Performance Criteria for Quality in Problem Solving

Donald Elger[*], Jennifer Beller[+], Steven Beyerlein[*], Barbara Williams[#]

[*]Mechanical Engineering, University of Idaho, Moscow, ID./ [+]Ed. Leadership & Counseling Psych., Washington State Univ., Pullman, WA./ [#]Ag. and Bio. Systems Engineering, University of Idaho, Moscow, ID.

Abstract

Many educators believe that our educational system teaches students to solve problems using cook-book procedures, instead of teaching students how to solve problems in an effective way. In trying to raise issues of teaching and learning of problem solving, we have encountered significant resistance from both teachers ("*I need to cover content"*) and students ("*just tell me how to get the right answer"*). To address these problems, it is important to have a clear understanding of what quality looks like. Thus, we developed criteria for performance (see Appendix A), that are a set of 30 specific objectives that can be observed and measured as students engage in relevant tasks. Our work is limited in scope to problem solving that involves engineering calculations that are based on mathematical representations of scientific concepts. Our context is those engineering classes that involve significant amounts of engineering analysis.

To understand present conditions, we designed a pilot (first iteration) survey to assess student and faculty beliefs about 8 of the 30 objectives. The survey provided a concrete example (scenario) of each specific objective (or performance) considered. Each scenario was assessed by asking a set of four focus questions. In simple terms, these focus questions are (a) Is this objective emphasized in engineering science courses? (b) Is this objective important? (c) Can students realistically develop this performance? and (d) What is the present level of student performance? Reliability of the survey was estimated by using statistical analysis with the Cronbach-Alpha metric. Logical validity was established through the use of expert analysis of questions relative to the theoretical construct**.**

The survey was completed by 66 students and 15 faculty members at our institution. For each objective measured, the survey data showed similar trends that may be summarized as follows. On average, student and faculty believe (a) the objective is emphasized in engineering science courses, (b) the objective is important, (c) students can develop the requisite performance in the context of an engineering science course, and (d) present performance levels are satisfactory. These results provide evidence that performance criteria developed in this study are aligned with professor and student perceptions of quality. These results also provide a plausible explanation for the resistance that we have encountered when we have raised issues associated with teaching and learning of problem solving. Both professors and students (on average) believe that present educational practices are producing satisfactory outcomes—thus, there is no compelling need for change and efforts to promote change prompt opposition. We hypothesize that the root cause of the problem is related to assessment practices. Because most professors have had little opportunity to learn effective assessment methodologies, they tend to reach invalid conclusions about students' abilities.

Introduction

Many educators believe that our educational system teaches students to memorize canned solutions and to solve problems by a "plug-and-chug" approach, rather than by understanding concepts. Thus, a central problem facing engineering educators is to identify effective means to improve the problem solving abilities of our students. However, we have observed significant resistance to teaching and learning of problem solving skills. Many students, especially those at lower levels of cognitive complexity, state "*tell me how to get the right answer and quit wasting my time*." Many professors are similarly resistant--they state, "*I need to cover the content on my syllabus.*" While this transmission model ("covering content") is pervasive in engineering education, we believe that a new vision of teaching is needed.

Our vision of teaching is described by several outstanding ski coaches. Tejada-Flores[12] states *"Against all conventional wisdom, I claimed that most skiers—virtually all skiers—could ski like experts, and the only reason they did not was that they did not know how expert skiers did it. I also claimed that expert skiing was not simply an improved, polished version of intermediate skiing. It was something else; not harder, just different."* Another outstanding coach (Harb[5]) describes how most skiers have learned "dead-end" skills, which are skills such as skidding one's skis that lead to a plateau, not to life-long learning. This teacher espouses a philosophy that teaches each skier life-long skills (i.e. the skills of the expert) regardless of whether the skier is a beginner, an intermediate, a racer or an instructor. The point of the analogy is that effective teachers avoid practices that reinforce dead-end skills and embrace practices that reinforce life-long (i.e. expert) skills.

We believe that quality in problem solving is defined by those approaches that are truly effective--that is, the approaches used by experts. The present work has two main goals:

- Create an operational definition of effective problem solving. That is, define quality in problem solving by listing objectives that can be observed and measured, thus creating goals for teaching, learning and assessing.
- Gather data from educators and students that provide insights about present beliefs. For example, do professors and students agree with or disagree with our concepts of quality? Do professors emphasize learning of the objectives? What do students believe?

Regarding the scope of study, we focus on engineering analysis, which we define to be reasoning and calculations that are performed using mathematical representations of scientific concepts. Our context is teaching and learning in engineering classes that emphasize analysis and calculations. We label such classes as engineering science classes--in our curriculum, engineering science classes include most of the classes taught by engineering professors, except those specifically designated as lab or design classes. Representative examples include Statics, Circuits, Heat Transfer and many electives at the senior and graduate levels.

Our notion of problem solving follows Bransford[2]: "*A problem exists when there is a discrepancy between an initial state and the goal state and there is no ready-made solution for the problem solver. The initial state is where you are as you begin the problem; the goal state is where you want to end up when you solve it.*" In other words, problem solving is the complex set of actions taken by an engineer as they navigate from an initial state to a goal state on an **unfamiliar problem**.

Literature Review

Quality in problem solving is an ill-defined concept. Are students learning problem solving or are they learning how to repeat memorized procedures? How can we measure growth in problem solving ability? How do we know if a method of teaching truly improves our students' problem-solving abilities? Do our students believe that effective problem solving is the same thing that we believe? Essential to social research is the need to provide concrete definitions of abstract constructs (Trochim[13]). Such a definition is known as an operational definition. To define a term operationally is to specify how this term will be measured. A well-crafted operational definition is specific and unambiguous, thereby facilitating a common understanding of what one means when they use the term. In conclusion, the use of an operational definition greatly improves research on an abstract construct like quality in problem solving.

Before considering quality, it is useful to review issues with student learning. Resnik[9] describes problem solving by students as manipulation of symbols and equations, with very little understanding of the underlying concepts and meanings. Schoenfeld[10] stated that "*Most textbooks present "problems" that can be solved without thinking about the underlying mathematics, but by blindly applying the procedures that have just been studied. Indeed, typical classroom instruction subvert understanding even further by providing methods for solving problems that allow students to answer problems correctly, without making an attempt to understand them."* Woods[25] notes that during a four-year degree program, engineering students observe professors work 1000 example problems, or more, and the students themselves solve more than 3000 problems. However, Woods[25] reported that the students "*show negligible improvement*" in problem solving skills--meaning that "*if they were given a related but different problem situation, they were not able to bring any new thinking or process skills to bear."*

Our operational definition of quality in problem solving is founded on knowledge of how experts solve problems. Wankat and Oreovicz[22] present an excellent review—they provide many details and summarize the finding with a side-by-side comparison of novice and expert performances. Resnik[9] (paraphrasing Larkin et al.[6]) presents a lucid summary of expert performance "*Recent research in science problem solving, for example, shows that experts do not respond to problems as they are presented—writing equations for every relationship described and then using routine procedures for manipulating equations. Instead, they reinterpret the problems, recasting them in terms of general scientific principles until the solutions become almost self-evident."* Experts see each new problem through the lens of scientific concepts and they develop a meaningful interpretation, a process that cognitive psychologists call creating an internal representation. The process of representation also involves finding and evaluating information—what is relevant and what is not relevant, and the degree to which this information is reliable (Matlin[7]). Attendant with the process of creating an internal representation (a visualization within one's head) is the process of transferring this internal imagery onto paper, thereby creating an external representation. External representations provide an effective means to deal with the limitations of short-term memory. Experts use external representations to keep track of the quantity of information that often accompanies a complex problem (Bransford and Stein[2]).

Two hallmarks of expert thinking are described by the concepts of schemas and metacognition (Pellegrino et al.[8]). Metacognition or "think about one's own thought processes" involves knowledge, awareness and control of one's thinking. Metacognition involves an active and purposeful monitoring of one's problem solving process. Schema refers to the way people organize knowledge in long-term memory. Pellegrino et al. report "*experts in a subject matter domain typically organize factual and procedural knowledge into schemas that support pattern recognition and the rapid retrieval and application of knowledge.*" Schema can involve discipline knowledge. The schema of disciplinary knowledge is a mental structure (think of a spider web) that connects or link relevant concepts and facts. The structure is hierarchical, with overarching concepts at the top of the hierarchy, secondary concepts in the middle and facts/details towards the lower part of the hierarchy.

Schema can involve procedural knowledge; that is, a schema can organize common patterns that facilitate problem solving. A few engineering educators have recognized the importance of schema for procedural knowledge. Wales et al.[14-21] used a schema (named the professional decision making process) and a teaching method (guided design) to teach thinking skills to freshman students. To assess the effect of guided design, Wales[15] analyzed ten years of data (5 year pre-guided design and 5 years post). The data showed that when thinking skills were taught, the number of students who ultimately graduated increased by 32%. Also, the average GPA at graduation was up by 25%. While Wales[15] was clever with percentages, the results of this study indicate a positive effect.

Woods et al.[23-27] have spent 20+ years developing a schema to organize and teach problem solving. Their schema, called the McMaster Problem Solving (MPS) program, represents problem solving using a hierarchical structure in which the big picture ideas (i.e. the stages) are at the top and the details (specific skills and attitudes) are associated with each stage. There is strong evidence that the MPS program has improved outcomes.

A Method for Defining Quality in Problem Solving

For the past fifteen years, we have worked on developing a process for teaching quality in problem solving (Elger et al.[4]). During fall semester 2002, we decided to put our understandings about quality on paper. To reach this goal, we selected a method from the formative assessment literature (Arter and McTighe[1]). This method is summarized below.

1. Students worked problems. Many of the problems were difficult and unfamiliar to the students. Much of the teaching process involved active learning in a team environment.
2. Students were asked to purposefully review their performance. While each review was different, the basic aim was to have students identify (a) what was strong or effective about their problem solving approach, and (b) specific ways to improve their approach.
3. Using a variety of methods, students were given individualized feedback on their performances.
4. As the course progressed, steps 1 to 3 were repeated in many different contexts.
5. To form the initial draft of the operational definition, the collection of student responses was organized into 6 main categories, with each category containing short statements (objectives) that describe specific details of quality.
6. The operational definition was improved by adding knowledge from our experiences and from the literature. We assessed each objective using the following criteria:

a. **Clarity**. Is this objective clear? Is this objective specific? Can this objective be measured?
b. **Essential**. Is this objective needed to describe quality? If this objective is not met, will quality be decreased?
c. **Connections**. Will the language appeal to most engineers? Is the language consistent with the research literature? Does the language communicate the idea of a community of engineers?
d. **Attainable**. Is this objective attainable by students?
e. **Goal; not a Method**. Is this objective a result and not a method? Is this objective independent of a specific engineering class?

Discussion of the Operational Definition

Description of the Operational Definition

The operational definition is presented in Appendix A. It is organized into six main categories (traits). For example, the second trait is "Scientific Concepts and Ongoing Learning." The meaning of each trait is amplified by focus questions. Each trait is separated into measurable units labeled as objectives. An example of an objective is

> **Scientific Concepts**. Engineers interpret the world using scientific concepts such as Ohm's law, equilibrium, and the ideal gas law. When faced with an unfamiliar problem or situation, engineers use scientific concepts to guide their actions—that is they use concepts to create understanding, to make predictions, to make decisions, to solve problems and to perform other similar actions.

Each objective is given a label and a text description. Each objective is designed for measurement. For example, one way to measure the "scientific concepts" objective would be to give students an unfamiliar problem and then interview them. For example, one can ask students how they would figure out the rotation rate of a yo-yo that is dropped and allowed to spin freely (i.e. to "sleep"). Students who are far along (i.e. a high performance level, meaning they solve problems like engineers) will likely apply scientific concepts—e.g., they might balance the change in gravitational potential energy with the change in rotational kinetic energy and then include work done by the human hand at the start of the motion. Students who are not far along (low performance level) will give trite answers, usually not involving scientific concepts.

Each objective is written in language that reflects the idea of community. For example, notice the wording: *Engineers interpret… engineers use …* This wording is intended to invite students to join our community, and suggest that successful practioners follow common practices. We wish to avoid reinforcing the common student view that "*We are playing a game and my first task in this game is to figure out what this professor wants ….*"

Limitations of the Operational Definition

The operational definition is not intended to be universally acceptable. Professors will find omissions and ideas that they disagree with. This is anticipated and acceptable because we are striving to describe a complex and abstract performance. Also, people are different and one size does not fit all.

The operational definition attempts to balance brevity and clarity with detail. We chose to restrict the amount of detail by omitting information about how each objective can be measured and by omitting information about levels of performance. We also tried to make the objectives specific enough to be useful. For example, an objective written like "engineers communicate clearly," is too broad.

The operational definition, if it is simply handed out to students or presented in a lecture format, will have no value. Problem solving, like other performances (think of skiing), cannot be effectively learned solely by listening to someone talk about it.

Design of an Instrument to Measure Beliefs about Problem Solving

Development of Four Focus Questions

Over the last 15 years, we have made many observations of engineering performance. Collectively, these observations have led us to conclude that engineering students, on average, are not effectively learning many of the performances that are described in the operational definition. This experience led us to a central question: *Why is it that students cannot demonstrate strong performances on objectives that are prized by members of the engineering community?* This question led us to ask more specific questions (focus questions) that guide the present study:

a. **Level of Emphasis**. Do people believe that present engineering science courses emphasize learning of the objectives that are in the operational definition?
b. **Importance**. Should present engineering science courses result in the learning described in the objectives? That is, are these objectives important?
c. **Attainable**. Are the objectives attainable? That is, is it realistic to expect that students can learn the performances described by the objective?
d. **Level of Performance**. What are the beliefs about how well students can demonstrate performances described by the objectives?

Design of a Survey Instrument

To provide insights that address the four focus questions, we considered a variety of methods, both qualitative and quantitative. After considering alternatives, we selected a survey method, based primarily on cost versus benefit issues. Also, we decided to create a pilot survey, which is a first iteration that is intended to guide the design of a more detailed survey.

To design the survey, we decided to provide concrete examples. These examples are described using "*scenarios*," that are short vignettes that illuminate the performance of interest. For example, Table 1 presents a scenario to examine metacognition. The scenario presents a concrete example in which a student has demonstrated the performance. The response block presents the four focus questions.

In the design of each scenario, we purposefully attended to several important details. To avoid issues with reading comprehension, we crafted prose at a simple reading level. To avoid confusion and different interpretations, we avoided jargon such as metacognition. To create ownership, we personalized questions using words like "you" and "your."

We designed separate surveys for students and professors. Notice in Table 1 that the wording of the focus questions is relevant to the perspective of a professor. Table 2 shows how we modified the wording of the focus questions for students. The only significant difference between the faculty and student surveys is the wording of the focus questions.

To keep the size of the survey manageable, we elected to assess 8 of the 30 objectives. These eight objectives and corresponding scenarios are

1.) **Scientific Concepts**: Assessed using Scenarios 1, 5 and 10
2.) **Units**: Scenario 2
3.) **Schema for Procedural Knowledge**: Scenario 3
4.) **Review**: Scenario 4
5.) **Metacognition**: Scenario 6
6.) **Representation**: Scenario 7.
7.) **Documentation**: Scenario 8
8.) **Verbal Explanation of a Scientific Concept**: Scenario 9

In the above list, "Scientific Concepts," was assessed using three scenarios in order to provide additional data for estimating reliability.

Scenario selection was based on our prior experiences. In particular, we selected specific performances that have been problematic for students. For example, we have used several design projects that involve a battery (see scenario 6 in Table 1) and noted that most junior level mechanical engineering students fail dismally (unless we tell them how) in their attempts to perform calculations that will allow them to size a battery to meet power requirements and life-time needs. Similarly, the "yo-yo" problem (scenario 10 in Table 2) came as a technical question during a job interview. A group of our brightest graduating seniors were discussing the problem—they were struggling, and they considered the problem quite difficult. Subsequent use of this question has revealed that many students will not even consider (unless prompted) applying scientific concepts to the "yo-yo" problem. Instead, they reply with answers such as "I have no idea" and "measure it."

To test the survey prior to implementation, we used about nine beta testers split into three groups. The first group, non-engineers at both the pre- and post-undergraduate level, was used to assess readability and interpretation. The second and third groups, engineering students and faculty, were used to provide disciplinary feedback. Based on feedback from the beta testers, the survey was modified significantly.

To facilitate data acquisition and processing, we designed each survey for web-based administration. The completed surveys are online[28, 29]. Each survey has approximately 44 questions (4 questions to gather information on background of the respondent; and ten scenarios, each containing the four focus questions).

The population used for the survey was comprised of most of the engineering faculty and undergraduate students at the University of Idaho in Moscow. The population size is approximately 100 faculty and 1,500 students. To gather responses, we sent a brief email to faculty and about half of our students. This email explained the purpose of the survey and provided a hyperlink to the web where the survey was located. This email was sent five days prior to the start of Spring Semester 2003. When a person responded to the survey, we did not record names of individual respondents.

Table 1 Example of a Scenario From the Survey (questions worded for a professor respondent)
Scenario Six. Jane is talking about how she will create a math model of a D-cell battery. "I will begin by finding information so that I can figure out how batteries work. Next, I will figure out how to model the battery using concepts from my circuits class. Then I will set up some basic equations and check my results with a simple experiment." **Outcome or Performance.** The student is keenly aware of their personal approach to solving problems and can verbalize this process. **Questions for Scenario Six.** **a**.) To what extent have you emphasized this outcome in engineering science courses? O Major emphasis O Significant O Medium O Slight O No Emphasis **b**.) Engineering science courses should result in this outcome. O Strongly Agree O Agree O Neutral O Disagree O Strongly Disagree **c**.) Students such as those in my classes can learn this performance. O Strongly Agree O Agree O Neutral O Disagree O Strongly Disagree **d**.) On average, how well do students who have completed their junior-year engineering science courses carry out this performance? O Masterful O Very Well O Satisfactory O Poor O Incompetent

Table 2 Example of a Scenario From the Survey (questions worded for a student respondent)
Scenario Ten. Michael has completed his sophomore year. During a job interview, he is asked how to figure out how fast a toy "yo-yo" will spin if it is allowed to drop and then spin freely ("sleep"). Michael thinks back to his dynamics class and then describes how he would balance the gravitational potential energy (before the yo-yo is dropped) with the rotational kinetic energy (when the yo-yo is spinning). **Outcome or Performance**. Faced with a problem that is unlike those they have solved in the past, the student applies scientific concepts to reason out a solution. **Questions for Scenario Ten.** **a**.) To what extent have your Professors emphasized this outcome? O Major emphasis O Significant O Medium O Slight O No Emphasis **b**.) Engineering science courses should result in this outcome. O Strongly Agree O Agree O Neutral O Disagree O Strongly Disagree **c**.) Students such as myself can learn this performance. O Strongly Agree O Agree O Neutral O Disagree O Strongly Disagree **d**.) On average, how well do your engineering peers carry out this performance? O Masterful O Very Well O Satisfactory O Poor O Incompetent

Results of the Survey

Response

The survey was completed by 66 students (51 mechanical, 13 civil, 1 electrical, 1 unknown), for a response rate of about 8% of students we contacted. The survey was completed by 15 faculty members (7 mechanical, 3 civil, 3 chemical, 1 agricultural, 1 other), for a response rate of about 15% of the college of engineering faculty.

Data

Fig. 1 presents data for scenarios 6 and 10. On the horizontal axis is the question number (S6a, F10b, etc.), where the capital S (e.g. S6a) denotes a student response, and the capital F (e.g. F10a) denotes a faculty response. Letters a to d correspond to the four focus questions. Thus, question S6a involves the first focus question, "is this objective presently emphasized in engineering science courses?" Similarly, question S6c involves the third focus question.

On the vertical axis of Fig. 1 is the level of response, averaged for all respondents. The numerical coding scheme assigned a value of 5 to the left-most selection on the survey and a value of 1 to the right-most selection. Thus, in Fig. 1 the response to question S6a has a numerical value of about 2.8, indicating that students believe that professors emphasize metacognition with a degree that is slightly less than "medium."

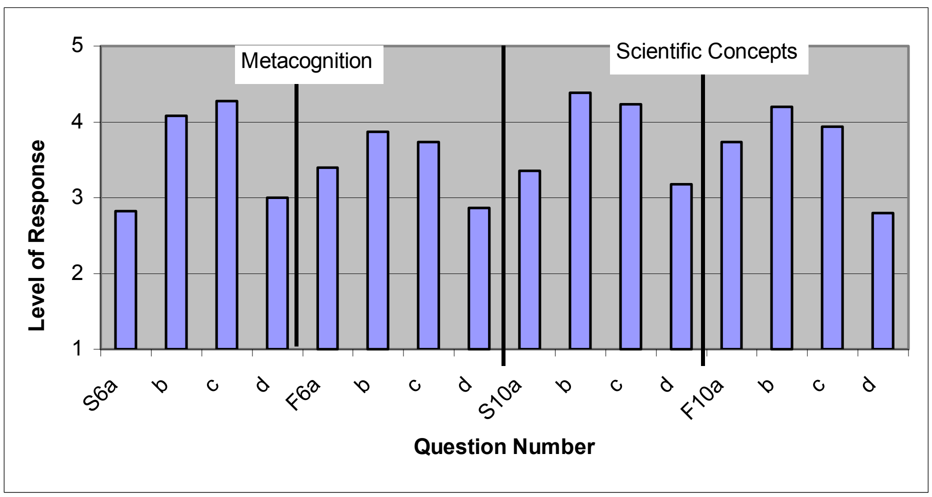

Figure 1 Average response values for scenario 6 (metacognition) and scenario 10 (scientific concepts).

Fig. 1 allows side-by-side comparison of student and faculty response to scenario 6 (metacognition) and 10 (scientific concepts). To interpret Fig. 1, consider the data for metacognition:

6a. **Level of Emphasis**. Both students and faculty believe that metacognition is emphasized. However, faculty members report a higher level of emphasis.
6b. **Importance**. Both students and faculty believe that metacognition is important. However, students rate the level of importance as higher
6c. **Attainable**. Both students and faculty believe that students can learn to be metacognitive. However, students report a higher level of attainability.

6d. **Level of Performance**. Both students and faculty report that present performance of metacognition is basically satisfactory. However, students report a slightly higher level of performance.

Figs. 2 and 3 show the complete data set for faculty and students, respectively.

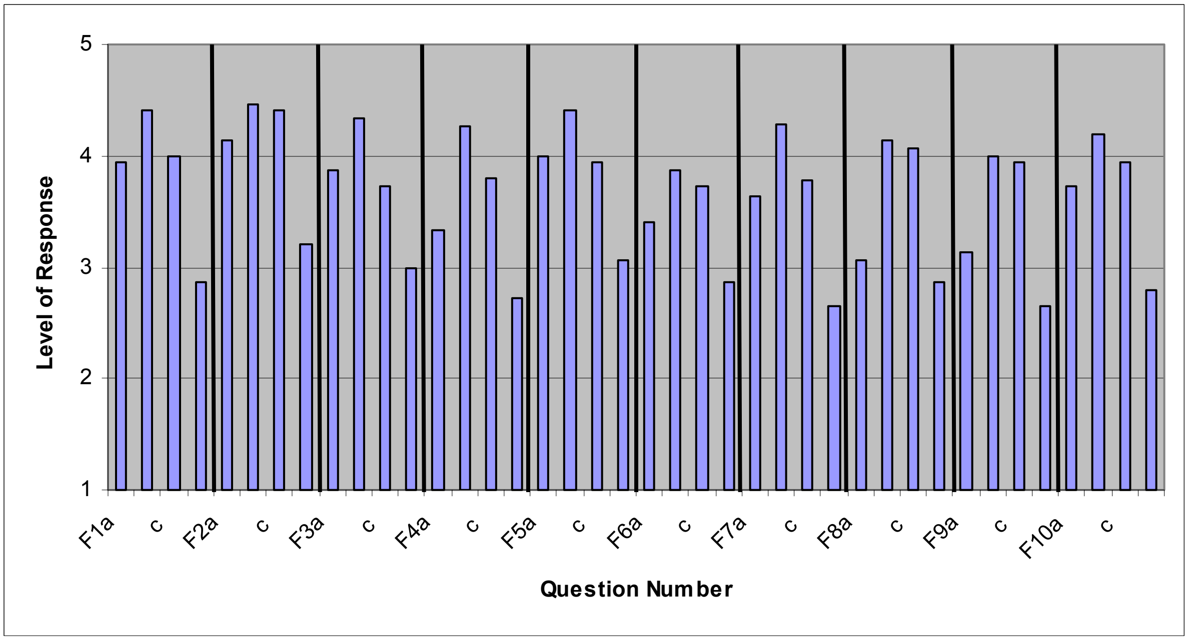

Figure 2 Complete data set for all scenarios (averages from faculty respondents)

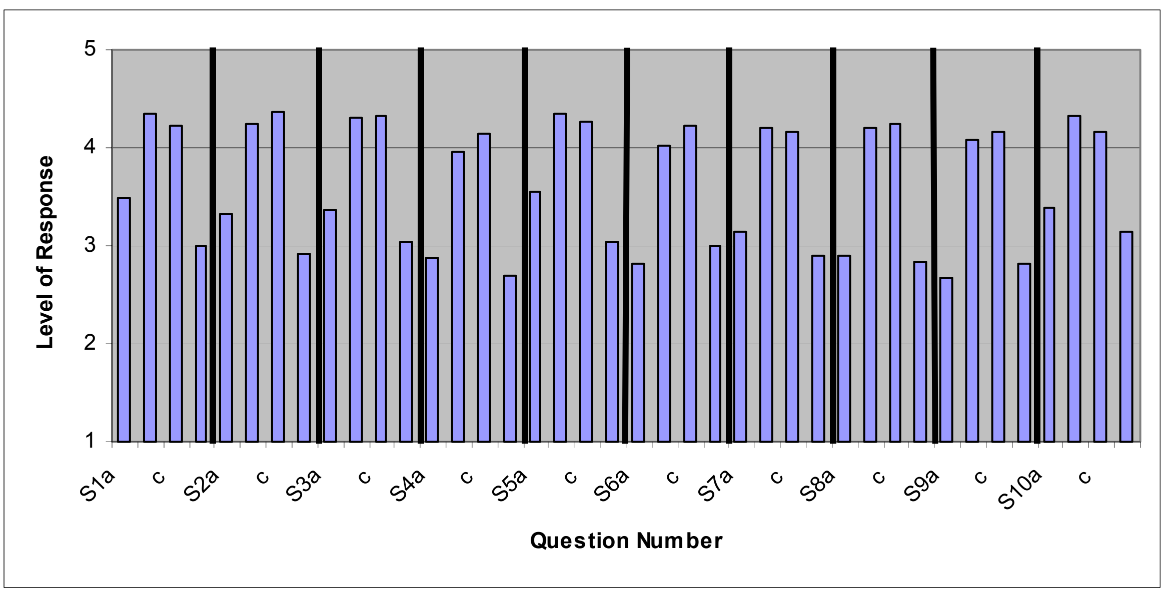

Figure 3 Complete data set for all scenarios (averages from student respondents)

Discussion of the Survey Results

Overarching Trends

The survey reveals some "big-picture" beliefs about engineering science classes.

a. **Level of Emphasis**: On average, students and faculty believe that each objective is emphasized (about a medium degree of emphasis).
b. **Importance**: On average, students and faculty agree that each objective is important.
c. **Attainable**. On average, students and faculty agree that each objective is attainable.
d. **Present Level of Performance**. On average, students and faculty believe that present levels of performance are nearly satisfactory.

Taken together, these results suggest that both students and faculty believe that present engineering science classes are meeting the basic goal of developing competent problem solvers.

Reliability

Statistical analysis was used to estimate reliability. This estimate was made using SPSS (Statistical Package for the Social Sciences Version 11.5). In particular, we applied the Chronbach Alpha procedure to analyze the reliability of each focus question on the survey.

Table 3 Summary of statistical data for the survey

Focus Question	**Cronbach Alpha**	
	Faculty	Students
a.) Level of Emphasis?	0.58	0.83
b.) Importance?	0.71	0.86
c.) Attainable?	0.85	0.87
d.) Level of Performance?	0.93	0.78

To establish reliability for the type of survey we used, Cronbach Alpha should be 0.70 or better, with values from 0.80s to 0.90 considered the best. As shown in Table 3, only one statistic fell below 0.70, which provides evidence that the data are reliable.

There are several interesting contrasts evident in Table 3. On the level of emphasis question, the faculty Cronbach Alpha is low (0.58), while the student value is high (0.83). Examination of the frequencies shows that faculty members are much less consistent in their belief that this is emphasized in their courses. Similarly on the importance question, faculty Chronbach Alphas were low (0.71) with student values high (0.86). In contrast, faculty are much more consistent in their belief that students have an acceptable level of performance (0.93) while students are less consistent (0.78).

Validity

Logical validity was established through use of engineering and measurement experts. These experts analyzed scenarios and questions relative to the theoretical constructs and where needed, questions were rewritten or reworded.

During the experiment design, we worked hard to reduce or eliminate many threats that could affect the instrument's validity. Because the instrument is in its first iteration, validity questions remain and will be addressed in future iterations. Our assessment and thoughts of the three major threats to validity (in order of priority) are summarized below.

1. **Threat One--Emotional Reactions.** Faculty members take tremendous pride in and ownership of their teaching methods. Perhaps some of the survey questions caused faculty to believe that their approaches to teaching were being questioned. If a faculty member made this inference, perhaps this person became emotional and this reaction clouded their ability to objectively reflect on the question being asked.
2. **Threat Two—Detailed Context.** During the course of this study, we discovered that faculty members have many different definitions of common terms such as problem solving, engineering science courses and engineering analysis. In general, faculty members seem to want a great deal of contextual information before making a judgment. Perhaps we did not give the faculty members enough details.
3. **Threat Three—Details of the Survey Design.** There are a variety of details that influence validity. Some of these details that we address in future studies are social desirability, the overall pattern of the questions (ours are highly patterned), and lack of a response choice to indicate that the respondent is unsure of an answer.

<u>Design of the Next Generation Survey</u>

The present survey is a pilot study--future iterations are planned. To improve the survey, we have several ideas. To improve reliability and validity, additional testing with users will be implemented. In particular, we want to observe users as they take the survey, and then question the users in order to understand their reactions and interpretations. In order to improve reliability consistent with other questions, the "level of emphasis" and "importance" questions for faculty will be re-examined and reworked as necessary. To improve the response rate, we will use several prompts via email and we will seek to enlist administrative support (e.g. an email from the Dean of the College) to encourage faculty participation. Similarly, we will enlist professors to encourage student support. To address issues of sample size and diversity within the sample, we will plan to broaden our population to include a more representative sample of disciplines and perhaps several (two to four) institutions.

<u>Interpretation of the Data</u>

The results of the survey provide evidence that the operational definition of quality in problem solving (i.e. the objectives in appendix A) is aligned with professor and student perceptions. In particular, for each objective studied, both student and faculty data consistently support the idea that engineering science courses should result in attainment of this objective. Similarly, the data consistently supported the notion that students can develop this performance in the context of an engineering science class. This finding is very valuable because clear objectives that are agreed upon by various stakeholders are essential for many purposes such as assessment, curriculum design, and collaborative efforts among professors.

Our most interesting and challenging interpretive task arose because our starting hypothesis was not confirmed! In particular, one of our initial beliefs was that faculty and perhaps students would recognize that most of the objectives are not being effectively learned. We also thought that students and perhaps faculty would recognize that most of the objectives are not being emphasized (e.g. How many professors focus on teaching students how to be metacognitive?

How many professors assess metacognition? How many professors grade students based on their performance of metacognition?).

Of course, we spent significant time trying to understand why the survey came out the way it did. Three possible explanations (that we rejected) are summarized below.

1. Engineering science course, on average, are truly producing students with acceptable abilities to problem solve. We rejected this explanation because there is simply too much evidence (e.g., see Woods et al.[23-27] and Wales[15]) to the contrary.
2. The survey design is flawed—that is, we are not measuring what we think we are measuring. While we have not fully established the validity of the survey, we have enough data to believe that this cause is unlikely.
3. Professors cannot really assess problem solving performance. We rejected this explanation because it is inconsistent with our experiences. For example, at oral exams of graduate students faculty perspectives on quality of problem solving are nearly always aligned. This suggests that faculty can judge quality and they tend to agree on quality when they see it.

Our interpretation of the survey results implicates present practices of assessment. The survey results suggest that our community (engineering educators) believes that present engineering science classes are producing students who are acceptable problem solvers. This belief is heartfelt and strong. We think that the explanation for present beliefs relates to present practices of assessment. Making valid assessments of problem solving performance is complex and requires a variety of methods, many of which are unlike traditional in-class exams and homework problems. Thus, we believe that the root cause of the problem is that faculty (on average) use poor assessment practices in the context of engineering science classes. These poor practices lead the faculty to reach inappropriate conclusions. To ameliorate this problem, the most pressing issue is to develop more effective practices for classroom assessment (i.e. by classroom assessment, we mean a continuous and ongoing assessment/learning process that is imbedded in the day-to-day practices of the professor—this process continually informs both students and professors of the results of the learning; see Stiggins[11]). Of course, the speculations inherent in this paragraph need to be tested by additional research (this is ongoing).

The results of the survey suggest that educators who are trying to improve educational practices need to carefully think about how to proceed. In particular, our research group is trying to create transformational change at our institution. The basic goal of this change is to dramatically improve learning by aligning teaching practices with modern knowledge of teaching and learning (e.g. Bransford et al.[3], Pelligrino et al.[8]). However, the survey data suggest that the community at our institution does not believe there is a need for change (i.e. present results in engineering science course are basically satisfactory). This suggests that the most important next step for our research group is to gather research-quality data that illuminates the present performance levels of our students. These data may help faculty members understand why they might want to consider changing their teaching practices.

Conclusions

To guide learning and assessment of problem solving, we have developed criteria for performance (see Appendix A). The performance criteria are a set of 30 specific objectives that can be observed and measured as engineers engage in relevant tasks. When an individual or team achieves high levels of performance on these objectives, then they have achieved quality in problem solving. Our work did not consider problem solving in general—instead it was constrained to problem solving that involves calculations based on mathematical representations of scientific concepts. Also, the primary focus was on engineering courses that emphasize calculations, analysis and modeling.

The performance criteria represent our operationalization of quality; which may or may not be aligned with beliefs of other professors and students. Thus, we designed a survey to assess community beliefs about our operationalization. This survey measured student and faculty beliefs about 8 of the 30 objectives. The results of the survey provide evidence that we have created objectives that align with both student and faculty perspectives. That is, both students and faculty believe that each objective that was tested should be emphasized in engineering science courses and that students can successfully learn this performance.

Overall, this study suggests a major problem that faces our research team. We are trying to create major change at our institution. This change involves increasing learning outcomes (dramatically) by aligning teaching practices with the best available scientific knowledge of how people learn. However, the results of the survey suggest that both students and faculty believe that our current engineering science courses are producing satisfactory levels of student performance in the area of problem solving. If the community (faculty and students) believe that the system is working, there is no reason to change teaching practices. We have inferred that the root cause of the problem is a lack of effective formative assessment practices. In simple terms, faculty members assess student problem solving and conclude that performance is satisfactory. However, the assessment practices are flawed and these flawed practices lead faculty members to draw invalid conclusions. Of course, the inferences in this paragraph are speculative. Future work will address these emerging hypotheses.

Acknowledgements

This work was funded in part by the National Science Foundation: Grant Numbers EEC-0212293 and DUE-0088591. This work was funded in part by the University of Idaho Department of Mechanical Engineering, College of Engineering and Office of the President.

Bibliographic Information

1. Arter, J., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Thousand Oaks, California. Corwin Press.
2. Bransford, J. D., & Stein, B. S. (1993). *IDEAL problem solver: A guide for improving thinking, learning, and creativity* (2nd ed.). New York, NY: W.H. Freeman and Company.
3. Bransford, J. D., Brown, A.L, & Cocking, R.R, eds. (2000) *How people learn: Brain, mind, experience and school.* Washington D.C. National Academy Press.
4. Elger, D., Armstrong, T., Beyerlein, S., Felicione, C., Fulcher, K., Rousseau, P. (2001, June) A structured problem solving model for developing high-level skills, Proceedings of 2001 Annual Meeting of American Society for Engineering Education.
5. Harb, H.R. (1997). *Anyone can be an expert skier.* Dumont, CO: Harb Ski Systems.
6. Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert vs. novice performance in solving physics problems. *Science*, 208, 1335-1342.
7. Matlin, M.W. (2002) *Cognition*. Fort Worth, TX: Hartford College Publishers.

8. Pellegrino, J.W., Chudowsky, N., Glaser, R., eds. (2001). *Knowing what students know: The science and design of educational assessment.* Washington D.C. National Academy Press.
9. Resnik, L.B., ed. (1987). *Education and learning to think.* Washington DC: National Research Council.
10. Schoenfeld, A.H. (1988) When good teaching leads to bad results: the disasters of "well-taught" mathematics courses. *Educational Psychologist,* 23(2), 145-166.
11. Stiggins, R. (2000). *Student involved classroom assessment.* NewYork, NY: Prentice Hall.
12. Tejada-Flores, L. (1994). *Breakthrough on skis: How to get out of the intermediate rut.* New York, NY: Vintage Books.
13. Trochim, W.M.K. (2001). *The Research Methods Knowledge Base.* Cincinnati, OH: Atomic Dog Publishing.
14. Wales, C. E., Nardi, A. H., & Stager, R. A. (1986). *Professional decision-making.* Franklin KY: The Writers Block.
15. Wales, C. E. (1979). Does how you teach make a difference. *Engineering Education*, 69(5), 81-85.
16. Wales, C. E., Nardi, A. H., & Stager, R. A. (1987). *Thinking skills: Making a choice.* Franklin KY: The Writers Block.
17. Wales, C. E., Nardi, A. H., & Stager, R. A. (1988). Do your students THINK or do they MEMORIZE?. *Engineering Education*. 582-588.
18. Wales, C. E. (1989a). *Teaching your subject.* Morgantown, WV: West Virginia University, Center for Guided Design.
19. Wales, C. E. (1989b). *Thinking about thinking about thinking.* Morgantown, WV: West Virginia University, Center for Guided Design.
20. Wales, C. E., Nardi, A. H., & Stager, R. A. (1989c). *Teaching thinking skills: Research supports experience*. Morgantown, WV: West Virginia University, Center for Guided Design.
21. Wales, C. E., & Stager, R. A. (1990). *Thinking with equations: Problem solving in math and science.* Franklin KY: The Writers Block.
22. Wankat, P. C., & Oreovicz, F. S. (1993). *Teaching engineering: Problem solving and creativity.* New York, NY: McGraw-Hill.
23. Woods, D. R., Crowe, C.M., Hoffman, T.W., & Wright, J.D. (1979, December). Major challenges to teaching problem-solving skills. *Engineering Education*, 277-284.
24. Woods, D. R., & Crowe, C. M. (1984, February). Characteristics of engineering students: Their first two years. *Engineering Education*, 289-295.
25. Woods, D. R. (1987). How might I teach problem solving?. In J.E. Stice (Ed.), *Developing critical thinking and problem-solving abilities* (pp. 55-71). New Directions for Teaching and Learning, Jossey-Bass.
26. Woods, D. R. (1992, Winter). Ideas about curriculum. *Chemical Engineering Education*, 34-37.
27. Woods, D. R. (2000). An evidence based strategy for problem solving. *Engineering Education*, 89(4), 443-459.
28. Faculty survey is at http://www.webs1.uidaho.edu/enrich/elger_results/asee_2003/faculty_survey.htm
29. Student survey is at http://www.webs1.uidaho.edu/enrich/elger_results/asee_2003/student_survey.htm

Biographical Information

DONALD ELGER. Professor Elger specializes in fluid mechanics and design. He has been at the University of Idaho for 16 years.

JENNIFER BELLER. Professor Beller specializes in assessment and evaluation. Her passion is social science research. During the past 10 years, Professor Beller has completed extensive studies on ethics in competitive environments. Recently she joined the faculty of Washington State University, and has been in her present position for 1 year.

STEVEN BEYERLEIN. Professor Beyerlein specializes in combustion and design. He has been at the University of Idaho for 16 years.

BARBARA WILLIAMS. Professor Williams specializes in porous media flow. After many years in industry she joined the faculty at the University of Idaho and has just completed her 2nd year.

Appendix A: An Operational Definition of Quality in Problem Solving (Context: Engineering Analysis)

Quality is achieved when a person follows the standard practices of the engineering community (i.e. what experts do.) To measure quality, one can observe performance, and compare these observations with practices of our community that are described by the thirty-two objectives listed below.

Trait 1: Understanding and Communication. Does the engineer have a rich understanding of scientific concepts? Is the engineer able to communicate and document this understanding in a way that is effective for relevant stakeholders?

Schema--Deep Understanding. Engineers organize their knowledge using structures (schema) in which the various pieces of knowledge are linked to one another (think of a spider web). Schemas are hierarchical—at the top of the pyramid are broad disciplinary concepts, which link to sub concepts, which link to facts and details.

Language of the Discipline. When communicating to others, engineers apply the concepts and ideas of our discipline. These concepts and ideas are communicated in a way that is valid and accurate and in a way that reflects an underlying schema.

Effective for Audience. Engineers match the communication to the audience. For example, when communicating with a layperson, engineers translate the engineering language into common words and ideas. For communicating with a peer, engineers use technical terms that are understood by the peer.

Verbal, Visual and Math Representations. To enhance the efficiency and effectiveness of communication, engineers blend three types of communication (visual, verbal and mathematical)

Effective Process of Documentation. Engineers document their work on paper as they proceed (they avoid backfilling). Documentation is organized and a peer can quickly skim and understand the work. Documentation is complete, meaning it is appropriately annotated and labeled, and presents essential details. Documentation is archival, meaning it can be located and used long after the project or task is completed. Documentation is cost effective, meaning the engineer does not spend unnecessary time creating the documentation.

Trait 2: Scientific Concepts and Ongoing Learning Does the engineer see the world through the lens of scientific concepts? Does the engineer combine problem solving with learning?

Scientific Concepts. Engineers interpret the world using scientific concepts such as Ohm's law, equilibrium, and the ideal gas law. When faced with an unfamiliar problem or situation, engineers use scientific concepts to guide their actions—that is they use concepts to create understanding, to make predictions, to make decisions, to solve problems and to perform other similar actions.

Scientific Honesty. Engineers "bend over backwards" to convince themselves that their work accurately reflects phenomena in the real world. That is, engineers seek to discover what is true by carefully checking their results against real-world observations and well-establish scientific laws. Engineers report their results honestly, and they reveal relevant concerns and uncertainties. Engineers do not manipulate their results in order to prove that a certain point-of-view is valid.

Assumptions. Engineers continually assess the validity of the underlying assumptions. Engineers communicate and document major assumptions.

Self-Assessment Engineers assess their knowledge by continually questioning and probing their own understandings and conceptualizations.

Ongoing Learning. Engineers learn in an ongoing and continuous process that is embedded in problem solving. Engineers assume responsibility for their learning and understanding. When understanding is in question, the engineer takes appropriate action such as asking a question or self-study.

Strong Learning Process. Engineers use effective means to learn. Engineers relate newly-learned knowledge to what they already know and to the physical world. Engineers organize their knowledge into schema. Engineers continually test and modify their schema.

Trait 3 Engineering Process. Does the engineer have a way to proceed on each new problem they face? Does the process used by the engineer work robustly and effectively?

Recognizing Opportunities. Engineers recognize opportunities to apply analysis based on math and science to enhance quality or reduce cost. Engineers avoid applying analysis when it is inappropriate, unnecessary or too expensive.

Cost Effective. Engineers match the precision and accuracy of an analysis to the needs of the task and the quality of the data available. Simple models (e.g. lumped parameter, algebraic equations) are used for rough estimates. Complex models (e.g. ordinary and partial differential equations) are used when accuracy is attainable and when the needs justify the cost.

Metacognition. Engineers have self-knowledge, awareness and control as they problem solve. Engineers can describe in words and pictures how they think and monitor their thinking.

External Representations. Engineers sketch diagrams and put information on paper as a means to reduce the demands on short-term memory.

Framework (Schema) for Procedural Knowledge. Engineers have a basic process (schema) that they apply to each new problem. Characteristics of this process are

a) Personal. Each engineer has adapted the process to fit their unique style and they have ownership.
b) Robust. The same basic process works for many different types of problems.
c) Flexible. The process is continually adapted to fit each new problem. The process is continuously improved.
d) Active. Engineers are active: talking, questioning, sketching, etc. Engineers avoid "sitting and thinking."
e) Effective. The process usually guides the engineer to a solution. The process is time efficient.

Stages of this process are:

a.) Problem Formulation. This is the process of interpreting the problem, creating a representation, and creating a specific goal.
b.) Strategy and Planning. This is the process of considering alternative ways to solving the problem, identifying a potential solution path and creating a simple plan.
c.) Action. This is carrying out the calculations and other actions associated with a solution.
d.) Review. This is reviewing the process of solution and the solution itself.

Trait 4 Problem Formulation. Does the engineer create a personal interpretation (representation) of each new problem? Is the representation founded on scientific concepts? Does the representation capture the relevant physics while neglecting non-essential details?

How Things Work. Engineers create representations that are founded on knowledge of how things work in the physical world. When this knowledge is lacking, engineers locate and learn the information.

Representation using Scientific Concepts. Engineers create representations that are founded on scientific concepts. Engineers simplify real-world problems by retaining factors that strongly influence the physics and eliminating extraneous details that have small effects on accuracy.

Representation of a Text Problem. Engineers create their own interpretation of a text problem. This interpretation is complete (considers all the given information) and valid. Engineers see text problems as they might exist in the physical world.

Concrete Goal. Engineers create a well-defined goal. This goal satisfies criteria such as specific, strategic (i.e. useful to the overall project), unambiguous, meaningful, and accurate.

Finding Information. Engineers locate relevant technical information such as material property data, standards data, and empirical parameters. Engineers assess this information, and make appropriate and valid decisions considering factors such as accuracy and applicability.

Trait 5: Units, Calculations and Concept of Estimation. Does the engineer realize that math models are approximate and that all calculations involve estimation? Is the engineer skilled with calculations and unit practice?

Units. Units are carried and canceled on each calculation, and this process is well organized and documented. When practical, numbers and symbols are assigned units. The engineer knows the meaning and relative size of each unit.

Calculations. Calculations follow the rules of effective documentation (organized, complete, simple, archival).

Concept of Estimation. Engineers make estimates that balance accuracy requirements with cost. Engineers seek an accuracy that is acceptable for the task—that is, they do not make a complex analysis when a simple analysis will suffice. Engineers can make valid estimates of the uncertainty of a calculation. Engineers can make quick estimates of complex problems.

Trait 6 Accuracy and Reviewing. Does the engineer have processes to produce valid and error-free work? Does the engineer continually grow their process for problem solving?

Accuracy. Completed work is free of errors. Throughout the solution process, decisions, communication and other factors are based on valid interpretations of engineering knowledge.

Error Checking. When proceeding through a problem, engineers routinely double-check most steps of their work.

Reviewing a Solution. After completing a calculation, engineers use a simple method such as a quick estimate to evaluate the solution. Engineers review their results and make inferences about the problem as it exists or might exist in the physical world.

Troubleshooting. Since getting stuck is a natural part of problem solving, engineers develop methods for troubleshooting that work in most cases. Once stuck, engineers systematically apply their methods, proceeding in a calm and deliberate fashion.

Kaizen (continuous improvement). To improve performance, engineers purposefully review. This review focuses on identifying strengths (practices that are effective for problem solving) and deltas (changes will bring about improvements in the future). To take advantage of the process of social learning, engineers involve others in these reviews.

Balance. Engineers balance persistence and perspective; they work hard, but they will take appropriate action if the effort to solve the problem becomes too much. By appropriate action, we mean performances such as asking for help, letting it go, or redefining the problem so that it is simpler.