-Efficacy Instrument (see Carberry, Lee, & Ohland, 2010). There were significant differences in scores on the scale between low, medium and high experience students. Associations with other variables validity evidence was provided with motivation(.78) and outcome expectancy (.92) and anxiety(-.60).

Figure 2. Narrative Summary of a Sample Instrument

Under the Reliability subsection is a description of evidence for the instrument's reliability. As appropriate, this paragraph may identify specific types of reliability evidence available and ranges of reliability reported for different measures. The presence of an icon by the Reliability subheading provides a popup window with an explanation of reliability rating standards used.

Similarly, the Validity subsection provides a brief description of evidence for validity of the instrument. This paragraph may identify the types of validity for which evidence is available and the ranges of validity reported for different measures in contexts for which validity was evidenced.

DESCRIPTION Tab. Under the Instrument Description tab are details on the instrument (many of which are included in the summary), presented here in tabular format for the purpose of making comparisons among instruments. These data are often descriptive terms, numbers, or items selected from a list of options. Items in the description include the following: instrument name, instrument purpose, constructs or outcomes evaluated, authors, year published, publisher, intended audience, administration time, cost, type of instrument, and many other descriptive details. These items of information relate to the instrument's usefulness.

SPECS Tab. Under the Technical Specifications Tab are tabular data related to the instrument's technical credibility. Included here are indications of reliability and validity as well as technical and utility standards addressed by the instrument. Information provided includes: versions of the instrument, languages in which it is available, availability of a user's manual, evidences of reliability (e.g., internal consistency, test-retest, generalizability theory, alternate form, inter-rater), evidences of validity (e.g., content, internal structure, associations with other variables, predictive), utility standards of program evaluation (e.g., stakeholder identification, information scope and selection), accuracy standards of evaluation (e.g., valid information, reliable information, analysis of quantitative information, analysis of qualitative information), and any ABET criteria addressed.

REVIEWS Tab and LINKS Tab. The Reviews tab is planned to display external reviews of the evaluation instrument for others to read. This feature has not been developed to date. Similarly, the Links tab will contain hyperlinks to sites at which the evaluation instrument or articles about the instrument may be found.

Software System Development

The ASSESS software development has advanced through early stages of the Rapid Application Development (RAD) process needed to prototype components of the evaluation instrument database and the web-based user interface. The Engineering Work Group has largely completed (a) needs analysis that produced software specifications to guide programming, and (b) coding (separate coding of database and user interface), integrating parts, and defining a testing plan for the system. The software has passed through five different prototypes, each refined by feedback from the ASSESS project team and/or project consultants. Continued refinement of specifications for the system (and subsequent software revisions) will occur as a formal needs analysis is conducted as part of the adoption research for this project.

The current website prototype is focused mainly on reports for viewing details of different instruments and a search page used to find desired instruments. These functions were deemed critical to the usability and ultimate acceptance of the website. These features also drive the structure of the database as the information to be viewed defines the data that must be captured.

The focal point of the current prototype is the search page, shown in Figure 3. The search features currently available include a two-level search: (1) keyword search and (2) filtering by fields of information that characterize the instrument.

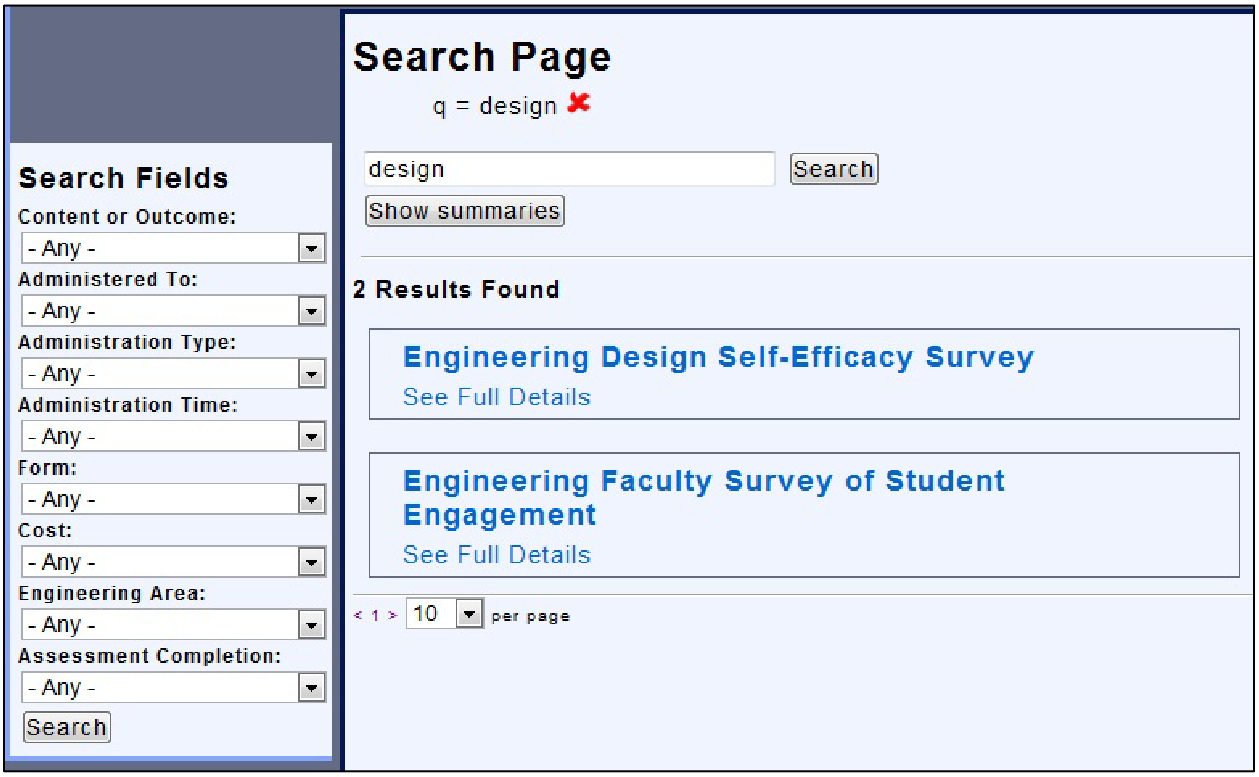

Figure 3. ASSESS Database Search Page

The keyword search shown at the top of the page allows users to search through the names and summaries of instruments in the database based on a specified word or phrase. Any instruments containing the search terms are displayed below the keyword search area. As shown, two instruments in the development database contained the word “design” in the content searched.

On the left of the search page is a panel of drop-down boxes to filter instruments by specific criteria. One may search on information stored in database fields such as content or outcome, administered to, administration type, or engineering area. For example, one may choose “undergraduate” as the criterion under “administered to” and the search drops one of the two instruments in the list shown. When the search on criteria is conducted (by clicking the corresponding “search” button), instruments identified by the keyword search are filtered according to the criteria specified. These filter terms were refined through input and critique by consultants and experts on the development team.

Search results are displayed in the center of the search page. Details displayed can be controlled by the user. By default, the instruments are displayed by their instrument name. The “show summaries” button causes the one-paragraph narrative Summary to be displayed below the instrument name. In addition, the “show details” button toggles to show all of the Summary tab information below the instrument title.

To aid in documenting the search history, a 'bread crumb' of the search keywords and filters used in the current search appears at the top of the page. As shown in Figure 3, the keywords appear first in the list, followed by filter terms and parameters specified. The search may be altered by mouse clicking the red “X” by a keyword or filter term to remove it. Additional search items may be specified by adding them to the appropriate keyword list input box or filter pull-down options.

Simple search algorithms are used for early stages of ASSESS system development, while search terms are being evaluated. Currently the keyword search looks for the exact phrase in the title or summary. Future versions will recognize multiple separate words and phrases surrounded in quotes as a single search parameter. The drop-down field searches currently filter for a very limited set of options within a field; these options will undergo continued refinement to match user expectations. Future plans include adding intelligence to the searches that will allow ‘novice’ terminology to be equated with the terminology used by professionals in the evaluation field.

Instrument Selection Process

The selection process for instruments to be included in the ASSESS database allows different paths of instrument nomination but a common review process for approval. Nomination of instruments will begin through an initial search of archival literature, grant databases, government and regional repositories for educational research, and websites for educational resources. In addition, individuals may nominate (or self-nominate) instruments felt worthy of inclusion in ASSESS.

Criteria for inclusion will be established and published by the ASSESS Team with input from project consultants and interviewees in adoption studies. These criteria will include: importance to engineering education, stage of instrument development, evidence for validity and reliability, and usability (including quality of documentation). Rubrics will be developed to define acceptable levels for each of the criteria. Publication of these criteria will serve as a guide for instrument developers and for nominators to locate and promote instruments needed by the engineering education community.

Reviews of candidate instruments will be conducted by the Measurements Work Group, led by Drs. Brian French and Michael Trevisan of the Learning and Performance Research Center, Washington State University. Both hold Ph.D.'s in educational psychology – measurement, have reviewed numerous tests and measures for the Buros Institute of Mental Measurements at the University of Nebraska, and have engaged in several engineering education research projects. Using test industry standards and an eye toward user needs, this group will determine which instruments meet the criteria for inclusion. For example, when reviewing reliability of an instrument, Table 5 can be used to judge its qualifications regarding reliability criteria. This group will also communicate with instrument owners to announce decisions and to encourage instrument development.

Table 5. Reliability Standards for Evaluation Instruments (Source: Nunnally & Bernstein[29])

Reliability Rating	Reliability Values	Score Use
Low Reliability	$r < 0.70$	Early stages of construct validation work
	$0.70 < r < 0.80$	Basic research, decisions about groups
	$0.80 < r < 0.90$	
	$0.90 < r < 0.95$	Decisions about individuals-low stakes
High Reliability	$r > 0.95$	Decisions about individuals-high stakes

The Measurements Work Group will prepare an instrument review for each instrument contained in the ASSESS database. Their reviews will be written in an accessible manner and style that provides the developer information about deficiencies to be addressed before inclusion and offers the user adequate general and technical information to determine instrument suitability for an intended application.

Outreach Activities

In October 2011, the project team presented a work-in-progress paper at the Frontiers in Education Conference in Rapid City, South Dakota. At this presentation, the audience was made aware of the need for capacity building in educational evaluation and was informed about the ASSESS project's goals and achievements to date. Audience members were invited to contribute to this development.

Next Steps

The ASSESS project has laid a foundation for establishing a robust, user-friendly website for superior evaluation instruments for engineering education projects and innovations. The project team has plans for adoption research, software development and testing, and populating the database.

Adoption Research. Conduct adoption research with prospective users to determine the attractiveness of the envisioned ASSESS resource and actions required for its broad utilization. Steps planned include:

- Interview individuals who have some familiarity with the project – participants in the precursor Inventory of Evaluation Instruments workshop
- Interview others who might be potential adopters – at engineering education conferences
- Apply diffusion of innovations theory to develop a model that describes adoption for ASSESS
- Inform ASSESS developers of needs that require refinements in the ASSESS system
- Continue interviews to obtain feedback on ASSESS website and database refinements and additional refinements needed for adoption

Software Development. Develop the website and database to achieve greater functionality and user satisfaction, including:

- Add other prototypical instruments to the database to span the types of instruments that are of interest to our customers. This will allow ASSESS developers to exercise the full range of the database structure.
- Refine the instrument data model based on experience with more data and feedback from users. More fields may be needed to describe a broader range of characteristics, and similar fields may be merged to simplify searching.
- Build administrative pages to support instrument entry, review and editing functions.
- Add home page and other informative pages to better inform users about the site's purposes and functionality and to guide users in proper operation of the site.
- Build a system for registering users with the site, logging in and out, and managing user roles. Some of the roles being considered include tool users (who predominantly use the search features), tool developers (who may edit instrument data), visitors, and administrators.

Database Expansion. Add high quality evaluation instruments of varied instrument types to the ASSESS database.

- Conduct literature and report reviews to identify suitable instruments.
- Develop and validate with stakeholders the criteria for inclusion of evaluation instruments in the ASSESS database
- Screen candidate evaluation instruments using established criteria
- Prepare reviews needed for evaluation instruments
- Enter into ASSESS the instrument review and other information needed for the instrument.

Opportunities for Collaboration

The ASSESS project can benefit from collaborations with evaluation instrument users and developers to guide the project's activities in producing resources valuable to the engineering education community. Persons aware of other developments or resources available to aid in the current development of ASSESS should contact one of the project team's members. Everyone will benefit from a high quality, robust, user-friendly database for evaluating the impact of engineering education programs and innovations.

Acknowledgements
The authors gratefully acknowledge support for this work through the National Science Foundation, Division of Undergraduate Education, grant number DUE 1065486.

Bibliography

1. Boeing Company Desired Attributes of an Engineer. http://www.boeing.com/educationrelations/attributes.html
2. National Academy of Engineering, *The Engineer of 2020: Visions of Engineering in the New Century*. The National Academies Press: Washington, DC, 2004.
3. National Academy of Engineering Grand Challenges for Engineering. http://www.engineeringchallenges.org/ (November 15).

4. Sheppard, S.; Colby, A.; Macatangay, K.; Sullivan, W., What is Engineering Practice? *International Journal of Engineering Education* **2006,** 22, (3), 429-438.
5. Tadmor, Z., Redefining engineering discipline for the twenty-first century. *The BRIDGE* **2006,** 36, (2), 5.
6. Vest, C. M., Context and challenge for twenty-first century engineering education. *Journal of Engineering Education* **2008,** 97, (3), 235-236.
7. ASEE *Engineering education for the global economy: Research, innovation, and practice*; Washington, DC, 2008.
8. Bransford, J., Preparing people for rapidly changing environments. *Journal of Engineering Education* **2007,** 96, (1), 1-3.
9. Committee on Maximizing the Potential of Women in Academic Science and Engineering *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*; 2007.
10. Lattuca, L. R.; Terenzini, P. T.; Volkwein, J. F.; Peterson, G. D., The Changing Face of Engineering Education. *The BRIDGE* **2006,** 36, (2), 9.
11. National Academy of Engineering, *Educating the Engineer of 2020: Adapting Engineering Education to the New Century*. The National Academies Press: Washington, DC, 2005.
12. Sheppard, S. D.; Macatangay, K.; Colby, A.; Sullivan, W. M., *Educating Engineers: Designing for the Future of the Field*. Jossey-Bass: San Francisco, CA, 2008.
13. Vest, C., Educating engineers for 2020 and beyond. *The BRIDGE* **2006,** 36, (2), 7.
14. National Research Council, *How People Learn: Bridging Research and Practice*. National Academy Press: Washington, DC, 1999.
15. Newstetter, W., Designing cognitive apprenticeships for Biomedical Engineering. *Journal of Engineering Education* **2005,** 94, (1), 207-213.
16. Pellegrino, J. W. *Rethinking and redesigning curriculum, instruction and assessment: What contemporary research and theory suggests*; 2006.
17. Rayne, K.; Martin, T.; Brophy, S.; Kemp, N. J.; Hart, J. D.; Diller, K. R., The development of adaptive expertise in biomedical engineering ethics. *Journal of Engineering Education* **2006,** 95, (2), 165-173.
18. Redish, E. F.; Smith, K. A., Looking beyond content: Skill development for engineers. *Journal of Engineering Education* **2008,** 97, (3), 295-307.
19. Schon, D. A., Knowing-in-Action: The new scholarship requires a new epistemology. *Change* **1995**, (November-December), 27-34.
20. Smith, K. A.; Sheppard, S. D.; Johnson, D. W.; Johnson, R. T., Pedagogies of engagement: classroom-based practices. *Journal of Engineering Education* **2005,** 94, (1), 87–101.
21. Svinicki, M. D., *Learning and Motivation in the Postsecondary Classroom*. Anker Publishing: San Francisco, CA, 2004.
22. National Science Foundation In *Project Impact: Disseminating Innovation in Undergraduate Education*, National Science Foundation, Arlington, Virginia, May 31-June 3, 1994, 1994; McGraw-Hill: Arlington, Virginia, 1994; p 116.
23. National Science Foundation *Evaluation of the Division of Undergraduate Education's Course and Curriculum Development Program* Arlington, VA, 1998.
24. Davis, D. C.; Trevisan, M. S.; Zhang, M.; Lebeau, J.; Fairbrother, D. *Inventory of Evaluation Tools for Engineering Education Projects*; Washington State University: Pullman, WA, September, 2009.
25. Trevisan, M. S.; Davis, D.; LeBeau, J. E.; Zhang, M. In *Work-in-Progress -- The Evaluation Tools Database for Assessing Engineering Education Innovations*, Frontiers in Education Conference, Rapid City, SD, 2011; Rapid City, SD, 2011.
26. Martin, J., *Rapid Application Development*. MacMillan College Division: 1991.
27. Rogers, E. M., *Diffusion of Innovations*. 5th Edition ed.; Free Press: 2003.
28. Sahin, I., Detailed review of Roger's diffusion of innovations theory and educational technology-related studies based on Roger's Theory. *Turkish Online Journal of Educational Technology* **2006,** 5, (2).
29. Nunnally, J.; Bernstein, I., *Psychometric Theory*. Third Edition; McGraw-Hill Humanities/Social Sciences/Languages: 1994.