# ASSESSING PROBLEM-SOLVING SKILLS

## Part 2: Assessing the Process of Problem Solving

Donald R. Woods, Theodora Kourti, Philip E. Wood, Heather Sheardown, Cameron M. Crowe, and James M. Dickson
*McMaster University • Hamilton, Ontario, CANADA L8S 4L7*

Our graduates should be skilled at solving problems.[1,2] In Part I,[3] we distinguished between "problem solving" and "exercise solving," summarized five principles of assessment, listed four example goals and criteria for skill in problem solving, and gave seven forms of evidence. These forms were basically tasks in the subject discipline that are assessed for the answer, comprehension of the subject knowledge, and skill in problem solving. One of the major difficulties in trying to use this as evidence is that it is hard to make certain the assignment is a "problem" that tests higher-order thinking and not an "exercise." Other difficulties include the lack of published marking criteria for the process and the student's inability to perform well on exams because of exam anxiety (or because of the lack of study skills or the inability to display the problem-solving process well).

In this paper, we describe other forms of evidence besides the seven options given in Part I.[3] In Section 1 we list eight forms of evidence related more directly to the problem-solving process. In Section 2 we analyze the relationship among some of these suggested forms of evidence. In Section 3 we offer suggestions for evaluating a program's effectiveness in developing skill in problem solving.

***Donald R. Woods*** *is Professor Emeritus of Chemical Engineering at McMaster University. He is a graduate of Queen's University and the University of Wisconsin.*

***Theodora Kourti*** *is Research Manager in the McMaster Advanced Control Consortium. She received her bachelor degree in chemical engineering from the Aristotle University of Thessaloniki, Greece, and her PhD from McMaster University.*

***Philip E. Wood*** *is Associate Dean of the Faculty of Engineering at McMaster University and is a graduate of the University of Waterloo and California Institute of Technology.*

***Heather Sheardown*** *is Assistant Professor of Chemical Engineering at McMaster University. She received her BEng from McMaster and her PhD from the University of Toronto.*

***Cameron M. Crowe*** *is Professor Emeritus of Chemical Engineering at McMaster University. He is a graduate of McGill University and the University of Cambridge.*

***James M. Dickson*** *(BASc, MASc, University of Waterloo; PhD, Virginia Tech) is Professor of Chemical Engineering, McMaster University.*

### OPTIONS PRIMARILY FOR THE PROBLEM-SOLVING PROCESS

Once the goals and criteria are developed and published, consistent with assessment Principle #4,[3] a wide range of evidence can be gathered. Option 1[3] (mark the answer) and options 2 through 7[3] focused on marking the answer, subject knowledge comprehension, and the problem-solving process. Here, we continue and consider eight additional options where we gather evidence of the problem-solving process.

***Option 8: Written or Talk-Aloud Scripts of the Problem-Solving Process for a Variety of Chemical Engineering Questions*** • In Option 2,[3] the written script (or problem solution) tends to be marked mainly for comprehension of the subject knowledge. In Option 3,[3] providing published criteria about the process makes it easier to assess the problem solving, although some criteria may be intimately connected with comprehension of the subject knowledge. Option 8 is a variation on protocol analysis used by researchers to analyze problem solving. Students are asked to write their thought processes as they solve a problem. We are still plagued by the question, "Does the student find this to be a problem or an exercise?" We can usually determine this from the script, however. Question 1 in Table 1 illustrates the approach. Some example criteria used for marking the script include that the students demonstrate

- *That they are aware of the process and can write it out or describe it (the script should be complete and intuitive processes should be made explicit)*

- *That they are organized and systematic*
- *That they apply a strategy flexibly*
- *That they monitor frequently (by asking, "Why am I doing this?" or saying, "If I do this, then...")*
- *That they check and double check, show a concern for accuracy*
- *That they are active by underlining, writing things down, and making charts (to overcome the limitations of short-term memory)*

The script is relatively easy to mark and the students' marks are broadly distributed. For example, in the MPS program[4] for a chemical engineering problem taken from a text by Felder and Rousseau,[5] the average mark was 58% with a standard deviation of 23%. For the chemical engineering questions we selected, we found no evidence to suggest that the students (sample size, n=73) were exercise solving.

A second observation relates to the style of script produced by students under Option 2,[3] (mark the answer, the comprehension of the subject knowledge, and the problem-solving process), compared with Option 8. In our program, the same students were registered in both the mass balance course and the problem-solving course. Scripts produced by the same students on similar problems from the text by Felder and Rousseau[5] showed that under the conditions for Option 2, few details about the process were included. Under Option 8, students described mistakes made, corrections and adjustments, and attitudes.

Angelo and Cross[6] describe similar options, CATs 21 and 22, and suggest that these be used to monitor the classroom learning experience. Both CAT 21 and 22 could provide evidence for assessment if the criteria are published and consistent with the goals, as described in Part I,[3] and if the students are aware of the goals and the purpose of the activity. Remember, however, that for CAT 21, in step two when you solve the problem yourself, you are modeling "exercise solving." Teachers are expert in their subject areas and have a rich set of correctly-solved problems in memory. In "problem solving," the steps taken are more likely to be similar to those described by Bodner[7]

> *". . . read the problem, read it again, write down what you hope is the relevant information, draw a picture, make a list, write an equation to help you begin to understand, try something, try something else, see where it gets you, read the problem again, try something else, see where it gets you, test intermediate results to see if you are making any progress toward an answer, read the problem again, when appropriate strike your head . . . and so on."*

Teachers rarely model such a process.

***Option 9: Test and Exams that Probe Problem-Solving Skills (TEPS)*** • Tests are created in the context of public, observable, and unambiguous learning goals together with measurable criteria that are created explicitly for problem-solving skills.[3] Table 1 gives some sample TEPS.

Angelo and Cross' CATs 19 and 20, "problem recognition tasks" and "What's the principle?,"[6] are suggested as methods to monitor classroom learning. These could be converted into a TEPS if the criteria are published and consistent with the goals, as described in Part I,[3] and if the students are aware of the goals and the purpose of the activity. Otherwise the CATs fail to meet assessment Principle 4, which says the judgment should be done in the context of published goals, measurable criteria, and pertinent, agreed-upon forms of evidence.[3]

***Option 10: Self-Assessment of Achievement of the Target Skills for Problem Solving*** • Individuals can periodically reflect on and self-assess their skills relative to the published target skills for problem solving.[4,8-10] Similarly, the list of target skills for trouble shooters[3,4] can be used to document growth through self-assessment.

***Option 11: Self- and Peer Assessment of Problem-Solving Activities*** • Traditionally, the evidence is assessed by the teacher. The assessment can be made, however, by outside evaluators, by peers, or by each student as self-assessment. Alverno College[11] emphasizes the importance of self-assessment as learning. In the MPS program,[4] students receive between six and eight hours of workshop training on assessment and self-assessment.[12] Students practice applying the principles of self-assessment through resume writing, personal enrichment projects, and their written reflective journals. Each has a personal interview (similar to a performance appraisal) twice during the course with the instructor. In these interviews, the student and the instructor consider the evidence and the "skill mark based on evidence" and agree on the degree to which the goals of the course have been achieved.

***Option 12: Written Evidence from Workshops on Problem Solving*** • As part of the classroom activities to develop students' skill in problem solving,[4] the student completes a range of written activities. The written evidence includes peer feedback, personal reflections, individual and team tasks, and self-assessment.

***Option 13: Written Reflections of the Process*** • Researchers have noted the importance of reflection as a means of developing skill.[13-15] Kimbell, *et al.*,[13] and Schon[14] have shown the merits of reflection. Students can be asked to stop often and write reflections on how they did the task, what they are discovering, or what went on.[15] Such written reflections provide a rich set of evidence about the problem-solving processes.

***Option 14: Reflective Assessment Journals of the Process and the Application*** • Students can keep a record of their

**TABLE 1**
**Sample Questions Based on Objectives for Subset Problem-Solving Skills, TEPS**

## Question 1

**Awareness of the overall process:** For the next 20 minutes, write out your thought processes as you solve the problem. Please show me the process you are using. *(Insert a chemical engineering problem.)*

Time: 20 minutes

## Question 2

**Goal setting, criteria generating and self-assessment:** A friend wants to (improve stress management, lose weight, stop smoking, improve her self-image, etc.) She asks your advice. For the goal "to improve stress management:"

- *Break the goal into two subgoals*
- *Rewrite one subgoal in "observable terms"*
- *Create at least one measurable criteria that could be used to measure success in achieving the subgoal selected in the second step above*
- *Write out the type of evidence you would collect to show progress toward the subgoal given in the second step above*

Time: 25 minutes

## Question 3

**Analysis, classification:** Draw a concept map for the information given in Table 1. *(Provide Table 1 from a textbook.)*

Time: 30 minutes

## Question 4

**Creativity, classification, and awareness:** For the trouble-shooting problem given in Figure 1 *(provide a trouble-shooting situation as Figure 1)*:

- *Brainstorm fifty possible causes and write these down within ten minutes*
- *Analyze your list, note the basis of classification, and divide your ideas into at least seven different categories (within thirteen minutes)*
- *Select four ideas which are technically feasible (within two minutes)*
- *Select the "craziest" idea and write out your thought processes as you use this idea as a trigger or stepping stone to obtain a "new" technically feasible idea (within 20 minutes)*

Total Time: 45 minutes

## Question 5

**Analysis and classification:** For the outline for a written report *(an example is given below)*:

- *Identify the basis of classification and indicate any faulty coordination or faulty subordination*
- *List any other faults in the classification*
- *Write out a corrected classification*

Audience: Engineer who needs to know information about the potential development of a hydroelectric dam.

1. Introduction:
    A. Water power and its significance for Canada's development
2. The Crane Canyon Dam and Resevoir
    A. Information about the dam site
        1. Suitability of the dam site
        2. Character of the subsoil for a resevoir
        3. Cost on constructing a dam on this site
        4. Cost of preparing the area for a resevoir
        5. Cost of constructing the power generation facility
        6. Cost of transmitting the power
    B. Information about the water obtainable
        1. Minimum run-off available
        2. Probable water loss from leakage and evaporation
    C. Information about the community
        1. Results from public input meeting number one
        2. Results from public input meeting number two
        3. Summary of "No Damn Dam Here" rally
        4. Input from the chamber of commerce
    D. References
    E. Appendices
        1. Consultant's report on the soil
        2. Minutes of public input meeting number one
        3. Minutes of public input meeting number two
        4. Newspaper clippings about the "Damn" rally

Time: 20 minutes

## Question 6

**Self-awareness and understanding of inventories about personal style:** You are on a team consisting of the people listed in the chart below. Also listed are the Jungian typology and Kirton inventory for the other people. Add your own scores. What are the implications of the results of these inventories on:

- *Potential conflict—be concrete and identify the people, their behavior and explain the conflict and the degree of conflict*
- *Team "blind spots" where the team may not be as effective*

Time: 20 minutes

| Member | I | S | T | P | KAI |
|---|---|---|---|---|---|
| YOU: | ___ | ___ | ___ | ___ | ___ |
| Andy | 23 | 25 | 15 | 19 | 85 |
| Lisanne | 15 | 28 | 13 | 20 | 91 |
| Jean | 28 | 21 | 26 | 13 | 82 |
| Tom | 10 | 30 | 35 | 26 | 70 |

## Question 7

**Self-awareness and understanding of inventories related to personal style of studying and learning:** For the "Approaches to Study" inventory and the "Strategic Measure," (first number):

- *Describe the meaning of a score of 5*
- *Describe the meaning of a score of 24*
- *What is your score?*
- *Describe the implications of your score? What does your score mean?*
- *Compared to the scores on the other two measures ("Rote" and "Search for Meaning"), is the score on the strategic measure a score that represents you? Explain.*

Time: 20 minutes

**TABLE 1 (continued)**
**Sample Questions Based on Objectives for Subset Problem-Solving Skills, TEPS**

Question 8

**Classification, "define the stated problem," and visual thinking:** For the problem in Table 1 *(provide Table 1 to students, the problem statement should be very long and complex)*:

- *Identify the stated goal, task to be done, or unknown to be determined*
- *Draw a "good" diagram(s) to represent the situation*
- *Indicate the system by drawing (on the diagram drawn in the previous step) a dotted line around it*
- *Identify the knowns*
- *Identify the stated constraints*
- *Identify the stated criteria*

Do not continue to solve the problem!

Time: 30 minutes

Question 9

**Stress management:** a) Monitor your own stress level before you read question eight, after you have read it, and again after you have completed it. Use a rating scale of zero to ten (zero is negligible stress and ten is high stress) to record your stress levels at these times. Distinguish between positive and negative stress. Use the following chart to summarize you ratings.

b) List the stress management techniques that seem to work for you under "test" conditions.

Time: 5 minutes

| | Positive Stress | Distress |
|---|---|---|
| **Before reading question 8** | | |
| **After reading question 8** | | |
| **After completing question 8** | | |

Question 10

**Learning:** Use a Larkin Checklist for "Henry's Law." Use the form given *(provide a form)*:

Time: 25 minutes

Question 11

**Stress management:** For stress management, select the task that is most appropriate for you from those given below and describe how this relates to how you manage stress and anxiety in a problem-solving situation (select only one):

- *List the traditions that are important to you to provide touchstones of stability in your life*
- *List the personal put-down talk that you used to use on yourself and write out how you have changed*
- *Write down the stress management techniques that you used to use and those that you use now when you are writing an exam*

Time: 10 minutes

Question 12

**Self-assessment:** Figure 2 *(provide Figure 2)* records a person's description of what he or she did in a summer job. Based on this information:

- *Write a paragraph describing the job and create additional information if you need to. Rationalize your need to add this information*
- *Consider your answer to the above direction to be "evidence" of your skills. Based on this evidence, list the skills that you can claim to have*
- *Write the introductory statement of, "Skills I bring to this job"*

Time: 10 minutes

Question 13

**Experience factors for chemical engineers:** Use SI units to record reasonable, order of magnitude values for the following:

- *The density of liquid water*
- *The vapor pressure of benzene at 100°C*
- *The heat of vaporization of water*
- *The density of air at room temperature*
- *The compressibility factor for "nonideality" for a gas at a pressure of 10 atm*
- *For a reactor producing compounds like ammonia and methanol, compute the fraction of the reactor exit stream that would be recycled*
- *Compute the fraction of a recycle stream that is usually sent to purge*

Time: 7 minutes

Question 14

**Diagrams and visual thinking:** Figure 3 *(provide Figure 3)* shows a section of a process to make ammonia. In this process, natural gas and steam react in the primary and secondary reformers to make ammonia. On this section of the flow diagram:

- *Add any missing streams that would be needed for a mass balance*
- *For the recirculating pump attached to the steam drum on the primary reformer, is the symbol for the pump consistent with the direction of the arrows?*
- *For the heat exchanger on the exit gas from the secondary reformer, is the exit gas on the shell side or the tube side of the exchanger?*

Time: 10 minutes

daily use and application of problem-solving skills in the context of published, explicit goals and criteria. In the MPS program,[4] students submit a written reflective journal for each unit. Based on the written evidence, students self-assess the degree to which they have achieved the goals for the unit. The written evidence can include materials from Options 7,[3] 8, and 10-13. They add reflections of personal, outside-of-class efforts to bridge and extend the application of the target skills to homework assignments in other courses and to solve problems in their everyday lives. In reference four (Table 5 by Woods, *et al.*) is a list of some example activities. Examples of written reflective journals have been published.[16,17] These reflective journals can be marked with two sets of criteria: 1) the quality of the written assessment, and 2) the degree of problem-solving skill development as substantiated by the submitted evidence.

***Option 15: Portfolio About Skill in Problem Solving*** • A portfolio is a collection of evidence of "best" work. Whereas Option 14, Reflective Assessment Journals, is structured around faculty-generated goals for skill in problem solving, the portfolio allows students to select materials that represent their best work. Ideally, the student should write in the portfolio a clear summary of the skills achieved, as demonstrated by the materials included in the portfolio. Consistent with assessment Principles 2 (judgment is based on evidence) and 4,[3] the goals and marking criteria, whether generated by the faculty or the students, should be published at the beginning of the semester.

## AN ANALYSIS OF SOME FORMS OF EVIDENCE FOR THE ASSESSMENT OF PROBLEM-SOLVING SKILLS

Of the seven options listed in Part I and the eight options listed in Part II, we focus on Option 2 from Part I (because it is traditionally used in many programs and because some measures of the answer and the subject knowledge are needed) and Options 8-14 (where the emphasis is on evidence in the context of published goals and criteria for problem solving).

Options 8-14 have been used in the McMaster problem-solving program,[4] which consists of a sequence of four required courses to develop (as program outcomes) the student's skill in problem solving, communication, team and group work, lifetime learning, self-assessment, and change management. The focus in this paper is on the sophomore course that addresses development of problem-solving skills.

To assess the development of problem-solving skills, we use the following four forms of evidence:

- *Tests and examinations for components of problem solving, called TEPS in the McMaster program, (Option 9). The TEPS, including scripts, were used for the two-hour written examinations held at the end of the problem-solving course. The resulting mark we call the "exam mark."*
- *A script TEPS that requires a script to be written, as described in Option 8, and as illustrated in Question 1 of Table 1.*
- *Reflective journals (Option 14) that use evidence from Options 10-13, and which were marked for the quality of the written assessment.*
- *Self-assessment (Option 11) based on a personal interview and the accumulated evidence of skill development.*

These four are independent measures. The course in problem solving was taught by two of the coauthors, HS and DW.

Using multivariate statistical analysis, we examined two issues. Is there any correlation between the results from Options 8, 9, 11, and 14? Are there any correlations between exam marks in chemical engineering courses and the measure of problem-solving skills? In multivariate analysis, the sample correlation coefficient between $y_1$ and $y_2$ is defined as[18]

$$\rho(y_1, y_2) = \frac{\sum (y_1 - \bar{y}_1)(y_2 - \bar{y}_2)}{s_1 s_2} \qquad (1)$$

where: y=individual variable, $\bar{y}$=mean, s=standard deviation, and n=sample size

► *Are the results from Options 8, 9, 11, and 14 correlated?*

For two relatively large sample populations, n=112 and n=256, the journal (Option 14) and the self-assessment (Option 11) are highly correlated ($\rho$=0.73) and these are moderately correlated with TEPS (Option 9) ($\rho$=0.40). For smaller samples, n=49 and n=72, where data were available on the script (Option 8), the journal (Option 14) and the self-assessment (Option 11) are highly correlated ($\rho$=0.78). They are moderately correlated with TEPS, excluding the script (Option 9) ($\rho$=0.45). The journal and self-assessment are not significantly correlated with the script (Option 8) ($\rho$=0.29). An in-depth principal component analysis, PCA, confirmed these findings. We used PCA to eliminate outliers and to confirm robustness of the results.

For those who seek a single "gold measure" of problem solving, according to assessment Principle 5,[3] the measure should include many different forms of evidence and we wanted the students to have some input as to the weighting of the components. Here is one approach to obtaining a single measure of problem-solving skill that combines different forms of evidence. The TEPS (including the script mark when this is available) we refer to as the "exam mark." The "term mark" is a fifty-fifty mix of the "sum of the marks for the quality of the reflective journals" (Option 14) based on published measurable criteria, and the "self-assessment mark," (Option 11) based on published and measurable criteria, written evidence, and confirmed by the personal interview.

Students could select any weighting between 90 and 10 for the exam mark relative to the term mark. The selection was made in writing by the end of week nine in a twelve-week semester. The default contract was that the student would receive the highest of a sixty-forty spilt. The results would be the "contract mark." Some illustrative contracts chosen by two different cohorts of students (n=37 and n=32) were:

| | | |
|---|---|---|
| 90% term work | 24% | 0% |
| 80% term work | 5% | 46% |
| 75% term work | 0% | 2% |
| 70% term work | 0% | 16% |
| Best of 40-60 split | 71% | 35% |

Occasionally a student will select seventy, eighty or ninety percent "exam mark." The results of the longitudinal study suggesting that the contract mark is a valid and reliable single measure are published elsewhere.[4]

► *Are exam marks in chemical engineering courses correlated with the measures of problem-solving skills?*

In addressing this question, we were sensitive to the issues raised earlier[3] related to our ability to gather evidence about problem-solving skills through exams. We wanted to select conventional, three-hour written exams in courses in chemical engineering where

- *The courses were taken in the sophomore year, either concurrently or immediately after the problem-solving course, so that we could capture the performance of students as their problem-solving skills were improving.*
- *The instructors in the courses were experienced in teaching the subject discipline and were aware of, and skilled in, teaching problem solving. Coauthor CMC has been co-investigator in the MPS program since its inception in 1974 and he developed the second of four courses in the MPS program and taught a sophomore course in mass balances. PEW was co-developer of the third problem-solving course and taught both mass balances and heat transfer. JMD has used elements of the MPS program in his approach to teaching the sophomore course on energy balances. Each instructor had taught these courses often.*
- *The exams included questions requiring higher-order thinking skills.*
- *The instructors assigned part marks for the process.*

The courses that satisfied these criteria were sophomore courses in mass balances (three classes taught in different years by CMC or PEW), energy balances (three classes taught by JMD), and heat transfer (two classes taught by PEW). The eight exams included in our analysis tended to have five questions, all of which we rated as addressing greater than level three in Bloom's taxonomy.[19] The marking scheme used by each different instructor focused primarily on subject knowledge comprehension, although marks for the process were assigned where the process could be deciphered. In all courses the exam mark was an average over all questions in the exam.

For the sample set of students for the study we wanted students who were more likely to be problem solving (rather than exercise solving), who were motivated and skilled in studying for exams, and who did not suffer from exam anxiety. Data were available from 266 chemical engineering students from six cohorts of students who took the problem-solving course and the three courses. We deleted students in the following categories from the pool :

- *24% who had marks better than B- on their first year performance (to remove potential exercise-solvers from the sample)*
- *31% of the students who probably had high exam anxiety (those with scores greater than 60% on exam debilitation from the Alpert-Haber inventory,[20,21] those who had greater than thirteen on the Kellner-Sheffield inventory,[21,22] or those who had greater than 30% difference between homework scores and exam performance scores)*
- *8.5% of the students who seemed to be poorly motivated or lacked skill in studying for tests[3]*

Some students removed from the sample were in more than one category. Student A was excluded because marks were above B-, exam anxiety was greater than sixty, and the difference between exam and homework marks was greater than 30%. Student B was removed because motivation and test skills were low, exam anxiety was greater than sixty, and the difference between exam and homework marks was greater than 30%. In other words, the resulting cohort of 119 students represented students who should perform well on exams (Option 2) and whose performance on an exam should give the instructor an opportunity to assess problem-solving skill.

Not all the students in the sample took courses together, so our analysis considers the three subjects separately. Both a multivariate analysis and PCA were used in this study. The correlation between the exam marks in the mass balances course (three classes with n=47 that had two different instructors) and the marks for the journal (Option 14) ($\rho$=0.4), the self-assessment (Option 11) ($\rho$=0.24), and TEPS (Option 9) ($\rho$=0.17) is "very low" to "not significant." The correlation between the exam mark in mass balances and "gold standard" mark in problem solving is "not significant," ($\rho$=0.26).

The correlation between the exam marks in the energy balances course (three classes with n=77, one instructor) and the marks for the journal ($\rho$=0.3), the self-assessment ($\rho$=0.1), and TEPS ($\rho$=0.2) is very low to not significant. The relationship between the exam mark in heat transfer and the gold standard mark in problem solving is not significant ($\rho$=0.235). The in-depth PCA confirmed these findings.

In summary, we

- ▶ *Believe that Options 8-14 can be used as evidence of problem-solving skill.*
- ▶ *Believe this suggests that the TEPS, in the forms that we used, is not as effective a form of evidence as self-assessment and journal writing. The script, in the form that we used, was not effective.*
- ▶ *Recommend using a variety of forms of evidence. We also recommend the use of a contract mark in which the goals and criteria are published and regarding which the students have some choice in the relative weighting of the evidence.*
- ▶ *Elaborated on the challenges of using conventional subject discipline exams as a measure of problem-solving skill. Even addressing these challenges as best we could, we found no relationship between the exam marks in subject discipline courses and our measures of problem-solving skills.*
- ▶ *Found, in a very limited part of this study, that an assessor tended to focus more on the comprehension of the subject knowledge than on the problem-solving process when assigning part marks.*
- ▶ *Found, in this study, that students working under conventional exam performance conditions rarely display the problem-solving process well. The quality is improved when students are asked to write a script.*

These results contradict the assumption that student performance on the standard chemical engineering exam questions is a reliable measure of problem-solving ability.

## PROGRAM EVALUATION

Is the program developing students' skills and confidence in problem solving? The evidence for program evaluation can include

- *Assessment data of student performance (by measures such as Options 8-14)*
- *Examples of products such as student journals (Option 14)*
- *The TEPS complete with goals and criteria*
- *The marking templates used in Options 2 and 3[3]*

Other useful data to benchmark performance include pre- and post results from commercially available tests (that have been developed and validated) and exit surveys of graduates.

### *Published tests*

The criteria for selecting tests might be

- *The test should relate to attitude or skill in problem solving*
- *Preference would be given to tests that are available at minimum cost*
- *Benchmark data should be available for comparable populations*

The book *Tests in Print*[23] lists a wide variety of psychologically proven tests that might add insight into skill in problem solving. Some tests assess "skills" and ability to apply the problem-solving skill, while others assess "attitude and/or confidence" in one's skill in problem solving. Other options with some benchmark data are published.[24] Examples of such tests include the Torrance Test of Creativity,[25] Basadur's Ideation,[26] Heppner's Problem-Solving Inventory,[27] and the Billings-Moos test.[28] The Torrance test for skill in creativity[25] has been used by other investigators.[29] The test asks participants to solve content-independent problems whereas we preferred to design TEPS, like Problem 4 in Table 1, that used chemical engineering problems.

Basadur's test[26] relatd?∏to attitude toward creativity through two measurers, willingness to postpone judgment and attitude toward "ideation." Values for the former range from 8 to 45 and a low number is desired. The measure of ideation has values from 6 to 30 with a large value as the target. The Heppner Problem-Solving Inventory[27] measures attitude toward problem solving.

The Billings-Moos test[28] asks people to identify a particularly difficult problem they addressed in the past year and, in the context of solving that problem, to rate the degree to which they applied the 32 listed tactics. From this the test developers have extracted statistically significant measures of problem-solving skill and of avoidance to solving the problem. Results of pre- and post-test results for the MPS program, and for some control groups, are given elsewhere.[4,24]

### *Use of Exit Surveys*

Queen's University has developed an excellent survey questionnaire.[30] The key question is, "Please indicate the degree to which your education at Queen's contributed to your learning and development in each of the following areas: "creative thinking," "problem-solving skills," etc. Rating for answers ranges from one (very little) to five (a great deal).

Data for engineering from 1994 are available as a benchmark.[30] Here are data for the alumni response to the question regarding development of creativity. The numerals in square brackets represent the standard deviation. The first four results are from graduates at McMaster University. Details of the unique five-year Engineering and Society program are available.[31] The last result is from all branches of engineering at Queen's University.[30]

- *For graduates of the MPS program:[13] 4.48 [0.82]*
- *For graduates of materials engineering: 3.57 [0.79]*
- *The contribution of the "home" department for graduates of the Engineering and Society Program: 2.97 [0.97][31]*
- *The contributions of the Engineering and Society Program courses for graduates of that program: 4.42 [0.73][31]*
- *For graduates of all branches of engineering: 3.27*

Here are some data for the responses to the question on development of problem-solving skills.

- *For graduates of the MPS program:[13] 4.92 [0.28]*
- *For graduates of materials engineering: 4.43 [0.53]*
- *The contribution of the "home" department for graduates of the Engineering and Society Program:[31] 4.28 [0.70]*
- *The contributions of the Engineering and Society Program courses for graduates of that program:[31] 3.56 [0.88]*
- *For graduates of all branches of engineering: 4.49*

Such data are useful evidence from recent graduates about the effectiveness of programs.

## SUMMARY

Assessment is based on evidence of performance. Fifteen options for gathering evidence were described, seven in Part I[3] and eight in this paper. Data from the students' performance in the McMaster Problem-Solving program were used to suggest that

- ▶ *Tests and Examinations of Problem-Solving Skills (TEPS) can be developed that are consistent with the published goals and criteria. These results, however, are least highly related among the measures.*
- ▶ *Self-assessment and journal writing were highly-related measures of problem-solving skill.*
- ▶ *The use of a written script gathered for the sole purpose of displaying ten to twenty minutes of thought processes used in problem solving was not a statistically significant measure.*
- ▶ *We recommend using a variety of forms of evidence. We also recommend the use of a contract mark in which goals and*

*criteria are published and regarding which the students have some choice in the relative weighting of the evidence.*

- *The results from exam scripts in subject discipline—where correctness of the answer, subject knowledge, and problem solving were being marked (Option 2[3])—gave one form of evidence, but it should not be used as the sole form of evidence. The challenges of using conventional exams as a measure of problem-solving skills were described in Part I.[3] Even addressing these challenges as best we could, we found no relationship between the exam marks in chemical engineering subject discipline courses and any of our measures of problem-solving skills.*

For program evaluation, the results from such published tests as Basadur's attitude toward ideation, Heppner's measure of confidence in problem solving, and Billings-Moos' measure of problem-solving performance and avoidance are useful for pre- and post data about the effectiveness of programs. We recommend that instruments be used, including the items on exit surveys, where benchmark data are available for comparable populations.

## ACKNOWLEDGMENTS

We thank the reviewers for their useful suggestions.

## REFERENCES

1. Rugarcia, A., R.M. Felder, D.R. Woods, and J.E. Stice, "The Future of Engineering Education, Part I: A Vision for a New Century," *Chem. Eng. Ed.,* **34**(1), 17 (2000)
2. ABET 2000 <http://www.abet.org> May (1999)
3. Woods, D.R., *et al.*, "Assessing Problem-Solving Skills, Part I: The Context for Assessment," *Chem. Eng. Ed.,* **35**(4), (2001)
4. Woods, D.R., *et al.*, "Developing Problem-Solving Skill: The McMaster Problem Solving Program," *J. Eng. Ed.*, **April**, 75 (1997) or <http://www.chemeng.mcmaster.ca/innov1.htm> and click on MPS for a summary of the research findings and some details for the many units
5. Felder, R., and R. Rousseau, *Elementary Principles of Chemical Process*, 3rd ed., John Wiley and Sons, New York, NY (2000)
6. Angelo, T.A., and K.P. Cross, *Classroom Assessment Techniques*, 2nd ed., Jossey Bass, San Francisco, CA (1993)
7. Bodner, G.M., "Toward a Uniform View of Problem Solving: A View from Chemistry," in *Toward a Unified Theory of Problem Solving,* M.U. Smith, Ed., Lawrence Erlbaum Associates, Hillsdale (1990)
8. Woods D.R., *Problem-Based Learning: How to Gain the Most from PBL*, Woods Publisher, Waterdown, ON, Canada, distributed by McMaster University Bookstore, Hamilton, ON (1994)
9. Woods, D.R., R.M. Felder, A. Rugarcia, and J.E. Stice, "The Future of Engineering Education, Part III: Developing Critical Skills," *Chem. Eng. Ed.*, **34**(2), 108 (2000)
10. Target skills for problem solving, <http://www.chemeng.mcmaster.ca/innov1.htm> and click on MPS and then on target skills for each MPS unit
11. Alverno College, *Student Assessment-as-Learning at Alverno College*, Alverno College Publications, Milwaukee, WI (1994)
12. Woods, D.R., et al., "Self-Assessment in the Context of the McMaster Problem-Solving Program," *Evaluation and Assessment in Higher Education*, **12**(2), 107 (1988)
13. Kimbell, R., et al., *The Assessment of Performance in Design and Technology,* Assessment of Performance Unit, National Examination and Assessment Council, Newcombe House, London, UK (1991)
14. Schon, Donald, *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*, Jossey-Bass, San Francisco, CA (1987)
15. Woods, D.R., and H. Sheardown, *Reflecting and Journal Writing*, Department of Chemical Engineering, McMaster University, Hamilton, ON (1999)
16. "MPS 3, Self-Assessment," *PS News*, **84**, 12 (1993). "MPS 5: Stress Management," *PS News*, **85**, 14 (1993). "MPS 6: Analysis, Classification," *PS News*, **92**, 14 (1994). "MPS 11: Personal Uniqueness," *PS News*, **93**, 17 (1994). "MPS 17: Time Management," *PS News*, **97**, 11 (1994)
17. Woods, D.R., "How to Assess," Chapter F in *Problem-Based Learning: Resources to Gain the Most from PBL*, 2nd ed., Woods, Waterdown (1999)
18. Box, G.E.P., W.G. Hunter, and S.J. Hunter, *Statistics for Engineers*, John Wiley and Sons, New York, NY (1978)
19. Bloom, B.S., *et al., Taxonomy of Educational Objectives Handbook I: The Cognitive Domain*, Addison Wesley, New York, NY (1954)
20. Alpert, R., and R.N. Haber, "Anxiety in Academic Achievement Situations," *J. Ab. and Soc. Psych.*, **61**(2), 207 (1960)
21. Roney, S.D., and D.R. Woods, "Ideas to Minimize Exam Anxiety," *J. Eng. Ed.*, accepted
22. Kellner, R., and B.F. Sheffield, "A Self-Rating Scale of Distress," *Psych. Med.*, **3**, 88 (1973)
23. Buros, O.K., *The Mental Measurements Yearbook*, Gryphon Press, Highland Park, NJ (each year) and Murphy, L.L., *Tests in Print IV: An Index to Tests, Test Reviews and Literature on Specific Tests*, Buros Institute of Mental Measurements, University of Nebraska-Lincoln, Lincoln, NB (1994)
24. Woods, D.R., "How to Select Instruments for Assessment and Program Evaluation," Chapter E in *Problem-Based Learning: Resources to Gain the Most from PBL,* downloadable from <http://www.chemeng.mcmaster.ca/innov1.htm> and click on PBL and download from the resources book
25. Torrance, E.P., *Torrance Test of Creative Thinking: Directions Manual and Scoring Guide*, Prentice-Hall, Englewood Cliffs, NJ (1974)
26. Basadur, M., and C.T. Finkbeiner, "Measuring Preference for Ideation in Creative Problem-Solving Training," *J. Applied Behav. Sci.*, **21**(1), 37 (1985)
27. Heppner, P.P., *The PSI Manual*, University of Missouri-Columbia, Columbia, MO (1986)
28. Billings, A.G., and R.H. Moos, "The Role of Coping Responses and Social Resources in Attenuating the Stress of Life Events," *J. of Behav. Med.*, **4**(2), 139 (1981)
29. Kvashny, A., and J.T. Sears, "Measurement of Creativity," *Eng. Ed.*, **68**, 269 (1978)
30. Office of the Registrar, *Undergraduate Learning Experiences at Queen's: Results from the Exit Poll*, Queen's University, Kingston, ON, Canada (each year since 1994)
31. Hudspith, R.C., *Improving the Bridge: Making Engineering Education Broader and Longer*, Engineering and Society Program, McMaster University, Hamilton, ON ❐