RUBRIC DEVELOPMENT FOR ASSESSMENT OF UNDERGRADUATE RESEARCH

Evaluating Multidisciplinary Team Projects

JAMES A. NEWELL, HEIDI L. NEWELL, KEVIN D. DAHM
Rowan University • Glassboro, NJ 08028

Experts agree on the importance of involving undergraduates in research-based learning[1-3] and teamwork.[4-6] The Boyer Commission suggested that research-based learning should become the standard for undergraduate education.[7] Many universities are responding to this challenge by introducing multidisciplinary laboratory or design courses.[8,9] At Rowan University, we developed a method of addressing these diverse challenges while also implementing valuable pedagogical hands-on learning experiences[10,11] and technical communications.[12-14]

At Rowan University, all engineering students participate in an eight-semester course sequence known as the engineering clinics.[15] In the junior and senior years, these clinic courses involve multidisciplinary student teams working on semester-long or year-long research projects led by an engineering professor. Most of the projects have been sponsored by regional industries. Student teams under the supervision of chemical engineering faculty have worked on emerging topics that included enhancing the compressive properties of Kevlar, examining the performance of polymer fiber-wrapped concrete systems, advanced vegetable processing technology, metals purification, combustion, membrane separation processes, and many other areas of interest. Every engineering student participates in these projects and benefits from hands-on learning, exposure to emerging technologies, industrial contact, teamwork experience, and technical communications.

Difficulties arise in trying to assess student learning and performance in project-based team settings, however. Angelo and Cross[16] provided significant suggestions for assessing the attitude of students toward group work, but provided little insight into distinguishing individual and team performances. One difficulty is that evaluating the semester-long performance of teams working on projects involves a substantial number of variables. Clearly, successful completion of the project's technical aspects is an essential component for demonstrating student understanding, but Seat and Lord[17] observed that while industry seldom complains about the technical skills of engineering graduates, industrial employers and educators are concerned with performance skills (*i.e.*, interpersonal, communication, and teaming). Lewis, *et al.*,[18] correctly observed that if students are to develop effective teaming skills, teaming must be an explicit focus of the project.

It is unreasonable to expect students to achieve specific learning objectives from a series of courses when the faculty members themselves are unclear about what the learning objectives are and how to measure them. Young, *et al.*,[19] dis-

James Newell *is Associate Professor of Chemical Engineering at Rowan University. He currently serves as Secretary-Treasurer of the Chemical Engineering Division of ASEE and has won both ASEE's Ray Fahien Award for his contributions to engineering education and a Dow Outstanding New Faculty Award. His research interests include high-performance polymers, outcomes assessment, and integrating communication skills through the curriculum.*

Heidi Newell *is currently the assessment coordinator for the College of Engineering at Rowan University. She previously served as the assessment consultant for the University of North Dakota. She holds a PhD in Educational Leadership from the University of North Dakota, an MS in Industrial and Organizational Psychology from Clemson University, and a BA in Sociology from Bloomsburg University of Pennsylvania*

Kevin Dahm *is Assistant Professor of Chemical Engineering at Rowan University. He received his BS from Worcester Polytechnic Institute in 1992 and his PhD from the Massachusetts Institute of Technology in 1998. His primary technical area is in chemical kinetics and mechanisms. His current primary teaching interest is integrating process simulation throughout the chemical engineering curriculum, and he received the 2003 Joseph J. Martin Award for work in that area.*

cussed development of a criterion-based grading system to clarify expectations to students and to reduce inter-rater variability in grading, based on the ideas developed by Walvoord and Anderson.[20] This effort represented a significant step forward in course assessment; however, for graded assignments to capture the programmatic objectives, a daunting set of conditions would have to be met. Specifically,

- *Proper course objectives that arise exclusively from the educational objectives and fully encompass all of these objectives must be set*
- *Tests and other graded assignments must completely capture these objectives*
- *Student performance on exams or assignments must be a direct reflection of their abilities and not be influenced by test anxiety, poor test-taking skills, etc.*

There should be a direct correlation between student performance in courses and the overall learning of the students only if all of these conditions are met every time. Moreover, much of the pedagogical research warns of numerous pitfalls associated with using evaluative instruments (*e.g.,* grades on exams, papers, etc.) within courses as the primary basis for program assessment.[21]

Obviously, a more comprehensive assessment method for a team-oriented, research-project based course must be developed. Woods[22] listed the following five fundamental principles for assessment of teams:

1. *Assessment is based on performance*
2. *Assessment is a judgment based on evidence rather than on feelings*
3. *Assessment must have a purpose and have clearly defined performance goals*
4. *Assessment is done in the context of published goals and measurable criteria*
5. *Assessment should be based on multidimensional evidence*

TABLE 1
Summary of Specific Indicators for Areas of Importance

Area of Importance	*Specific Indicators*
Technical	• Defined objectives • Demonstrated technical awareness • Obtained and interpreted appropriate results • Formulated supportable conclusions • Properly considered error • Provided recommendations for future work
Logistical	• Organized project • Met deadlines • Executed project plan • Kept detailed records
Laboratory Operation	• Maintained safe practices • Developed hazardous operations (HAZOP) report • Dressed appropriately • Proper use/maintenance of equipment • Performed end-of-semester shut down
Teaming	• Division of labor • Professional conduct • Learning experiences for all team members

Part of the purpose of this pilot program was to clarify for the students the expectations in junior/senior clinic by providing specific information about their learning goals.

Rowan's Chemical Engineering Department is implementing the following strategy for improved assessment of student team projects: decide on the desired learning outcomes for the clinic, develop indicators that demonstrate whether or not the teams (and each member of the team) have achieved each of the outcomes, develop rubrics to evaluate student performance in each of the areas, and present all of this information to the students at the start of the project.

PILOT PROGRAM

In the junior/senior engineering clinic, each student team submits a final written report and gives an oral presentation, which allows the communication aspects of the project to be evaluated directly, but the remaining elements of a successful project experience had to be identified and measured. As a first effort to address the assessment of team performance in project-based research experiences, the faculty developed the following list of four learning objectives of primary importance that were common to all projects:

- *Technical performance*
- *Project planning and logistics*
- *Laboratory operation*
- *Teaming*

Once these objectives were identified, specific indicators were developed for each so the students would have clearly defined behaviors. Table 1 summarizes these indicators.

With the specific indicators determined, the next step involved developing descriptive phrases that would assist both students and faculty members in evaluating student performance. It became clear that specific descriptions of the level of performance in each area would be required. The goal of our rubrics was to map student work directly to the individual learning outcomes. As Banta[23] stated, "The challenge for assessment specialists, faculty, and administrators is not collecting data but connecting them." The assessment rubric also followed the format developed by Olds and Miller[24] for evaluating unit operations laboratory reports at the Colorado School of Mines.

The decision to frame the rubrics based on only three levels was significant and requires explanation. At one time, many of the other program-assessment instruments used by Rowan's Chemical Engineering Department used a 5-point Likert scale with qualitative labels (5=excellent, 4=very good, 3=good, 2=marginal, 1=poor), but the qualitative natures of the descriptive labels led to confusion in scoring. Some professors have different distinctions between "excellent" and "very good" and tended to use them more than the descriptive phrases that define the difference between levels for each indicator. More important, if the rubrics are well designed, the descriptive phrases should stand alone, without the need for subjective clarifiers such as "excellent" and "good." Ulti-

TABLE 2
Behaviors Corresponding to Technical Performance

Indicator	*An "A" Team*	*A "B" Team*	*A "C"-or-Lower Team*
Defined objectives	Is actively involved in defining aggressive and achievable objectives that thoroughly address fundamental project needs.	Aids in defining objectives. Some may be too simplistic or unrealistic.	Takes little initiative in defining the project.
Demonstrated technical awareness	Clearly demonstrates awareness of the work of others and establishes a context for their project. Shows an understanding of information from multiple literature sources.	Shows understanding of the work in the field, but has limited depth and breadth. Knowledge is limited to faculty-provided materials.	Fails to demonstrate an awareness of the work of others and the significance of of their project.
Obtained appropriate results	Obtained meaningful results with minimal wasted effort.	Produced some results but not enough (or too many).	Generated few meaningful results.
Interpreted data appropriately	Provided thorough and correct analysis of data.	Provided analysis but partially incorrect or insufficiently thorough.	Little meaningful analysis of data or blatantly incorrect.
Formulated supportable conclusions	Formulated and adequately supported meaningful conclusions.	Needed significant help in formulating meaningful conclusions or lacked sufficient support for their conclusions.	Conclusions are absent, wrong, trivial, or unsubstantiated.
Properly considered error	Used appropriate mathematical and technical skills to quantitatively express limitations of of the data.	Error analysis is largely qualitative or incomplete.	Sources of error and reproducibility issues are ignored or misinterpreted.
Provided recommendations for future work	Makes insightful recommendations about future work.	Makes broad or obvious suggestions for future work.	Makes no plausible suggestions for future . work.

TABLE 3
Behaviors Corresponding to Project Planning and Logistics

Indicator	*An "A" Team*	*A "B" Team*	*A "C"-or-Lower Team*
Organized project	Effectively organizes project tasks to minimize wasted time and effort.	Identifies relevant tasks but may struggle with setting priorities and planning.	Has difficulty converting broad objectives to specific tasks.
Met deadlines	Consistently meets deadlines.	Misses some deadlines despite reasonable effort.	Routinely ignores deadlines.
Executed project plan	Effectively and safely executes the project plan. Makes significant progress. Modifies the plan as necessary.	Executes the project plan but has difficulty overcoming setbacks.	Works haphazardly with little chance of achieving project objectives.
Kept detailed records	Keeps detailed records easily followed by others. These records include a laboratory notebook, computer files, purchase records, and others.	Keeps a lab notebook but records lack organization or contain omissions.	Keeps poor, sketchy, or no records.

mately, we decided to eliminate such descriptors and divide rubric elements by listing behaviors that demonstrated the level (1, 2, or 3) at which the student had obtained the desired learning outcomes.[25]

These previously developed rubrics, however, were programmatic assessment tools that were seen and used only by the faculty. Part of the purpose of this pilot program was to clarify for the students the expectations in junior/senior clinic by providing specific information about their learning goals. Students tend to be more focused on grades than on learning outcomes, so characterizations such as "level 1 vs. level 2" would be meaningless to them, and subjective phrases such as "excellent" and "good" would be subject to the same shortcomings described above. Further, if grading truly represents the measure of achievement of learning outcomes, it is not unreasonable to present the behaviors that demonstrate successful attainment of a learning outcome in terms of grades. Consequently, the rubrics were written for presentation to the students in terms of behaviors that an A-Team would demonstrate, a B-Team would demonstrate, etc., Tables 2 through 5 provide the rubrics.

Both the chemical engineering faculty at Rowan and the reviewers of this paper questioned if the "C-or-Lower" range was too broad. Some items were barely acceptable, while others could be dangerous. There was even a question about whether or not laboratory safety could be scaled at all. We decided to stay with three levels for several reasons. First, we did not want students bargaining about the lower-level

TABLE 4
Behaviors Corresponding to Laboratory Operations

Indicator	*An "A" Team*	*A "B" Team*	*A "C"-or-Lower Team*
Maintained safe practices	Develops and follows procedures that account for safety and clean-up. Lab is clean and neat.	Develops and follows procedures consistent with safe practices but sometimes misses minor safety issues or fails to clean up.	Fails to develop and follow safe procedures and/or clean up.
Developed Hazardous Operations (HAZOP) report	Conducts a thorough Haz-Op.	Performs a Haz-Op but focuses on obvious issues withouth depth (e.g., does not check MSDS sheets).	Fails to perform a Haz-Op or performs one inadequately.
Proper use/maintenance of equipment	Treats equipment with care and performs necessary maintenance.	Usually handles equipment properly but has an occasional lapse.	Uses equipment carelessly or fails to maintain it.
Performed end-of-semester shut down	Lab area is neat and clean. Lab notebook and electronic copies of all data and reports are provided to the faculty member. Samples and materials are labeled appropriately and are either stored or disposed of properly.	Must be pushed by the faculty member for the behaviors described previously.	Fails to accomplish some of the listed items.

TABLE 5
Behaviors Associated with Teaming

Indicator	*An "A" Team*	*A "B" Team*	*A "C"-or-Lower Team*
Division of labor	Has all members making significant contributions to a project that progresses . satisfactorily.	Progresses satisfactorily but some members feel that workload distribution was . disproportionate.	Internal conflicts result in team failing to achieve project goals.
Professional conduct	Consistently behaves in a professional manner (shows up for meetings prepared and on time; treats vendors, technicians, team members and staff with courtesy and respect; external communications are formal and. businesslike). Always dresses appropriately (long pants and safety glasses in labs; business attire for industrial meetings and presentations, etc.).	Usually behaves in a professional manner (shows up for meetings prepared and on time; treats vendors, technicians, team members, and staff with courtesy and respect; external communications are formal and businesslike). Usually dresses appropriately (long pants and safety glasses in labs; business attire for industrial meetings and presentations, etc.).Does not repeat errors.	Frequently fails to behave in a professional manner (shows up for meetings prepared and on time; treats vendors, technicians, team members and staff with courtesy and respect; external communications are formal and businesslike). Frequently fails to dress appropriately (long pants and safety glasses in labs, business attire for industrial meetings and presentations, etc.).
Learning experiences for all team members	Has all team members demonstrate a thorough understanding of the technical issues of the project.	Has all technical issues understood by someone on the team, but is segmented. Some members do not have the whole picture.	Has team members with significant gaps in their understanding of technical issues.

Faculty distributed the tables to the students at the beginning of the semester, referred to them throughout the semester in giving feedback on student performance, and used them to aid in assigning and justifying a final grade.

behaviors (*e.g.*, "I can be late for three meetings and still get a 'C,' but the fourth one gets me a 'D'."). The lowest-level behaviors were to be avoided entirely, so we chose not to put a distinction between "bad" and "really bad." The other important point to keep in mind is that the rubric items do not represent individual grades, but rather a holistic approach to evaluating all of the factors on a team. If the team has mostly A-level performances but also has some "C-or-Lowers," it would likely lower their project grade to a "B."

TABLE 6
Faculty Assessment of Grading Rubrics
(1=strongly disagree...4=strongly agree)

Statement	Mean Response
• The grading rubrics helped me explain the expectations of my project.	3.80
• The grading rubrics helped me determine how my team would be graded.	3.70
• The grading rubrics helped me consider project issues that I otherwise might not have considered.	3.30
• I referred to the grading rubrics during the semester.	3.40
• I think that clinic is more fair using grading rubrics.	3.70
• I would like to use the rubrics again next semester.	3.80

RESULTS AND DISCUSSION

The rubrics have two uses, each of which was piloted within the Chemical Engineering Department during the 2002-03 academic year. The first is that it will facilitate grading that is uniform, fair, and clearly understood by the students. Faculty distributed the tables to the students at the beginning of the semester, referred to them throughout the semester in giving feedback on student performance, and used them to aid in assigning and justifying a final grade.

The second use of the rubrics is assessment of the junior/senior clinic program as a whole. As mentioned above, simply using course grades as a primary assessment tool (even when the grades are fair and based on well-constructed criteria) has pitfalls. In the junior/senior clinic, for example, there is a danger that students will perform well overall but have widespread deficiencies in one or two areas. In such a case, the fact that most teams earned A's and B's for the semester would imply that students in the junior/senior clinic are meeting the desired learning outcomes, when in reality there is a need for specific improvement. As part of the pilot assessment program, faculty went through the eighteen indicators, one by one, and examined the level of performance demonstrated by each team with respect to each indicator. Through this process, specific problem areas were uncovered even when the overall student performance was objectively very good.

Faculty members were asked to assess the effectiveness of the rubrics. Table 6 indicates that the faculty clearly felt the rubrics were useful in improving fairness and linking the grading to the learning objective. In our annual assessment review, however, the faculty decided that it would be more valuable to have the students do a mid-semester assessment of progress based on the rubrics. Ideally, this should help both the team and the professor identify areas that need improvement while there is still time to adjust. Specific faculty comments about the rubrics included, "I felt much more confidant that my grade meant something," and "I was able to use items from the rubrics to drive my teams and help keep them on track."

Student comments about the rubrics were more mixed. They were discussed with a focus group of seniors who had participated in the clinic the previous year without the rubrics. Their consensus was that the rubrics were useful and probably the correct way to do things, but one student asked, "Couldn't you have waited until I graduated to implement these?" The students also expressed concern that the rubrics could be used as a basis for artificially lowering grades.

Ironically, part of the impetus for developing the rubrics was a concern that grading that seemed arbitrary might lead to grade inflation. In fact, more "A"s were given using the rubrics than had been given the previous year when no rubrics were used. The faculty attributed the change to improvement by the students. When we told the students what we expected them to do, more of them did it.

FUTURE WORK

Although development of the above rubrics represents a significant step forward, the results presented here describe a pilot study. Substantial work remains to be addressed. Meaningful assessment instruments must be developed to gauge student and faculty perceptions of these criteria. Are the critical learning objectives addressed in these rubrics and are the measurements accurate? Appropriate and meaningful weightings must be developed for each of the behaviors. While appropriate dress has been listed as an important part of the project, one would be unlikely to argue that it is as significant a learning objective as "drew meaningful and supportable conclusions."

Once the rubrics have been optimized, the next major task to be addressed is differentiating the performance of individuals from the performance of the team. It is possible that a team could have one (or more) member who fully attains the desired learning outcomes, but whose teammates fall substantially short of achieving those outcomes. Currently, the Chemical Engineering Department at Rowan University uses a peer-assessment technique modeled after a process described by Felder.[26]

Although this is a useful tool, it is somewhat over-reliant on student evaluation of peers. Our experience indicates that reasonably successful teams generally recommend an equal distribution of points, while the recommendation of less successful teams often are clouded with personal issues and resentments. Because students tend to focus on grades rather than on learning outcomes, their responses tend to be holistic (person X should get 50% of the points) and more about evaluation and grading, but less about achieving specified learning outcomes.

A major thrust of this effort is to develop evidence-based tools to complement the Felder survey, such that students could more meaningfully assess the performance of their teammates without defaulting to meaningless (*e.g,* "everyone contributed equally"), hierarchial (*e.g.,* "person X was terrible," but no reasons provided), or personal assessments. Moreover, the students will be required to cite specific evidence linking their evaluations to the specific desired learning outcomes. Ideally, in addition to aiding the faculty member in attempting to discern individual achievement from a group experience, forcing an evidence-based approach may help the students recognize the importance of the learning outcomes.

REFERENCES

1. Gates, A.Q., P.J. Teller, A. Bernat, N. Delgado, and C.K. Della-Pinna, "Expanding Participation in Undergraduate Research Using the Affinity Group Model," *J. Eng. Ed.,* **88**(4), 409 (1999)
2. Kardash, C.M., "Evaluation of an Undergraduate Research Experience: Perceptions of Undergraduate Interns and Their Faculty Mentors," *J. Ed.Psychology,* **92**, 191 (2000)
3. Zydney, A., J.S. Bennett, A. Shahid, and K. Bauer, "Impact of Undergraduate Research Experience in Engineering," *J. Eng. Ed.,* **91**(2), 151 (2002)
4. Guzzo, R.A., and M.W. Dickson, "Teams in Organizations: Recent Research on Performance and Effectiveness," *Ann. Rev. of Psychology,* **47**, 307 (1996)
5. Katzenbach, J.R., and D.K. Smith, *The Wisdom of Teams: Creating the High Performance Organization,* Harvard Business School Press, Boston, MA (1993)
6. Byrd, J.S., and J.L. Hudgkins, "Teaming in the Design Laboratory," *J. Eng. Ed.,* **84**(4), 335 (1995)
7. Boyer Commission on Education of Undergraduates in the Research University, *Reinventing Undergraduate Education: A Blueprint for America's Research Universities,* New York, NY (1998)
8. King, R.H., T.E. Parker, T.P. Grover, J.P.Gosink, and N.T. Middleton, "A Multidisciplinary Engineering Laboratory Course," *J. Eng. Ed.,* **88**(3), 311 (1999)
9. Barr, R.E., P.S. Schmidt, T.J. Krueger, and C.Y. Twu, "An Introduction to Engineering Through an Integrated Reverse Engineering and Design Graphics Project," *J. Eng. Ed.,* **89**(4), 413 (2000)
10. Heshmat, A.A., and A. Firasat, "Hands-On Experience: An Integrated Part of Engineering Curriculum Reform," *J. Eng. Ed.,* **85**(4), 327 (1996)
11. Schmalzel, J., A.J. Marchese, and R. Hesketh, "What's Brewing in the Engineering Clinic?" *Hewlett Packard Eng. Ed.,* **2**(1), 6 (1998)
12. Newell, J.A., D.K. Ludlow, and S.P.K. Sternberg, "Progressive Development of Oral and Written Communication Skills Across an Integrated Laboratory Sequence," *Chem. Eng. Ed.,* **31**(2), 116 (1997)
13. Van Orden, N., "Is Writing an Effective Way to Learn Chemical Concepts?" *J. Chem. Ed.,* **67**(7), 583 (1990)
14. Fricke, A.C., "From the Classroom to the Workplace: Motivating Students to Learn in Industry," *Chem. Eng. Ed.,* **33**(1), 84 (1999)
15. Newell, J.A., A.J. Marchese, R.P. Ramachandran, B. Sukumaran, and R. Harvey, "Multidisciplinary Design and Communication: A Pedagogical Vision," *Int. J. Eng. Ed.,* **15**(5), 376 (1999)
16. Angelo, T.A., and K.P. Cross, *Classroom Assessment Techniques: A Handbook for College Teachers,* 2nd ed., Jossey-Bass, Inc., San Francisco, CA (1993)
17. Seat, E., and S.M. Lord, "Enabling Effective Engineering Teams: A Program for Teaching Interaction Skills," *J. Eng. Ed.,* **88**(4), 385 (1999)
18. Lewis, P., D. Aldridge, and P. Swamidass, "Assessing Teaming Skills Acquisition on Undergraduate Project Teams," *J. Eng. Ed.,* **87**(2), 149 (1998)
19. Young, V.L., D. Ridgway, M.E. Prudich, D.J. Goetz, and B.J. Stuart, "Criterion-Based Grading for Learning and Assessment in the Unit Operations Laboratory," *Proc. 2001 ASEE Nat Meet.,* Albuquerque (2001)
20. Walvoord, B.E., and V.J. Anderson, *Effective Grading: A Tool for Learning and Assessment,* Jossey-Bass, Inc., San Francisco, CA (1998)
21. Terenzinis, P.T., and E.T. Pascarella, *How College Affects Students: Findings and Insights from Twenty Years of Research,* Jossey-Bass, Inc., San Francisco, CA (1991)
22. Woods, D.R., "Team Building: How to Develop and Evaluate Individual Effectiveness in Teams," Workshop at 2000 American Institute of Chemical Engineering (AIChE) National Meeting, Los Angeles, CA (2000)
23. Banta, T.W., J.P. Lund, K.E. Black, and F.W. Oblander, *Assessment in Practice,* Jossey-Bass, Inc., San Francisco, CA (1996)
24. Olds, B.M., and R.L. Miller, "Using Portfolios to Assess a ChE Program," *Chem. Eng. Ed.,* **33**(2), 110 (1999)
25. Newell, J.A., K.D. Dahm, and H.L. Newell, "Rubric Development and Inter-Rater Reliability Issues in Assessing Learning Outcomes," *Chem. Eg. Ed.,* **36**(3), 212 (2002)
26. Kaufman, D.B., R.M. Felder, and H. Fuller, "Accounting for Individual Effort in Cooperative Learning Teams," *J. Eng. Ed.,* **89**, 2 (2000) ❐