Session 3425

Assessing Student Learning about Engineering Design in Project-Based Courses

Reid Bailey, Zsuzsanna Szabo, Darrell Sabers
Aerospace and Mechanical Engineering/Educational Psychology/Educational Psychology
University of Arizona

ABSTRACT

Teaching first year engineering students in a meaningful way is an issue struggled with at schools around the world. At the University of Arizona, our Introduction to Engineering course focuses on engineering design, communication, and teamwork primarily through three design projects over the course of one semester. While the course is fairly mature (it has been in existence, evolving to its current state, for roughly the past ten years), currently we do not know what our freshman students are learning with respect to the learning objectives of the course.

We have embarked on a study to assess exactly what our students are learning in Introduction to Engineering. Still in its early stages, this study is aimed at understanding what the students learn about engineering design through their experiences in the course and using that knowledge to improve the course.

In this paper, we focus on the strategy for assessing our students' engineering design knowledge. The backbone of this strategy involves a pre and posttest where students critique a proposed process for designing a product (e.g., a shopping cart, a device for counting eggs). The development of the pre and posttests as well as the detailed analytic rubric used to assess student responses is addressed in this paper. Such an assessment strategy should have broad applicability to the growing number of project-based first year engineering courses.

MOTIVATION

A core learning objective of the first year ENGR 102 Introduction to Engineering course at the University of Arizona is for students to learn about engineering design as a process. We want students to learn how to identify needs, develop solutions to meet those needs, and implement those solutions. The course is geared towards this goal with three team-based design projects. At this point however, we have nothing more than anecdotal evidence that our students know more about engineering design after the course than they do before. Our strategy for assessing students' design knowledge more accurately is presented in this paper.

CONTEXT

Introduction to Engineering (ENGR 102) is a required course for all undergraduate engineering students at the University of Arizona. ENGR 102 has evolved over the years into a well-developed class with its own textbooks[4,5] and a structure centered around three team-based design projects each term. Other elements of the course that focus on helping students learn about engineering design include in-class activities, a video highlighting design at an innovative industrial firm, and lectures on design.

As taught in ENGR 102, engineering design is composed of three universal phases[i]: 1) Problem Formulation, 2) Problem Solving, and 3) Solution Implementation. In Problem Formulation, the needs of the project are identified and represented in terms of criteria and constraints. In Problem Solving, multiple concepts are generated and analyzed, and one to two are selected for implementation. The concept(s) remaining after Problem Solving are built and tested in Solution Implementation. These three main phases are shown in Figure 1 as being circular to emphasize iteration. Also in the figure, the three phases are subdivided into ten *possible* steps (the ten steps are presented to the students as *one* way to implement the three phases, not *the* way).

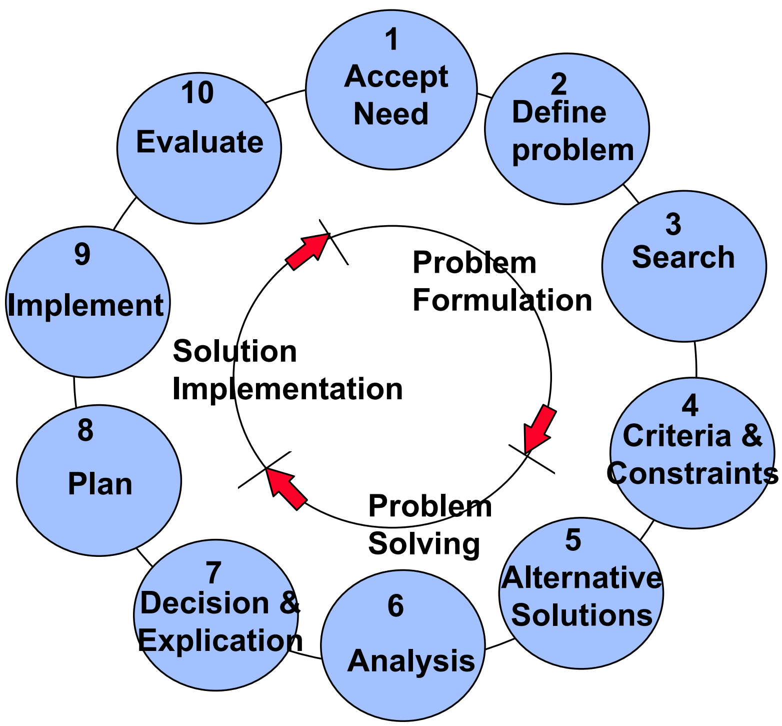

Figure 1 Design Phases as Taught in ENGR 102 [4]

The goal in this study is to assess how well students internalize the three main phases of engineering design shown in Figure 1. This assessment is aimed at providing better information so that the course can be improved.

---

[i] These three phases are common to nearly all representations of engineering design. One can find the same three phases (with variations in names and divisions, but not purpose) in, among others, Pahl and Beitz[7], Dym and Little[3], Ulrich and Eppinger[9], Dieter[2], and Pugh[8].

Previous Assessment of Engineering Design Knowledge at Arizona

In Introduction to Engineering, the previous assessment of engineering design knowledge relied on 1) team grades on design reports, 2) the performance of teams' built designs, and to a lesser degree, on 3) student's ability to answer design questions on tests. Design reports provided the most process-oriented grade, but at the team (not individual) level (individual knowledge of design could be very poor if a minority of team members did the majority of the work). Furthermore, design reports between different sections and different years of the course could not be compared due to inter-rater reliability problems. The performance of designs has fewer inter-rater reliability problems, but is strongly focused on the end result of design as opposed to the process used. Our interest is in understanding students' knowledge of design as a process. Test questions are problematic due to inter-rater reliability and because they are different among different sections of the class. A process-oriented measure of student's design knowledge that handled inter-rater reliability was lacking for ENGR 102.

Assessment of Student Learning

The use of essay responses in student assessment permits the measurement of students' ability to describe relationships, applications of principles, present relevant arguments, state necessary assumptions, describe limitations of data, explain methods and procedures, produce, organize, and express ideas, evaluate the worth of ideas, etc.[6]. For this reason, short essay responses were used in assessing engineering students' learning in project-based courses. The use of essay assessment is beneficial in assessing higher-order thinking skills. When using essay assessment it is very important to set well-defined criteria describing how the essays will be graded.

Nitko[6] mentions two general methods for scoring essays, the analytic method and the holistic method. A top-down method is used for crafting an analytic rubric[6]. An analytic scoring rubric requires first an outline containing a list of ideal points, major traits, and elements that a student should include in an ideal answer. The teacher would decide the number of points awarded for each element in the ideal response. When scoring students' responses, those who responded essentially correct would get the full credit of points, as compared to those who would respond totally incorrect and would receive no points for their answer.

The holistic rubric, on the other hand, assesses mostly an overall impression of the response in a less objective manner than the analytic rubric. In crafting a holistic rubric, a teacher would use a bottom-up method[6]. In the bottom-up method, the teacher begins using actual student responses of different qualities, sorting the responses in categories that would help identify the different levels of student responses. After students' responses are sorted, the teacher writes very specific reasons why each of the responses was put in the respective category. Then, for each category, the teacher writes a specific student-centered description of the expected response at that level. These descriptions constitute the scoring rubric to grade new responses.

The two methods (analytic and holistic) are not interchangeable, and the clear advantage of the analytic rubric, compared to the holistic rubric, is that it provides a more objective way of assessing students' strengths and weaknesses. Also, the analytic rubric can give teachers a clearer picture of the areas where students have more. The disadvantage of using essays and analytic rubrics for assessment, is that scoring might be a little slower than when objective items are used (e.g., true-false, multiple choice, matching).

In assessing students in the ENGR 102 course, the analytic rubric with a top-down crafting method was selected because this corresponds to the purposes of the course.

STRATEGY FOR ASSESSMENT

The goal of the assessment is to measure the change in our students' engineering design knowledge while enrolled in ENGR 102. This information will be instrumental in creating new approaches for teaching engineering design and for evaluating their effects. To accomplish this goal, the following strategy is employed.

- Students take a pretest on the first day of class
- Students take a posttest towards the end of the semester
- Students respond to a survey towards the end of the semester

The pretest-posttest experimental design is selected due to its ability to indicate change in a student's knowledge[1]. Because this research was conducted in a class in which engineering design is a core learning objective, we could not withhold the treatment from any of the students to create a control group for each semester. The lack of a control group each semester weakens the overall strength of the experimental design (e.g., all first year students may undergo a maturation process during their first term in college which results in changes in pretest and posttest scores). The treatment for this study is enrollment in ENGR 102, so it would be impossible to withhold treatment to students in the class. Additionally, there is not a realistic way to test a peer group of students not taking ENGR 102.[ii]

While only those students completing both pretest and posttest are included in the study, the results are analyzed on the entire population. That is, the pre and posttests are not paired to investigate effects on each individual.

Some of the effects of ENGR 102 on the students may not manifest themselves until after the semester is finished (this would be particularly true of a project-based class like ENGR 102 where several more design experiences are necessary for the seeds of knowledge planted during ENGR 102 to flourish). By only collecting data during the semester in which the course is taken, we are not measuring the long term effects of ENGR 102 on students' design knowledge.

Due to the requirement that each student sign a human subjects form, our sample will always have an element of self-selection. Only those students that choose to sign the human subjects form (approximately 47% of the students did) will be in the study. This weakness of the strategy for assessment is unavoidable.

The survey is an additional source of information from students. In the survey, students self report on the effect of ENGR 102 on their abilities, including the effect on their ability to design a system to meet a set of needs.

[ii] Interestingly, we are not as interested in seeing the exact effect of ENGR 102 on our students as we are in seeing how their knowledge changes while enrolled in ENGR 102 regardless of whether the change is solely due to ENGR 102 or if it is confounded with other factors. Other students in later years will have very similar confounding factors (e.g., moving to Tucson, moving into a dorm, exposure to a new set of friends, increased demand on time), so going to great lengths to separate effects is not useful.

Factors Considered in Building the Pre and Posttests

Testing a student's engineering design knowledge is a difficult task for several reasons. First, there is not one right answer that is being pursued in design. Two people can develop two very different designs to address the same need. Furthermore, each of these designs can be roughly equivalent in terms of how well it meets the needs of the problem. Therefore, evaluating students' design knowledge by comparing their final design to the "right" answer to a problem is inappropriate.

Instead of focusing on *end results* of design, the focus in ENGR 102 is on the *process* of design. Engineering design processes are highly adaptable to address different kinds of problems – this is a second reason that measuring design knowledge is difficult. Many aspects of the process used to design a space shuttle will look different than the steps used to design a toothbrush. The pre and posttests, therefore, must focus on those elements of design that are common across a wide range of problems. For instance, while a space shuttle's design process may use simulation intensively and a toothbrush's design process may not, the design of both involves some sort of analysis. Analysis of a design is the common element of engineering design in this example. Other common elements of engineering design include the identification of needs, the generation of multiple concepts, and the building and testing of a final design.

Building on the adaptability of engineering design, a third difficulty is that we cannot measure our students' engineering design knowledge by how well their designs performed when built and demonstrated. As an example, consider the waterwheel project in ENGR 102. This project involves the design of a waterwheel to lift a weight. The students are evaluated based on the work done lifting the weight divided by the cost of their waterwheel (which is based on material, manufacturing, and water costs). The best team has the highest work to cost ratio. Why not merely evaluate a student's design knowledge based on this ratio? Part of the answer is that we are not interested in their ability to design a waterwheel; we are interested in their ability to design a wide variety of systems. The exact steps followed to design their waterwheel effectively will be different than the exact steps to design a space shuttle or a toothbrush. Therefore, to measure their design knowledge (as opposed to their *waterwheel* design knowledge), we need a different approach. Additionally, using performance on a project as a measure of design knowledge does not provide information about students' knowledge of the individual parts of a design process (e.g., problem identification, idea generation, analysis).

This leads to the fourth difficulty in measuring design knowledge in ENGR 102: students work on teams in ENGR 102 and will likely work on design teams in their careers. It is difficult to measure design knowledge on an individual level when design typically occurs in teams. We address this by asserting that each engineer on a design team needs an understanding of the overall process through which a system is being designed. Therefore, our pre and posttests focus on the processes used by teams that each individual on the team needs to understand. The reason we have each student complete the pre and posttests, as opposed to the design teams, is that if two students out of the five on a team know a good answer and the other three are not involved in the response, then we could think our students are learning about design when in fact more than half of our students are not!

Finally, a method used in previous research to measure design knowledge involves videoing design teams and scoring their activities based on a rubric. With over 500 students in the fall semester and nearly 300 in the spring, the time intensiveness of such a method would be prohibitive. Furthermore, students work on the projects over the course of several weeks and do so in multiple locations (dorm rooms, meeting rooms, the library, etc.). Obtaining video to observe all the teams would be nearly impossible.

To summarize the factors considered in developing the pre and posttests:

- Comparing their designs to a “right” answer is not appropriate.
- Focusing on process is appropriate but difficult since the design process is not the same for each system.
- Focusing on design performance provides information too specific to the type of project and does not provide information about the student knowledge of individual parts of a design process.
- Design is typically performed in teams, but we need to measure design knowledge of individuals to ensure that our results are representative of the entire class, not just one or two people on each team.
- Measuring design knowledge by observing video of the teams is not realistic for a class with roughly 800 students per year.

Pre and Posttest

Based on these factors, the pre and posttest is focused on the common elements of an engineering design process, is completed by all students, and requires only twenty minutes of class time over the entire semester. The students are asked to critique a proposed design process as they would if a job interviewer asked them to critique it. The process that the students are asked to critique for the pretest, which involves designing a shopping cart, is shown in Figure 2.

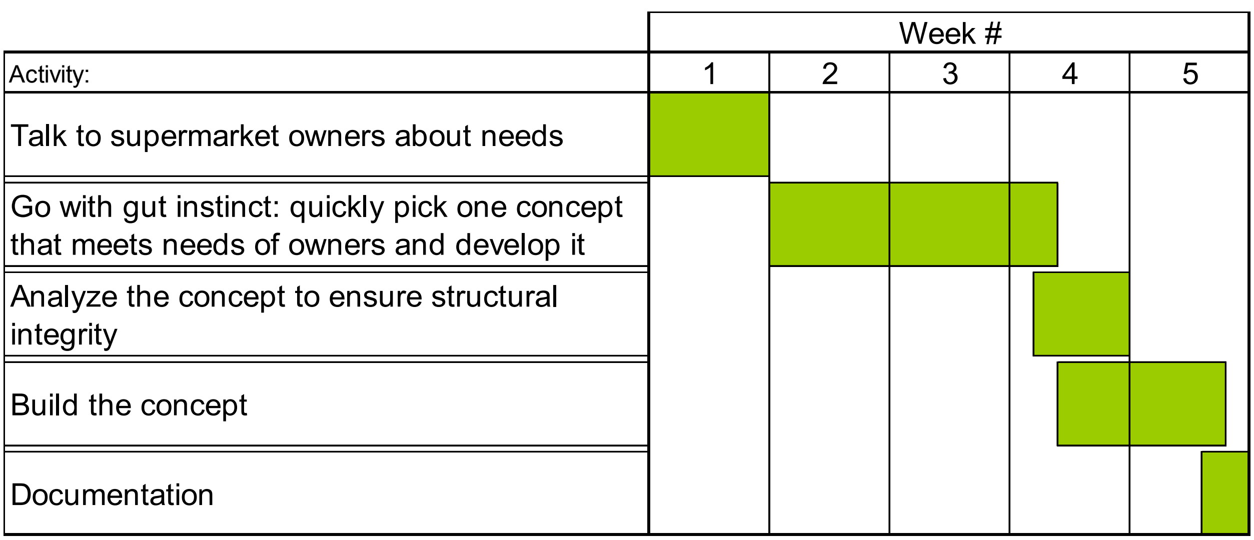

Figure 2 Pretest Proposed Design Process

The proposed process involves talking to store owners about what they need from a shopping cart. Then, a single concept is quickly selected and developed into a more detailed design. The structure of the cart is then analyzed before roughly a week is spent building the cart. Finally, two days are spent documenting the project.

Students are asked to comment on all the good parts of the proposed process and on all of the parts that need improvement or are missing. Students have ten minutes to complete their answer.

The posttest follows the same structure as the pretest, but involves the design of another system. Two posttests have been developed at this point: one is for the design of a device to count eggs on a conveyer belt and the other is for a canoe trainer (analogous to a rowing machine). The good and bad aspects of each test (the pretest and both posttests) are not identical (e.g., generating multiple ideas before moving forward with building a final design is a key element of engineering design - in the pretest, only a single design idea is generated, while multiple ideas are generated in the posttest).

ANALYTIC SCORING RUBRIC

An analytic scoring rubric is selected to score the pretest and posttest. We chose an analytic rubric because we want to measure students' knowledge of individual parts of an engineering design process. To do so, a holistic rubric (with one score for the entire response) would not be sufficient.

Development of the Rubric

The rubric was developed by first creating a general rubric for any of the questions. The general rubric was developed from the top-down with a strong foundation in engineering design methodologies such as Pahl and Beitz[7]. Next, the rubric was particularized to each pretest and posttest. This version of the rubric was applied to a pilot group of roughly 200 students in spring 2003. In fall 2003, six education students learned how to apply the rubric. Based on the pilot application and feedback from the education students, the rubric was changed and reworded significantly for two primary reasons:

1. Some items needed to be clarified so that non-engineers could understand the rubric. For example, the use of the word "analysis" was unclear and needed to be clarified to indicate that it refers to modeling of a design before building it (as opposed to "testing" a built final design).
2. Because the rubric did not fit all student responses well, it needed to be updated to better reflect a measure of a student's design knowledge. For instance, initially students received 2 points for indicating that more documentation was needed for the pretest. After scoring the pilot tests, the points were split such that indicating that documentation is needed throughout the design process receives 2 points whereas merely indicating that the time spent in documentation needs to be lengthened (a less specific answer) only receives 1 point. This bottom-up adjustment was needed to ensure that the theory-based rubric could be reliably applied to a wide range of student responses.

The end result of this process is the rubric presented in this paper in addition to a rubric guidebook (which includes a longer explanation of each part of the rubric along with examples of how to score specific student responses) used in teaching people how to apply the rubric.

The Rubric

The analytic rubric used to score the shopping cart pretest shown in Figure 2 is presented in this paper. Before doing so, a guide to the rubric is shown in Figure 3.

Design Phase	Step	Pts.	Description	Shopping Cart
II	4	3	**Analyze** ideas on **all** relevant criteria and constraints Possible means of analysis include (do not have to mention any of these, but these are key words to look for for analysis): • Experiments/Design of Experiments • Equations/Analytical Models • Simulation • Verbal analysis through group discussion of designs	Positive: Analyzed concept for structural integrity, or, indicates that more time is needed for analysis of structural integrity (+1 pt) Negative: other areas besides structural integrity (e.g., weight, steering, ergonomics) need to be analyzed +1 pt for noting that more analysis is needed +1 pt. for noting an additional type of analysis needed

Refers to Phase from Figure 1.

Unique Step I.D. to Identify this Level.

Total # of points for this level.

General description of this level. This column is the same for all pre and posttests.

Description and point allocation particularized for the shopping cart. "Positive" means that something is done well in the proposed process in Figure 2. "Negative" means that something is omitted or needs to be changed in the proposed design process

Figure 3 Rubric Layout

The rubric is divided into seventeen different levels, one of which is shown in Figure 3. The first four columns of the rubric are identical for all pre and posttests. The last column is particularized to the particular questions. In this case, Step 4, which is part of Design Phase II in Figure 1, is worth a total of three points. The point totals are allocated to roughly approximate the importance of that level with regards to what we are trying to teach our students about engineering design. The "Description" in the fourth column indicates that Step 4 is concerned with the analysis of a design on all criteria and constraints, where analysis involves activities such as using equations to model a system and simulating a system.

The column entitled "Shopping Cart" shows how to apply the rubric to the shopping cart question from Figure 2. In this case, the proposed process is good (as indicated by the word "positive") in that the shopping cart concept is analyzed for structural integrity. Recognizing this is worth one point (of the three total for this level). The proposed process is not good (as indicated by the word "Negative"), however, in that no other areas besides structure are analyzed for the shopping cart. Recognizing that other areas need to be analyzed is worth one additional point and giving at least one example of an additional analysis that would be needed is also worth one additional point. In Table 1, sample scores for possible student answers are shown.

Table 1 **Sample Responses and Scoring**

Points Earned	Answers
+1 pt total – analyzing structural integrity	◊ *"it is good that they checked that the cart wouldn't break before building it"* ◊ *"More time is needed to ensure the structural integrity" –* pts given because this answer clearly indicates that analyzing for structural integrity is good ◊ *"Doing analysis before building is good"*
+ 3 pts total – should analyze other areas, too.	◊ *"They should analyze the cart to see how heavy it will be, too."* ◊ *"In addition to structural integrity, they also need to make sure the cart is easy to use before they start building it"* (ergonomics)

Scoring an entire response involves evaluating student responses on seventeen different levels such as the "analysis" level. The levels of the rubric are grouped into eight that relate to the design process in Figure 1 and nine that deal with general issues relating to design such as time allotments for activities and documentation. The design process levels, labeled Steps 1 through 8 are shown in Table 2 while the general levels, labeled Levels A through I are shown in Table 3.

Table 2 Design Process Levels

Design Phase	Step	Pts.	Description	Shopping Cart
I	1	0.5	State that a **team** must be formed for the project.	Negative: +0.5 pts if stated that a team is needed
I	2	4 < 4 earned if this step is addressed but • multiple sources are not addressed • sources are not comprehensive	Gather information about project needs from **multiple sources**: sources should include: • **All** users (current and potential) of this type of device (e.g., shoppers, store owners, children) • Library and on-line **research** (e.g., information on injury statistics associated with shopping carts) • Existing designs – from **literature** (e.g., information from current manufacturers)and from **direct use** of existing designs (e.g., using a standard shopping cart) Information gathered is **used to form criteria and constraints** for the project.	Positive: information is gathered about needs (+1.5 pts) Negative: Only one source used to gather information (shop owners): +1.5 for noting that more sources are needed +0.5 for noting one additional source +0.5 for noting 2 or more additional sources Additional sources include: ➢ customers ➢ baggers ➢ research on injuries ➢ research on existing products ➢ children ➢ using the cart themselves
II	3	3	Generate **multiple ideas** to address the project needs through **brainstorming**	Negative: +3 pts if stated that they need to develop more than just one idea
II	4	3	**Analyze** ideas on **all** relevant criteria and constraints Possible means of analysis include (do not have to mention any of these, but these are key words to look for for analysis): • Experiments/Design of Experiments • Equations/Analytical Models • Simulation • Verbal analysis through group discussion of designs	Positive: Analyzed concept for structural integrity, or, indicates that more time is needed for analysis of structural integrity (+1 pt) Negative: other areas besides structural integrity (e.g., weight, steering, ergonomics) need to be analyzed +1 pt for noting that more analysis is needed +1 pt. for noting an additional type of analysis needed
II	5	2	**Based on the analysis**, **decide** which idea best meets the criteria without violating any constraints (may retain more than one concept if further iterations eventually reduce it to one final concept) Decision-making may include (do not have to mention any of these, but these are key words to look for for decision-making): • Voting • Selecting concept that maximizes a single objective • Reaching group consensus • Using a decision tool	Negative: They plan to go with "gut instinct" to choose which design to move forward with. +1 pt for stating that going with "gut instinct" is not good practice +0.5 pts for stating an alternative to going with gut instinct, such as voting, weighing strengths and weaknesses of multiple designs, considering multiple objectives +0.5 pts for stating that you should analyze your design before making decisions (i.e, before "going with gut instinct"
III	6	--	**Plan** how to build the selected concept	N/A for shopping cart
III	7	3	**Build** the concept	Positive: The concept was built: this must be directly addressed to get 3 pts.
III	8	3	**Test** the built concept to determine **how well criteria and constraints are met**	Negative: The built cart is never tested. +3 pts for stating this.

Table 3 General Levels

	Pts	Description	Shopping Cart
A	3	The 3 phases are each addressed in the appropriate order (will always be correct on sample, and should be mentioned).	Positive: This is done well here +3 pts total: clearly states that plan is "logical" or that each task flows from one to next +1.5 pts total: vaguely states that plan is "well organized"
B	Depends on how many problems	The 8 steps are each addressed in the appropriate order (will not be correct on some questions, this should only be mentioned for incorrect aspects).	N/A
C	2	Iteration should be planned into the process.	Negative: No iteration here. Must clearly state that time must be planned in for iterating back to earlier steps when problems are found. (+2 pts)
D	1.5	Relative time allotments should be reasonable: phase II with more time than phase I, phase III leaves enough time not only for building and delays but also for testing (roughly same amount of time as phase II, but depends on project)	Positive: "Getting needs from shop owners" time is reasonable (+0.5 pts) Negative: Too much time spent developing concept before building; "more time for building" (+1 pt)
E	1.5	Gantt chart must have sufficient detail to be useful.	N/A – detail of chart is fine, and comments to that effect should receive 0.5 pts under Step H
F	1.5	Criteria and constraints (i.e., the needs of the project) must be use in analysis, decision-making, and testing.	Positive: Needs are addressed in both concept development and in analysis (1.5 pts)
G	2	Project should be documented throughout (1 pt) with enough time left at end (1 pt) to compile and finish documentation	Negative: Not done well here. +1 pt only if stated that more time is needed for documentation +2 pts total if stated that documentation should occur throughout the process
H	+0.5	Extra credit for insights not listed on rubric	Examples include: "too many things happening in week 4" or "good to be doing more than one thing at a time"
I	-1	Answers that are directly incorrect. (e.g., saying that analysis is not necessary when it is)	

The following are notable aspects of the overall rubric:

- The rubric is split into levels so that a student's knowledge about specific aspects of design is measurable.
- The shopping cart case does some things well (e.g., Step 7, building the shopping cart), some things not well (e.g., Step 8, testing the shopping cart), and some things have both good and bad aspects (e.g., Step 2, identifying the needs). While this is true for the two posttests, the same things are not done well/poorly on each of the tests.
- Steps 1 through 8 are elements of engineering design processes that are general enough to be applicable to the design of a wide range of systems. That said, they are described in significant detail for the shopping cart case. A more detailed "rubric guidebook" is used by raters when the descriptions in Tables 2 and 3 are not sufficient.
- In addition to awarding points for knowing which steps should be performed, Level A in Table 3 awards points for students showing that they understand how all the steps fit together.

- ❖ Levels H and I in Table 3 are unique. Level H gives an extra half point for any response that shows design process knowledge but that is not elsewhere in the rubric. Level I removes an entire point if a student's response directly contradicts the rubric (e.g., a student receives 0 points if they are tacit on idea generation, but -1 point if they say that it is good that only one idea is generated for the shopping cart concept).
- ❖ Certain levels on the rubric (e.g., Step 6 and Level E) are not applicable to the shopping cart case, but are included in the generalized rubric because they are applicable to one or both of the posttests.

A sample student response is shown in Figure 4.

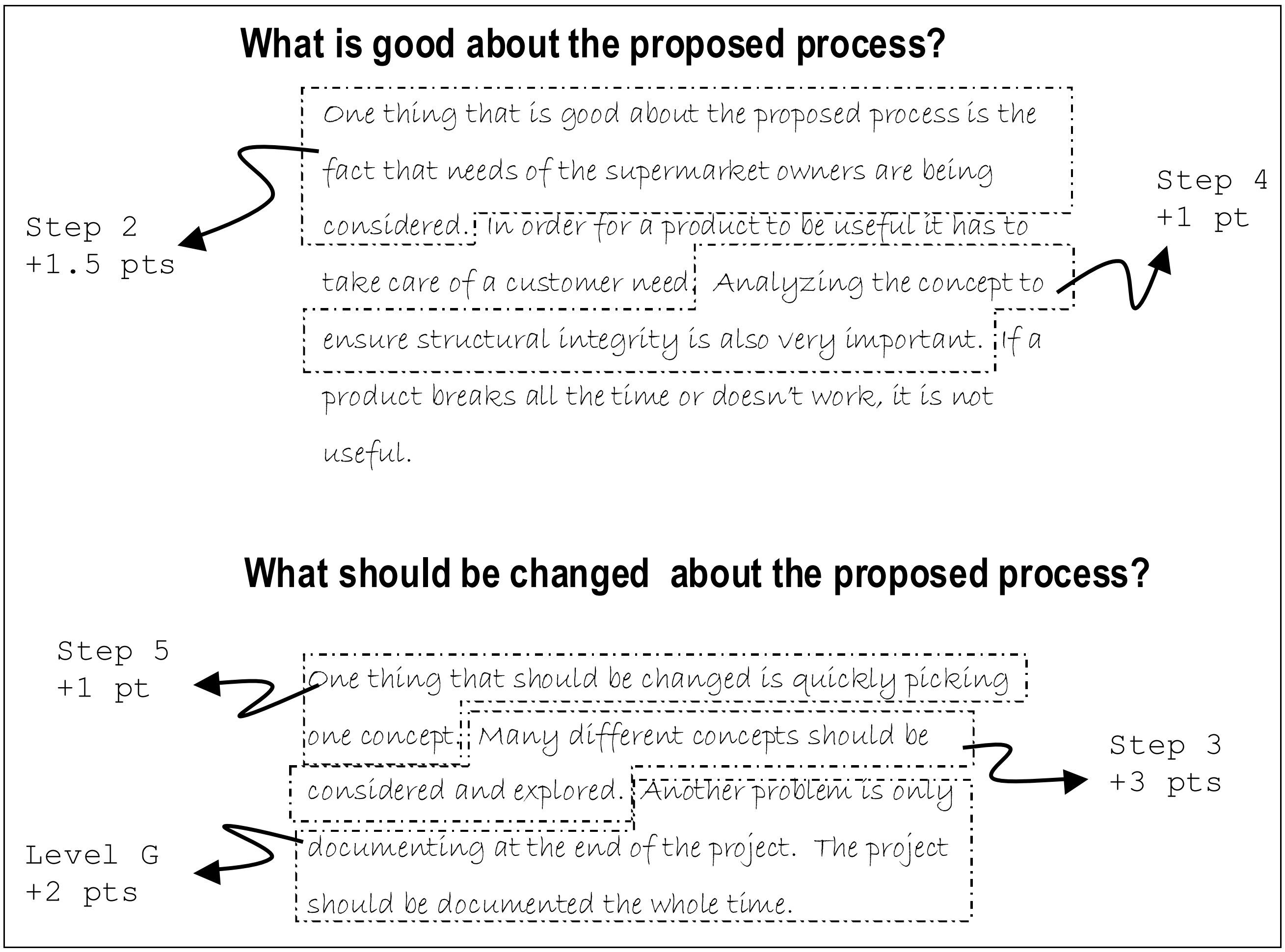

Figure 4 Sample Response and Scoring

The first statement receives +1.5 points on Step 2 because the student notes that talking to the supermarket owners about needs is important. Another point is gained by noting that analyzing the cart's structure is "also very important." The second and fourth sentences in the top half of the response, while true, merely support the other two sentences and therefore are worth zero additional points on the rubric.

The student receives six points for describing weaknesses of the proposed process. First, she asserts that quickly picking one concept is not a good way to design a system. Instead, she indicates that multiple ideas should be generated, thereby earning 3 points for Step 3. Finally, by noting that documentation should receive time allocations throughout the project instead of only at the end, the student earns two additional points. The total score for this person is 8.5.

LESSONS LEARNED

From the process of creating the pre and posttests, creating the rubric, and applying both, the following lessons have been learned.

- An analytic rubric, by allowing different aspects of a response to be scored individually, is appropriate for evaluating design knowledge. We can learn more about our students' strengths and weaknesses with it than with a holistic measure where everything is aggregated into one score. The teacher can then modify the curriculum and emphasize the topics that students have difficulties in learning.
- There are no shortcuts to making a good test and rubric for measuring engineering design knowledge. The construction of the rubric is iterative. A top-down initial approach followed by refinement of the rubric through having several people apply it to student responses was effective. Although the initial cycles of this were the most crucial, more iteration is needed for each new question to which the general rubric is applied.
- It is possible to teach people not familiar with engineering design to use the rubric effectively. Eight one-hour meetings were devoted to training the education students to apply the rubric. After being trained, the education students scored a response in an average of no more than 1.5 minutes.
- The eight hours of training was not enough to achieve desired levels of inter- and intra- rater agreement. Additional training and practice will be provided in the next phase of the study.
- The strengths and weaknesses of the proposed design processes in the pre and posttests were not the same. For instance, there is not any time allotted in the shopping cart process for iteration, whereas for the posttests iteration is part of the proposed process. Analysis of the results is not complete, but it appears as though this choice will bias the results. For example, because iteration is included in the proposed design process for the posttest, students are more likely to comment on it than in the pretest where it is omitted.

CLOSURE AND FURTHER WORK

We started with a question that few if any universities know the answer to for their students: What are our students learning about engineering design in a project-based course? The approaches applied to answer this question are not new to education, but are certainly not widespread in engineering. A pre and posttest is given to our students and scored with an analytic rubric that measures their knowledge about certain aspects of design.

While the iterative construction of the tests and the rubric was time consuming, the same tests and rubric can be reused for years. In fact, the applicability of the approach does not end with the freshman ENGR 102 course. Nothing about the rubric prevents its application to design classes throughout the curriculum or to design classes at other schools. This is due to the broad foundation of the rubric in common elements of engineering design.

Acknowledgements

The authors are grateful for support from the National Science Foundation through grant EEC-0338634.

References

1. Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
2. Dieter, G. E. (2000). *Engineering Design: A Materials and Processing Approach*. New York: McGraw Hill.
3. Dym, C. L., & Little, P. (2000). *Engineering Design: A Project-based Introduction*. New York: John Wiley & Sons, Inc.
4. Johnson, V. R. (2003). *Becoming a Technical Professional* (2nd ed.). Dubuque, Iowa: Kendall/Hunt.
5. Johnson, V. R. (2003). *Becoming an Engineer and Teaming on Design Projects* (4th ed.). Dubuque, Iowa: Kendall/Hunt.
6. Nitko, A. J. (2004). *Educational Assessment of Students* (4th ed.). Upper Saddle River, New Jersey: Prentice Hall.
7. Pahl, G., & Beitz, W. (1996). *Engineering Design* (K. Wallace, Trans.). New York: Springer.
8. Pugh, S. (1990). *Total Design*. New York: Addison-Wesley.
9. Ulrich, K. T., & Eppinger, S. D. (2000). *Product Design and Development*. New York: Irwin McGraw-Hill.

Biographies

REID BAILEY is an Adjunct Assistant Professor in the Department of Aerospace and Mechanical Engineering at the University of Arizona. His research interests include engineering design, environmental issues affecting design, and engineering education. He received his B.S. from Duke University and both his M.S. and Ph.D. from the Georgia Institute of Technology.

ZSUZSANNA SZABO is a doctoral student in the Educational Psychology Department at the University of Arizona. Her research interests include assessment, team learning, and gender issues in education. She received her B. S. in Civil Engineering from Technical University Cluj, B. S. in Psychology from University Babes-Bolyai, both in Romania, and her M. Ed. from SUNY at Buffalo, NY.

DARRELL SABERS is Professor and Head of the Department of Educational Psychology at the University of Arizona. His research specialty is applied psychometrics, especially focused on educational testing and research. He received his Ph.D. in Educational Measurement and Statistics at the University of Iowa.