

# Evaluating What Students Know: Using the RosE Portfolio System for Institutional and Program Outcomes Assessment Tutorial

—Feature by

JULIA M. WILLIAMS, SENIOR MEMBER, IEEE

**Abstract**—Currently, colleges and universities have developed assessment systems that can collect student work products for evaluation in an effort to make student learning transparent and ensure accountability in higher education. At the Rose-Hulman Institute of Technology, we have developed a digital portfolio system, the RosE Portfolio System (REPS), that allows for efficient data collection; the results of portfolio evaluations are used by academic departments and programs to improve curriculum and provide evidence to external accrediting agencies. The results of evaluations of student performance are also used to ensure the quality of academic curricula.

**Index Terms**—Accreditation, assessment, electronic portfolios, evaluation.

In the realm of higher education, faculty and administrators are searching for tools that can help them assess and evaluate their students' achievement of defined learning outcomes in fields as diverse as engineering, business, health professions, math, science, and technology (to name a few). These assessments and evaluations are part of a national trend toward transparency and accountability regarding the value added in education. Perhaps the most notable of these trends came in September 2005 with the announcement of the Secretary of Education's Commission on the Future of Higher Education. Margaret Spellings, then Secretary of Education in the Bush administration, charged the commission to develop a "comprehensive national strategy for postsecondary education" that would "meet the needs of America's diverse population and also address the economic and workforce needs of the country's future" [1]. The noble goal of the work was, however, undercut by what some educators saw as a potential threat; in her remarks at the press conference announcing the commission, Spellings stated

President Bush has proposed a plan to extend the benefits of high standards and accountability to our high schools. And we must act on it. Thanks to "No Child Left

Behind" [NCLB], we've already seen what a difference these principles have made for our younger students. [2]

When the commission made its final report in 2006 [3], the notion of NCLB at the college level had disappeared from the list of recommendations, perhaps shouted down by vocal critics like Brian Huot, who criticized the central premise of the project [4]. In the years since that report, however, a number of sources have emerged that offer information on costs, tuition, and students' self-reports of achievement of learning outcomes, such as the University and College Accountability Network (U-CAN) and the National Survey of Student Engagement. It is interesting, however, that these sources do not provide information if the college or university decides not to share it [5]–[7]. In addition, the only data provided on achievement of student learning outcomes comes from self-report surveys.

While the calls for accountability have reached a crescendo during the past several years, they are not new, particularly in the field of engineering education accreditation. Beginning in the 1980s, engineering educators responded to the call from industry for better-prepared students. In addition to asking for students who were well-prepared to solve problems and perform engineering design, industry also wanted students who could communicate effectively, work on cross-disciplinary teams, and demonstrate an awareness of global cultures [8]. This multifaceted call was translated into a focus on outcomes-based assessment and codified into the Engineering Criteria, a set of defined student learning outcomes used by the Accreditation Board for Engineering Technology Inc. (ABET), in conjunction with the American Society for

Manuscript received December 01, 2008; revised March 23, 2009.

The author is with the Office of Institutional Research, Planning and Assessment, Rose-Hulman Institute of Technology, Terre Haute, IN 47803 USA (email: Williams@rose-hulman.edu). Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

IEEE 10.1109/TPC.2009.2038737

Engineering Education (ASEE) and the Institute of Electrical and Electronics Engineers (IEEE) to measure US engineering programs [9]. This radical shift in focus—from course accounting to outcomes assessment—produced another radical shift. Engineering faculty would now need to document student learning beyond simply reporting course grades. They would need to define outcomes and assess student achievement, producing results that could then be used to improve curricula and pedagogy. And these results would ultimately need to convince accreditors of the quality of the engineering programs themselves.

The Rose-Hulman Institute of Technology responded to accreditation demands early in the 1990s by developing an assessment and evaluation process using online tools for data collection and evaluation. The RosE Portfolio System (REPS) was first used in 1998 to collect students' work products that were then evaluated against a set of established rubrics by teams of faculty evaluators. In 2008, a revised system, the RosEvaluation Tool (RET), designed with the same assessment approach, was used to evaluate students' work products within the campus learning-management system. The decade-long project continues to be used for institutional and program assessments, and it forms the bedrock of our preparations for program and institutional accreditation requirements. This paper presents an overview of the REPS, along with a detailed explanation of the RET, the online tool developed within the course-management system that provides faculty and administrators with the ability to collect, assess, and report on students' performances against a set of learning outcomes. This paper also presents information regarding how the data are used in conjunction with other assessment information and explains how these assessment practices have impacted faculty.

## BACKGROUND

The Rose-Hulman Institute of Technology is a private, undergraduate college of approximately 1,900 students located in Terre Haute, Indiana. Its emphasis is on educating undergraduates to pursue careers in the fields of mathematics, engineering, and science. It has a strong track record of creatively developing and rigorously assessing pedagogies for teaching in these fields. For example, it was the innovator of the Integrated First-Year Curriculum for Science, Engineering, and Mathematics, a curriculum designed to help students understand unifying ideas across seemingly disparate technical

disciplines; the Integrated First-Year Curriculum led to invited participation in the National Science Foundation-sponsored "Foundation Coalition," a nationwide coalition of schools applying current learning theories to revitalize fundamental engineering courses.

In addition to curricular innovations, Rose-Hulman has led the field of engineering, mathematics, and science education in the use of technology in the classroom. It was among the first colleges to require the use of laptop computers (beginning in 1995), and it was one of the first campuses to use Maple (a computer algebra system) in all first-year calculus classes. It continues to produce new technology-enabled "studio" courses (in, for example, physics and electrical engineering) that link hands-on learning in laboratory sessions with theories and concepts from traditional lectures. Since 2003, it has been implementing tablet PCs in engineering, science, mathematics, and humanities courses, and it has continued to rigorously assess these efforts at pedagogical innovation. For these and other education innovations, Rose-Hulman has been ranked first by engineering educators as the nation's best college or university that offers the bachelor's or master's degree as its highest degree in engineering for the eleventh-straight year; this ranking is published in the annual edition of "America's Best Colleges" guidebook by *U.S. News & World Report*.

By combining a tradition of curricular development with a dedication to the use of technology to enhance education, we at Rose-Hulman began in 1997 to develop an institute-wide assessment process. The centerpiece of the project included developing a defined set of institutional learning outcomes and the Rose-Hulman electronic portfolio project, the REPS. We initiated the process by discussing the various approaches to learning-outcomes development that were available, such as Bloom's Taxonomy and Gagne's Outcomes of Learning [10], [11]. From these approaches, we developed a set of institute-wide student learning outcomes—outcomes that would constitute the set of skills all Rose-Hulman students develop by the time of graduation. These outcomes were designed based on input from a wide variety of constituents: faculty, alumni, industry (those who hire our graduates), graduate schools, and other sources. By the end of the 1997–1998 academic year, we had a set of ten Institute Student Learning Outcomes. These ten learning outcomes were adopted by the faculty of the institute and subsequently published

TABLE I  
INSTITUTE DOMAINS AND STUDENT LEARNING OUTCOMES

Technical Knowledge Domain	<ul style="list-style-type: none"> <li>• Student learning outcomes in this domain are defined by each department as is appropriate to its majors and disciplines.</li> </ul>
Professional Skills Domain	<ul style="list-style-type: none"> <li>• Criterion 1. Leadership: Successfully motivating and enabling a group towards the achievement of a shared, articulated goal.</li> <li>• Criterion 2. Teamwork: Requires cooperative effort toward a common goal wherein each individual contributes in a particular role.</li> <li>• Criterion 3. Communication: Regardless of the media, requires unique skills whether communicating with individuals or with groups.</li> </ul>
World Citizenship Domain	<ul style="list-style-type: none"> <li>• Criterion 4. Cultural and Global Awareness: Requires perception and understanding of the cultural perspectives and social systems that define human communities.</li> <li>• Criterion 5. Ethics: Requires the use of appropriate moral theories, ethical principles, or professional standards to weigh alternatives and determine a proper, professional course of action.</li> <li>• Criterion 6. Service: Requires the use of one's time and skills to benefit an individual or community without cost to the recipient.</li> </ul>

in Rose-Hulman official documents, such as our course catalog and webpages.

The ten Institute Student Learning Outcomes were used for two ABET/Engineering Accrediting Council cycles for program accreditation, first in 2000 and again in 2006. After the 2006 accreditation cycle, when all Rose-Hulman engineering and computer-science programs were accredited, we conducted a review and revision of the ten learning outcomes. The review process resulted in a new set of outcomes organized into three domains. (See Table I.)

A complete list of the outcomes, performance criteria, and evaluation rubrics are available at the RosE Portfolio website (<http://www.rose-hulman.edu/REPS>).

## ASSESSMENT APPROACH FOR THE ROSE PORTFOLIO SYSTEM

The task of preparing for ABET accreditation includes documenting many different facets of the program and the institution, from total square

footage of laboratories and classrooms to faculty salary and experience. For the purpose of this article, however, the focus is on documenting students' achievement of learning outcomes, considered central to the program's success and sometimes the most difficult component of a program for faculty to demonstrate. In the past, ABET accreditors—the peer reviewers who examine the academic programs' self-study reports and go to campuses in teams for site visits—only counted the number of courses listed for each major; these counts were then verified to determine that a cross-section of students had taken the courses (through transcript analysis). The measure of student learning was merely in confirming that students took required courses and passed them. The move to outcomes assessment meant that programs could not demonstrate student learning as they had in the past; instead, they were required to use alternative assessment tools, in addition to transcripts and grades (often referred to as a focus on inputs), to prove that students could do what the program claimed they could do (often referred to as a focus on outputs or outcomes). To document achievement of student learning outcomes, engineering and computer-science programs across the US have tested and adopted a variety of assessment approaches to document outcomes.

A recent publication from the Association for Institutional Research presented a special-focus volume on evolving best practices in assessment for engineering programs [12]. The articles demonstrate how much variety has developed within assessment practices. Perhaps the most widely used assessment tool is the survey. With this tool, students may be asked to self-report on their growth and development during a course, a cooperative experience, or a project [13]. Student self-reports may then be compared to surveys of the course instructor, of co-op employers, or the design faculty team. A second frequently used assessment tool is the mapped exam question, what Estes et al. refer to as an “embedded indicator” [14, p. 135]. With this tool, for example, individual exam questions are mapped to specific learning outcomes. Students' scores on the questions are then compiled and compared to other embedded indicators, to surveys, and/or to course grades. Different from surveys, these embedded indicators are “direct measures of student performance based on assignments that are already in the curriculum” [14, p. 135]. While the tools described here offer efficient data collection of evidence of student

outcomes, they still present limitations. Students' self-reports must rely on their self-perception of their learning and growth. Mapping with more objective evaluators, such as faculty and employers, provides some rigor, but the limitation remains. Faculty who teach a student in class may have particular biases. Employers are often too busy to complete a lengthy survey, and small companies may encounter problems with maintaining employee confidentiality. Embedded indicators also contain a degree of bias. A professor teaching a course assigns a student a course grade. While the final grade in the course may not indicate exactly what the student can do, an embedded indicator is also evaluated by the faculty member or a team of faculty who teach the course. In both cases, however, the information provided to programs by ABET regarding how they should document achievement of student learning outcomes is left purposefully vague:

Explain the assessment and evaluation processes that periodically document and demonstrate the degree to which the program outcomes are attained. Describe the level of achievement of each program outcome. Discuss what evidence will be provided to the evaluation team that supports the levels of achievement of each program outcome. [15, p. 7]


Thus, each program can determine what assessment tool will be used and how the assessment results will be evaluated.

In the context of Rose-Hulman, we felt strongly that we needed to know more about our students' abilities than a survey could tell us. Instead, we selected an assessment tool that would allow for the direct assessment of authentic student work; this work would be judged by faculty evaluators who did not teach the course in which the work was completed. We knew that each engineering program and the computer-science program would need to document student learning. In addition, we are accredited as an institution by the North Central Association of the Higher Learning Commission, which also requires that we demonstrate achievement in student learning. We believed we could leverage the demands for program and institutional accreditation if we designed institutional outcomes in a way that could efficiently map to program outcomes. For example, ABET-accredited programs must show that students can demonstrate communication skills (only one of the 11 outcomes specified by ABET). By defining a communication outcome

for Rose-Hulman, we gained cooperation from all of our programs; they agreed to use the data-collection method (the REPS) and the portfolio-rating results in their own self-study reports for submission to their accrediting boards. The REPS is the data-collection and assessment mechanism for Rose-Hulman's six Institute Student Learning Outcomes: (1) leadership, (2) teams, (3) communication, (4) ethics, (5) cultural and global awareness, and (6) service. Each program defines technical knowledge as it is appropriate for its own majors, and each program also assesses and evaluates student learning for these outcomes. Many of these programs also use the REPS to assess technical outcomes because the process is efficient and produces useful results.

## DECISION TO DEVELOP A PORTFOLIO SYSTEM

Like other engineering and computer-science programs, we needed to determine proper assessment tools that would allow us to document achievement of student learning outcomes for the purpose of program and institutional accreditation. Our consideration of assessment tools initially covered a variety of alternatives, including surveys and embedded indicators. The decision ultimately to adopt a portfolio approach, with the subsequent effort to build an electronic portfolio, evolved from research we conducted on the uses of portfolios in fields other than engineering education.

As our initial research made clear, there were several definitions of "portfolio"—either as a  copy or electronic—in current use. Originally, a reference to a device for carrying leaves or sheets of paper (from the Italian "portafoglio"), the term "portfolio" has come to mean a selection of a student's work that is collected over a period of time and is often judged to determine a student's performance or progress. The image of the student portfolio seems characteristic of several fields, such as art or architecture. We can envision an art student who brings his or her portfolio to an interview for a graphic artist position or an architecture student who displays his or her portfolio for prospective clients. The portfolio has been incorporated into other academic fields as well. As early as the 1970s, portfolios were used to collect samples of student writing in English composition programs, and today we can see portfolios being used to assess competencies in such diverse fields as nursing, business, general education, and even engineering [16]–[18]. This

variety of applications of portfolios has produced a shift in the definition of a portfolio; the portfolio is a

purposeful collection of a student's work that exhibits the student's efforts, progress, and achievements. The collection must include student participation in selecting contents, the criteria for selection, the criteria for judging merit, and evidence of student reflection. [19, p. 60]

Even this definition suggests that there is only one kind of portfolio, and anything that claims to be a "portfolio" but doesn't include all of the elements listed above isn't a portfolio. Actually, as the use of portfolios expands and different fields adopt it as a data-collection method, the very nature of portfolios has changed. The approach employed in a portfolio depends on the primary purpose of assessment [20], [21]. If the primary goal of the portfolio is to measure an individual student or learner's performance improvement over time, a "growth-model" portfolio is adopted. The growth-model portfolio contains work samples collected at different stages of the learning process to illustrate differences in performance. For instance, a first-year student may have in her portfolio samples of calculus problems that she worked during her freshman year. By the end of her college career, she can include problem sets from advanced mathematics courses. A comparison of her work as a first-year student versus her performance as a senior will show her advancement within her chosen major.

On the other hand, if the primary goal of the portfolio is to measure competencies against certain standards or performance criteria, a "best-work" or "showcase" model captures work samples that demonstrate the achievement or fulfillment of the stated criteria. In this model, a student's portfolio can include her most exemplary work in a senior design project and an essay from an upper-level history course that has been written and revised several times. In the case of the best-work/showcase portfolio, the student is able to highlight her best work for the purpose of meeting program outcomes, for the purpose of career development, or for some combination of the two. The portfolio may contain reflective statements in which the student describes the significance of the work, what she learned, and so on.

In addition to the two models listed, we can further distinguish portfolios by the purpose a portfolio serves. Student or learner portfolios are typically

used to collect evidence of student achievements in the classrooms and/or throughout an academic career. As mentioned before, the growth-model approach most closely maps to the student or learner portfolio type. In contrast, colleges and universities use institutional portfolios in a manner very different from the growth-model portfolio type. While the content of the portfolio may indeed be student work, the student may or may not have had an input on the choice of the work that is included. Students also may have no role in the evaluation of the work, and they may not be required to include reflection on their learning as part of the portfolio.

Institutional portfolios have a very different purpose. Faculty and staff in higher education use them to communicate with accrediting agencies and with broader audiences when addressing topics such as institutional effectiveness and accountability. When students' institutional portfolios are evaluated, they are generally assessed on the basis of whether the students' work meets institutionally defined learning outcomes. The institutional portfolio is the type that was eventually selected by our institution. It is also used at the Colorado Schools of Mines [22].

Whether the approach is growth or best work, whether the type is student or institutional, the work that is gathered into an assessment portfolio will serve its purpose best if it is associated or mapped to one or more learning outcomes, or—even better—to specific performance criteria. The mapped work can then be assessed by using a set of predefined rubrics. The work of mapping portfolio contents is best done before the work is collected rather than after the fact. Mapping can also make evaluation much more manageable. For our institutional context, we were initially unsure if students were getting adequate opportunities to develop their skills in all of the outcome areas. For that reason, all departments on our campus create curriculum maps; the maps identify where students are given the opportunity to develop skills in each outcome area. These maps are created annually. Faculty members work within their departments to identify the student learning outcomes that are important to their disciplines and to their accrediting agencies. When programs appear to have gaps in curricula, a lack of opportunity for students to develop their skills, then curriculum changes can be made by identifying faculty who will provide these opportunities in their courses and require that students complete assignments that can serve as evidence of outcomes achievement. This information was especially important early

in our assessment-process development, since it highlighted that ethics had been added as an institutional outcome (and an ABET outcome as well), but in many curricula, students were not given the opportunity to develop their skills. In response, faculty members in all of our engineering programs began to add ethics modules to specific courses to provide students this opportunity.

In 1997, when work on this project began, we recognized that we needed to develop an effective and efficient data-collection method; portfolios were one method among many that we considered. As we debated portfolio adoption, we realized that we did not want to rely on a hard-copy portfolio. We decided to develop an electronic portfolio because, in 1995, Rose-Hulman had initiated a laptop computer requirement for all students (making it one of the first colleges to do so). Thus, all students used an institute-specified laptop computer with a preinstalled software suite. We believed we could make the portfolio-assessment process effective and efficient if all dimensions of the process—from student submission to portfolio evaluation—occurred within an electronic system. At that time, no electronic portfolios that reflected our assessment model were commercially available (discussed later in this paper). Therefore, we began to construct our own portfolio.

The REPS was first used during the summer of 1998 to evaluate a set of student submissions for a pilot project. Every year since then, we have used the REPS to collect, evaluate, and report achievement in student learning outcomes to students, faculty, employers, graduate schools, and various accrediting agencies. For the past two years, we have supplemented the REPS with the RosEvaluation Tool (RET), a plug-in component that we built for use with a course-management system, to evaluate student work products. Since our project began, numerous electronic-portfolio products have become available on the market. A review of these commercially available portfolios reveals a variety of approaches and methodologies [23].

### INSTITUTIONAL LEARNING OUTCOMES, PERFORMANCE CRITERIA, AND EVALUATION RUBRICS

The structure of the REPS is constituted by the Institutional Student Learning Outcomes that were developed early in 1997 and then revised in 2006. These outcomes define what we believe

TABLE II  
COMMUNICATION OUTCOME PERFORMANCE CRITERIA, REPS

Communication Criterion	Performance Description
1	Provide a substantive critique that includes recommendations for improvements.
2	Adapt technical information for a nonspecialized audience.
3	Convey information effectively through visual media.

all graduates of Rose-Hulman should be able to do once they enter their professions or graduate school. The challenge of the outcomes, however, is that they are not measurable; in other words, while we expect each student to demonstrate the skills necessary to work successfully on a team, the broad outcome does not provide measurable behaviors we could observe and then evaluate to determine if the student has met the outcome. For that reason, we developed a set of performance criteria and evaluation rubrics to define the required behaviors and to quantify the levels of performance that we expect. An example of this aspect of the system follows.

Each Rose-Hulman student is expected to demonstrate effective communication skills. Thus, Institutional Student Learning Outcome for Communication states that “Communication—regardless of the media—requires unique skills whether communicating with individuals or with groups.” This statement alone, however, is not measurable, meaning that the statement does not describe what the student should actually be able to do or the skills that he/she should possess. To measure this level of behavior, we developed a set of performance criteria (specific statements that explain exactly what the outcome means) and evaluation rubrics (descriptions of what successful performance means for each criterion) for each criterion. The nature of the performance criteria and rubrics should be noted. First, it would be possible to define “communication” and the expected behaviors in many different ways. For the purposes of our assessment project, we decided to focus on three primary performance areas. (See Table II.)

Student work products that can provide evidence of student learning are not specified; thus, a faculty member can determine which of his/her assignments provides the best evidence of student achievement. Example evidence documents for



these communication criteria include but are not limited to the following: criterion 1—peer reviews, performance evaluations, team evaluations; criterion 2—technical reports, product-design presentations for nonengineers; and criterion 3—PowerPoint presentation slides, charts, tables, and visuals from a technical report. For each criterion, there is a rubric that describes specifically how the student work product should be evaluated. For example, for criterion 1, the evaluation rubric states that

A passing submission for this criterion must  
(1) provide helpful/constructive criticism that gives recommendations for improvement and  
(2) justify recommendations.

The design of the rubric is supposed to offer students, faculty portfolio evaluators, and instructors making assignments with specific descriptions and examples that will help them understand exactly what is expected [24]. The outcomes, performance criteria, and rubrics were developed by a campus-wide committee made up of faculty from all disciplines. They are periodically reviewed and revised.

## ROSE PORTFOLIO RATING PROCESS

To determine students' success in achieving the learning outcomes, all student submissions to the REPS are assessed each year by a team of trained faculty raters. The purpose of the RosE Portfolio rating session is to assess evidence of student learning in the six outcomes. Student work products serve as evidence of student learning in these six outcomes, and the evidence is collected each year through assignments made by faculty in technical and nontechnical departments. For example, some engineering faculty members require that students submit documents from capstone senior design courses as evidence of the teamwork outcome. Humanities and social sciences faculty members require that students submit documents produced in their courses for evidence of the cultural and global awareness outcome. The definition of performance criteria and rubrics, collection of documents, as well as the assessment and evaluation of evidence for technical-learning outcomes is the province of technical departments (although many departments use the same portfolio collection and assessment methodology that will be described).

The process of rating submissions to the RosE Portfolio has followed the same basic

methodology since the system was initiated in 1998. Rose-Hulman faculty members (usually up to 14 each year) are hired as portfolio raters. Attempts are made to involve faculty from many different departments to ensure objectivity in rating and broad-based familiarity and participation in the process. Raters work together for two days in a computer laboratory and are compensated for their work. The rating session coordinator, usually a member of the faculty who has had a long association with the project, facilitates the process and assigns pairs of raters to rate student submissions for a particular outcome. For example, a faculty member from mechanical engineering and a faculty member from chemistry may work as a rating pair to assess the student files submitted for the communication outcome. The work of rating occurs within the RET interface that was developed inside the ANGEL Learning Management System. A screenshot of the RET rating screen is shown in Fig. 1.

The rating process consists of four steps.

- (1) First, portfolio raters review the rating rubric associated with the learning outcome. Each year, raters review the rating rubric and the comments of faculty who evaluated the same outcome in previous years. As part of its training, the rating team discusses the rubric while comparing it to student documents that were rated during previous rating sessions. The purpose of this paper is to ensure calibration between the two faculty raters as well as between the current faculty raters and each previous year's faculty-rater team. Calibration like this helps ensure consistency in ratings from year to year.
- (2) Second, the RET requires that each rater team rate a set of three shared documents against the established rubric. Raters answer "yes" or "no" for a single rating question: "Does this document meet the standard expected of a student who will graduate from Rose-Hulman?" Student achievement is measured as either "yes/pass" or "no/fail." Raters also have the opportunity to mark the document as "yes/pass/exemplary" to designate student submissions that represent superior achievement for a particular outcome. To ensure consistency in rating between the raters, the RET uses an inter-rater reliability (IRR) process. When raters read and evaluate the set of three shared documents, their ratings must agree. If their ratings are not identical, the RET prohibits them from

The screenshot displays the 'Evaluate Submissions' interface. At the top right, a status box indicates: 'You have rated 3% of this session', 'You have completed 0 of 11 sessions assigned to you', 'You are in set 1 of 11 in this session', 'You are rating submission 1 of 3 in this set', and a 'Jump to submission: 1' link. The main content area is divided into sections: 'Evaluation Criterion' with a link to 'Teamwork B1: Demonstrate how you reached a decision as a team.' and 'View Example Comments'; 'Submission Details' showing 'Assignment: "Gantt Chart" for EM103-04 Introduction to Design', 'Title of submission: Group 46 Gantt Chart (Stephanie Drenten, Steve Juszczak, Scott Eaton)', and 'Submission message: (none)'. It also states 'You are now rating submission 1 of 3 in this set.' and lists '1 file: Gantt\_Chart.docx'. The 'Evaluation' section features a 'Rating: Fail' dropdown, an 'Exemplary' checkbox, and a 'Comments:' text area. To the right, 'Rating codes:' are listed with checkboxes for various reasons like 'Deferred decision on exemplary', 'Document won't open', 'Duplicate submission', 'No highlight - couldn't locate best evidence', 'Other problem(s)', 'Document submitted to wrong criterion', 'Doesn't meet all components of the rubric', 'Mentions relevant issue but doesn't address it', 'Not project specific - insufficient detail', and 'Rating is borderline'. At the bottom are buttons for 'Go Back', 'Update Rating and Continue', and 'Return to Session Listing'.

Fig. 1. Portfolio rater screen.

continuing with the rating process. Raters then discuss their ratings, checking their evaluation against the rating rubric for the outcome; they then come to an agreement on how they will evaluate the shared document set. IRR is a key component of the RET; it ensures that raters look for the same qualities and features in order to rate documents. This helps the faculty raters to calibrate their ratings against each other and ensure consistency in ratings.

- (3) Third, if the raters agree in their IRR, the RET then allows them to proceed with a set of ten documents, with each rater reading and rating a different set of ten documents. As the faculty members assign ratings to documents, the RET records their ratings for each document. The system also introduces a shared file for every ten documents in order to check that the raters have maintained their IRR. Failure to rate the shared document identically will cause the system to stop the raters so that they can recalibrate their evaluation before moving on to another document set. Thus, IRR continues to validate ratings throughout the rating process.
- (4) Fourth, the raters can provide comments about the rating session or about the student submission in the comment boxes. In addition to the work of rating, faculty raters also record the rubrics they used and collect sample documents to provide next year's raters with material for calibration. They may also suggest changes to rating rubrics or learning outcomes, although revisions must be reviewed and approved by the Commission on

the Assessment of Student Outcomes (CASO) before they are implemented into REPS.

The issues of reliability and validity are important to note here. The methodology we adopted for our project relies on an integrative approach to portfolio scoring. Rather than working independently, portfolio evaluators "work together to construct coherent interpretations, continually challenging and revising initial interpretations" [25, p. 205]. This approach, outlined by Moss in 1998, is most applicable for faculty who are using portfolios for the purpose of examining an overall program, rather than individual students [25]. The IRR method used is a consensus estimate approach, defined by Stemler as

based on the assumption that reasonable observers should be able to come to exact agreement about how to apply the various levels of a scoring rubric to the observer behaviors. [26]

While we recognize the inherent advantages of the consensus estimate method (Stemler identifies these as a "strong intuitive appeal," "easy to calculate," and "easy to explain"), we also acknowledge its drawbacks. We are currently working to test other IRR methods as we develop the new RosEvaluation assessment tool inside of the ANGEL Learning Management System.

## USING PORTFOLIO RATING RESULTS FOR INSTITUTIONAL AND PROGRAM IMPROVEMENT

RosE Portfolio rating results from the annual Portfolio rating session are compiled for each



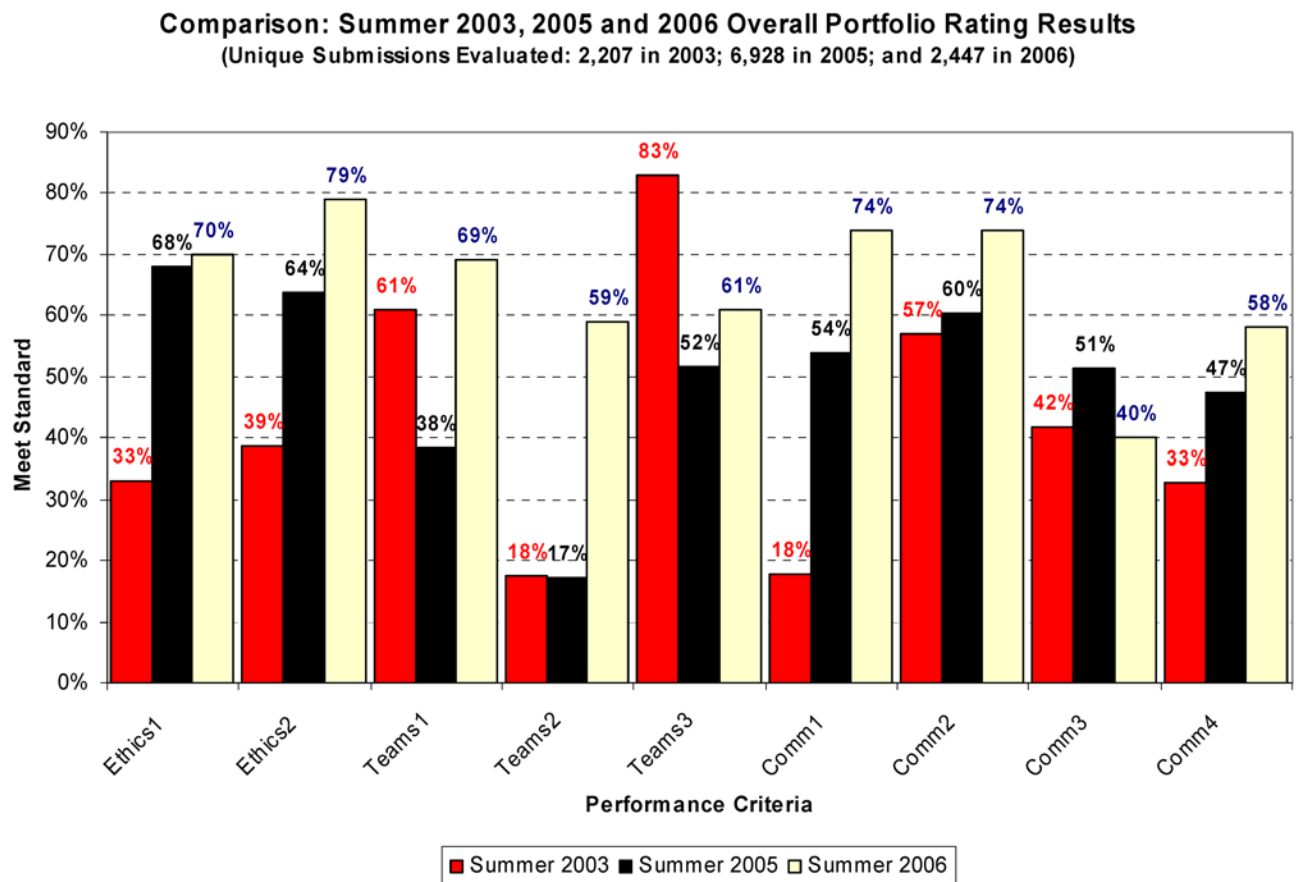


Fig. 2. Percent of submissions meeting the criteria in 2003, 2005, and 2006.

department and program on our campus. The students' achievements in learning outcomes are then used by each department to evaluate the effectiveness of curricula and to enact curricula change if necessary. These data are also provided to accrediting agencies to demonstrate successful achievement of student learning outcomes. At Rose-Hulman, our engineering and computer science programs are accredited by ABET. The programs' accreditation is noted on materials, such as departmental brochures and webpages.

The responsibility for evaluating students' achievement of learning outcomes is distributed across the institution. At each level, portfolio results produced through the REPS are assessed and evaluated to determine strategies for improvement. An example of portfolio results compiled for three rating sessions is shown in Fig. 2.

At the institute level, CASO reviews the results of the portfolio ratings each year and develops strategies to improve the portfolio process. CASO also maintains the assessment rubrics

used by the portfolio's raters. After the rating session for the summer portfolio is concluded, CASO members review the comments, ideas, and suggestions the portfolio raters provide for possible changes to the assessment rubrics. CASO then discusses all changes and implements them in the following year's session. In addition, CASO periodically reviews the list of learning outcomes to determine whether the outcomes should be revised. For instance, during the 2006–2007 academic year, CASO revised the institute outcomes and added new outcomes in leadership and service that reflected the changing nature of technical education. CASO provided feedback to Rose-Hulman and recommended the addition of the two new outcomes; the faculty gave final approval of the outcomes.

At the program level, departments are engaged in evaluating the REPS results and making changes to their own curricula. Academic departments review the results of the portfolio evaluations each year during their departmental retreats. At that time, they can evaluate the level of student achievement based on program and institute student-learning

outcomes. The portfolio results may indicate where changes need to be made in the curriculum. For instance, a review of the REPS results from the 2003 and 2005 rating sessions indicated that students were not achieving an adequate level of performance for the second performance criterion—the teamwork outcome. Teamwork criterion 2 stated that students should demonstrate that when they work with others on a team, they can analyze ideas objectively to discern feasible solutions by building consensus. In the Mechanical Engineering Department, members of the faculty analyzed the results by comparing them to their curriculum maps. This was done to ensure that students had been given adequate opportunities to develop their skills to reach this outcome. At that point, faculty members recognized that students had been given adequate opportunities; the problem seemed to lie in the assignments students were being asked to submit. The faculty reviewed the team assignment they had required students to submit and realized that the assignment—a final report from a team project—focused more on the product of the team's work rather than on the students' process of working together as a team. After reviewing the assignment, faculty members responsible for developing and requiring the assignment shifted their focus. Instead of requiring that teams submit the final project report, students were required to submit minutes from one of their team meetings. These minutes were meant to show the specific process the team used to decide among several design alternatives. Introducing a revised assignment provided students with the opportunity to show that they could evaluate different ideas and come to consensus as a team. Thus, the assignment changes ensured that students addressed the outcome in their work product. As indicated in Fig. 2, student performance in that performance criterion increased in the next rating session.



The institution also reports achievement in learning outcomes to its accrediting agency—the North Central Association (NCA) of the Higher Learning Commission. Rose-Hulman is participating in the Academic Quality Improvement Project (AQIP) of the NCA. AQIP is an accreditation program that focuses on quality-improvement processes within an institution. These processes must be ongoing and address all facets of the institution, from student learning and facilities to creating collaborative relationships and planning for the future.

## ADDITIONAL ASSESSMENT TOOLS

We are now in the eleventh year of using the REPS for the assessment of learning outcomes, but we do not rely on the REPS alone to provide us with data regarding student achievement. In addition to using the REPS, we employ a number of other assessment methods to determine whether students are achieving the stated learning outcomes.

First, each program at Rose-Hulman uses data collected from a number of sources to evaluate the effectiveness of curricula. The need for a curricular revision could be indicated by low ratings on student course-evaluation surveys or poor levels of performance on standardized tests, such as the Fundamentals of Engineering Exam. In that case, the program uses the information from these sources to make appropriate changes to the curriculum, checking the progress of the change at regular intervals in order to measure improvements.

Second, we collect information regarding alumni satisfaction through our annual alumni survey. This instrument, focusing on academic dimensions of the Rose-Hulman experience, asks alumni to evaluate two elements related to student learning outcomes: (1) How important is the outcome to the alumnus' current position (in industry, in graduate school, etc.)? and (2) How well did the alumnus' education at Rose-Hulman prepare him/her in this outcome? The outcomes listed on the alumni survey are the six institute learning outcomes as well as program-specific outcomes. Given the alumni data, program administrators can review curricula and propose revisions. For instance, in previous alumni surveys for graduates of the civil engineering program, respondents indicated that they thought their program needed to offer more courses in transportation. As a result of their responses as well as data from the program's advisory board and other sources, changes were made to increase the number of course offerings in transportation topics.

Third, we also use information regarding graduation rates, retention rates, and placement rates to gauge institutional and program effectiveness. Information regarding these rates is published on the website of the Office of Institutional Research. Since it is distributed to the public, the information is available to students, faculty, staff, alumni, prospective students and their families, as well as industry, graduate schools, and interested members of the community. These rates indicate that we are fulfilling our mission

to provide the best undergraduate education in engineering, mathematics, and science in an environment of individual attention and support. [27]

For instance, for the cohort of all full-time bachelor's (or equivalent) degree-seeking undergraduate students who entered our institution as freshmen in fall 2005 (or the preceding summer term), the percentage of students enrolled at our institution at the start of official enrollment in fall 2006 was 91.7%. In addition, our placement rate for graduates (into the industry, graduate school, etc.) remains consistently in the 97%–99% range, an indication that we are successfully preparing students for careers and further education.

## ROSE PORTFOLIO AND FACULTY PROFESSIONAL DEVELOPMENT

One important conduit for information to the public about students' achievement is the conferences at which faculty make presentations and give papers. As part of their professional development, faculty who have served as portfolio raters and/or been engaged in program assessment for their departments give presentations and papers at national conferences, such as the ASEE Conference, the Higher Learning Commission Conference, and the Association for Institutional Research Conference, to name only a few. As faculty report their research in engineering, mathematics, and science education, they present data gathered in the courses they teach. Rose-Hulman faculty members are also publishing articles in journals in their respective disciplines. Often, the topics address issues of student learning and pedagogical research.

## CONCLUSION

The call for accountability in higher education continues to be heard in the US and abroad.

## REFERENCES

- [1] Press release. (2005, Sep. 19). Secretary Spellings announces new Commission on the Future of Higher Education. [Online]. Available: <http://www.ed.gov/news/pressreleases/2005/09/09192005.html>
- [2] M. Spellings. (2005, Sep. 19). A national dialogue: Commission on the future of higher education. [Online]. Available: <http://www.ed.gov/news/speeches/2005/09/09192005.html>
- [3] Spellings Commission. (2006, Sep.). A test of leadership: Charting the Future of U.S. higher education. [Online]. Available: <http://www.ed.gov/about/bdscomm/list/hiedfuture/index.html>
- [4] B. Huot, "Consistently inconsistent: Business and the Spellings Commission report on higher education," *College English*, vol. 69, no. 5, pp. 512–525, 2007.
- [5] University and College Accountability Network. (2009). [Online]. Available: <http://www.ucan-network.org/>
- [6] National Survey of Student Engagement. (2009). [Online]. Available: <http://nsse.iub.edu/>
- [7] M. Marklein. (2007, Nov. 4). Beyond rankings: A new way to look for a college. *USA Today*. [Online]. Available: [http://www.usatoday.com/news/education/2007-11-04-nsse-cover\\_N.htm](http://www.usatoday.com/news/education/2007-11-04-nsse-cover_N.htm)

In a climate like this, assessment strategies will be increasingly important in identifying effective educational programs. The chorus for accountability, however, appears to be met by a competing set of voices—faculty in higher education who express exhaustion and frustration with having to maintain assessment systems within their own institutions even as they conduct research and teach. As a recent study of the impact of the ABET Engineering Criteria and the focus on outcomes assessment has shown, the criteria expanded the definition of engineering competencies to place much greater emphasis on “professional skills, such as solving unstructured problems, communicating effectively, and working in teams” and “shifted the basis for accreditation from inputs, such as what is taught, to outputs—what is learned” [28, p. 1]. These two changes were expected to be transformative:

program changes would reshape students' educational experiences inside and outside the classroom, which would in turn enhance student learning. [28, p. 2]

But these changes have come with a cost, and time will tell how sustainable these costs will be.



On the campus of Rose-Hulman, however, we have taken the sustainability issue and used it to calibrate our assessment efforts. The REPS has proven to be an effective and efficient tool for our purpose of documenting students' learning for accreditation. We use the data in all engineering programs on our campus to document student achievement in program accreditation and institutional accreditation. As we move into the next accreditation cycle (site visit 2012), we have made changes to our institute-student learning outcomes, but the data-collection method and its assessment methodology will remain the cornerstone of our efforts.

- [8] J. Williams, "Transformations in technical communication pedagogy: Engineering, writing, and ABET Engineering Criteria 2000," *Tech. Commun. Quart.*, vol. 10, no. 2, pp. 149–169, 2001.
- [9] ABET Inc. (2007, Dec. 7). Criteria for accrediting engineering programs. [Online]. Available: <http://www.abet.org/Linked%20Documents-UPDATE/Criteria%20and%20PP/E001%2008-09%20EAC%20Criteria%2012-04-07.pdf>
- [10] B. Bloom, Ed., *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. White Plains, NY: Longman, 1956.
- [11] R. Gagne, *The Conditions of Learning and New York: Theory of Instruction*. New York: Harcourt Brace, 1985.
- [12] W. Kelley, Ed., *Assessment in Engineering Programs: Evolving Best Practices*, ser. Assessment in the Disciplines, W. E. Kelley, Ed. Tallahassee, FL: Association for Institutional Research, 2008, vol. 3.
- [13] C. Malmberg, "Facilitating live assessment of outcomes in engineering education," in *Assessment in Engineering Programs: Evolving Best Practices*, ser. Assessment in the Disciplines, W. E. Kelley, Ed. Tallahassee, FL: Association for Institutional Research, 2008, vol. 3, pp. 103–121.
- [14] A. Estes, R. Welch, and S. Ressler, "Structured and systematic assessment: A successful and sustainable civil engineering example," in *Assessment in Engineering Programs: Evolving Best Practices*, ser. Assessment in the Disciplines, W. E. Kelley, Ed. Tallahassee, FL: Association for Institutional Research, 2008, vol. 3, pp. 123–141.
- [15] ABET, Inc. (2009, Jun. 10). Engineering Self-Study Questionnaire. [Online]. Available: <http://www.abet.org/Linked%20Documents-UPDATE/Program%20Docs/E003%20Self-Study%20Questionnaire%202009%206-11-09.doc>
- [16] P. Belanoff and M. Dickson, Eds., *Portfolios: Process and Product*. Portsmouth, NH: Boynton/Cook, 1991.
- [17] K. Yancy and I. Weiser, Eds., *Situating Portfolios: Four Perspectives*. Logan, UT: Utah State Univ. Press, 1997.
- [18] J. Williams, "The Engineering portfolio: Communication, reflection, and student outcomes assessment," *Int. J. Eng. Educ.*, vol. 18, no. 2, pp. 199–207, 2002.
- [19] F. Paulson, P. Paulson, and C. Meyer, "What makes a portfolio a portfolio?," *Educ. Leadership*, vol. 48, no. 5, pp. 60–63, 1991.
- [20] B. Cambridge, *Electronic Portfolios: Emerging Practices in Student, Faculty, and Institutional Learning*. Washington, DC: AAHE, 2001.
- [21] A. Jafari and C. Kaufman, *Handbook of Research on ePortfolios*. Hershey, PA: Idea Group Reference, 2006.
- [22] B. Olds, "Twenty years of assessment: A retrospective case study," in *Assessment in Engineering Programs: Evolving Best Practices*, ser. Assessment in the Disciplines, W. E. Kelley, Ed. Tallahassee, FL: Association for Institutional Research, 2008, vol. 3, pp. 159–172.
- [23] H. Barrett. (2006). Authentic assessment with electronic portfolios using common software and web 2.0 tools. [Online]. Available: <http://electronicportfolios.org/web20.html>
- [24] A. McKeever and E. Hobson. (2006, Jul. 5). Focusing faculty on key concerns: Using a primary-trait rubric to evaluate student drug posters presented at the Annu. Meeting of the American Association of Colleges of Pharmacy. [Online]. Available: [http://www.allacademic.com/meta/p\\_mla\\_apa\\_research\\_citation/1/1/9/1/2/p119125\\_index.html](http://www.allacademic.com/meta/p_mla_apa_research_citation/1/1/9/1/2/p119125_index.html)
- [25] P. Moss, "Rethinking validity for the assessment of teaching," in *With Portfolio in Hand: Validating the New Teacher Professionalism*, N. Lyons, Ed. New York: Teachers College Press, 1998, pp. 202–219.
- [26] S. Stemler. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*. [Online]. Available: <http://PAREonline.net/getvn.asp?v=9&n=4>
- [27] Rose-Hulman Institute of Technology. (n.d.) Rose-Hulman at a glance. [Online]. Available: <http://www.rose-hulman.edu/admissions/facts/facts.htm>
- [28] L. Lattuca, P. Terenzini, and J. Volkwein, *Engineering Change: A Study of the Impact of EC 2000*. Baltimore, MD: ABET, Inc., 2006.

**Julia M. Williams** (SM'XX) **[Please provide year of Senior Membership]** is Executive Director of the Office of Institutional Research, Planning, and Assessment and Professor of English at Rose-Hulman Institute of Technology, Terre Haute, IN. Her articles have appeared in *Technical Communication Quarterly*, *Technical Communication*, *Journal of the Society for Technical Communication*, *International Journal of Engineering Education*, *Journal of Engineering Education*, and *The Impact of Tablet PCs and Pen-based Technology on Education*. She is Pedagogy Editor of the IEEE TRANSACTIONS ON PROFESSIONAL COMMUNICATION.