Int. J. Engng Ed. Vol. 22, No. 3, pp. 508–518, 2006
Printed in Great Britain.

0949-149X/91 $3.00+0.00

Assessing Engineering Design Process Knowledge*

REID BAILEY
College of Engineering, University of Arizona, 1130 N. Mountain Avenue, Tucson, AZ 85721, USA
rrbailey@u.arizona.edu

ZSUZSANNA SZABO
Educational Leadership Department, Southern Illinois University, Edwardsville, Illinois, 62026, USA

Rigorously assessing students' design process knowledge is essential for understanding how to best create learning environments to facilitate the development of such knowledge. Such assessment is also quite difficult and hence there is a lack of assessment tools capable of measuring design process knowledge of every student in a large college. Faculty from both the Colleges of Engineering and Education at the University of Arizona are developing such a tool. The approach being developed aims at assessing if students can explain and analyze an engineering design process by having them critique a proposed process. Two versions have been developed so as to provide a pre- and post-test. An analytic scoring rubric is used to assess the design knowledge embodied by the student responses. Results from the 2003–4 academic year indicate that one of the two tests has sufficient validity and that the scoring rubric is too detailed for the nature of the tests. Hence, in the second phase of this work, a new test will replace the invalid one and a simpler rubric will be implemented.

Keywords: design process; assessment; Bloom's taxonomy.

MOTIVATION

A CORE LEARNING objective for engineering students from all disciplines at all universities is to learn about engineering design. To this end, capstone design courses populate nearly all curricula, while design courses in freshman and other years are becoming more commonplace. Despite the ubiquity of engineering design in curricula, little is known about what students learn in engineering design courses. Rigorously assessing students' design process knowledge is essential for understanding how to best create learning environments to facilitate the development of such knowledge.

Such assessment is also quite difficult and hence there is a lack of assessment tools capable of measuring the design process knowledge of every student in a large college. We are currently developing a tool to address this need. Starting in Spring 2003 as an unfunded pilot study and continuing from Fall 2004–Spring 2005 with funding from the National Science Foundation, this work has led to the development of a pre- and post-test along with scoring rubrics. The tests and rubrics have evolved as data from more than 400 engineering students has been analyzed.

* Accepted: 5 December 2005

CONTEXT

A process of engineering design is subjective in that there are no mathematical proofs or conclusive experiments to prove that one process is *the* process. That said, some common elements of engineering design have emerged over the course of centuries of engineering. These common elements are seen today throughout the disciplines of engineering and in practice throughout industry (albeit in varying forms). Engineers (1) clarify and articulate a need, (2) create a design to meet that need, and (3) implement that design. These three phases of design are typically iterated several times before a design is finalized. This process is shown in Fig. 1.

The process in Fig. 1 starts at the center of the spiral. 'Problem Formulation' steps relate to identifying and clarifying the needs to be met by the engineers. 'Problem Solving' steps involve developing designs on paper. Problem Solving includes the divergent process of creating several solutions to a problem and the convergent steps of analyzing this set of solutions and selecting the more promising solutions to implement in the third phase. 'Solution Implementation' is focused on turning ideas on paper into realized systems. The two primary activities within Solution Implementation are building the design and testing it. [Here the word 'building' is used broadly to include not only physically building a system, but also activities such as writing software code.] The spiral is used to represent iteration through these three phases.

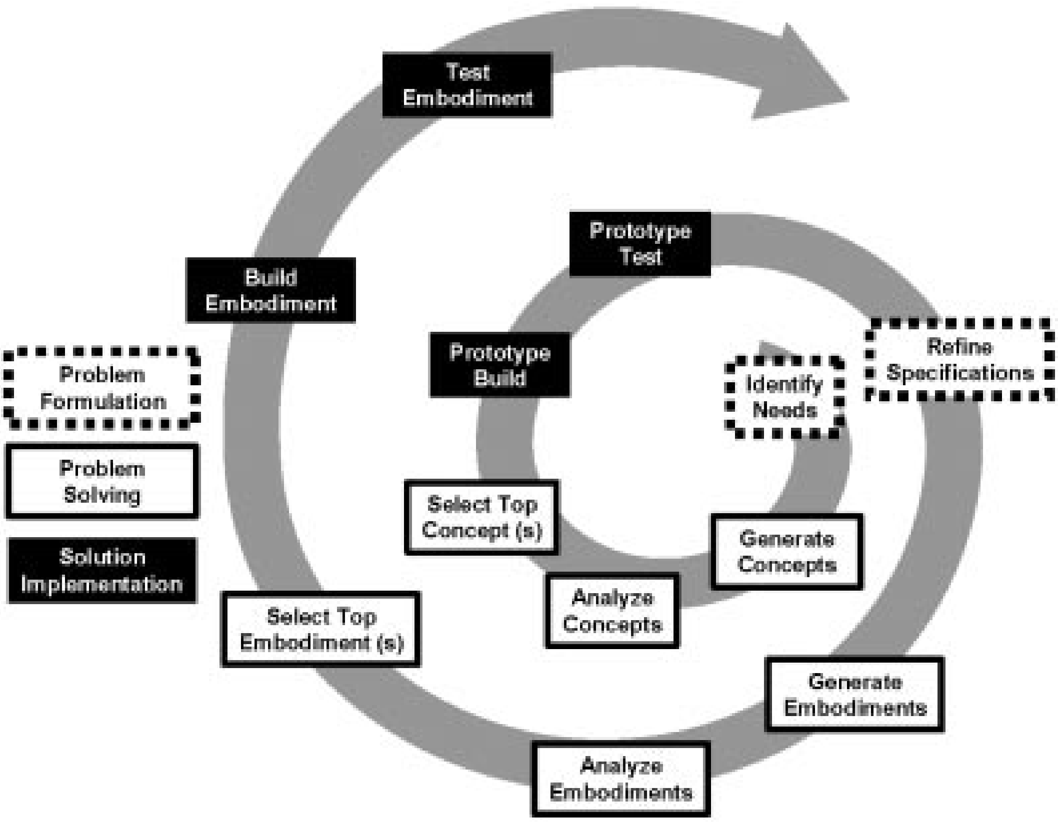

Fig. 1. Engineering design spiral.

While each iteration may not include every step of each phase, each of the phases is found in nearly any engineering design process at some point.

The work presented in this paper is aimed at assessing the following overall instructional objective: students should be able to explain and analyze a design process involving iteration through the three phases shown in Fig. 1. The target group includes both first year engineering students enrolled in an introduction to engineering design course and seniors in capstone design courses. This instructional objective is linked to multiple levels from Bloom's taxonomy, a set of six basic types of cognitive learning [1]. A revised version of Bloom's taxonomy contains the following six levels [2, 3]:

- *Remembering*: Being able to recite information from memory without necessarily understanding it.
- *Understanding:* Being able to explain material within its own domain.
- *Applying*: Being able to use a concept to solve a particular problem.
- *Analyzing:* Being able to parse something into its parts.
- *Evaluating*: Being able to judge different concepts and determine their value.
- *Creating:* Combining concepts to create something new.

These six levels of the revised Bloom's taxonomy are related to the engineering design model from Fig. 1 in Table 1.

It is expected that students in the introduction to engineering design class are able to remember, understand, and apply the engineering design process in Fig. 1. For seniors, students should also be able to analyze the steps being used and begin to evaluate the effectiveness of alternative

Table 1. Engineering design learning related to the revised Bloom's taxonomy

Remembering	Redrawing the spiral figure in Fig. 1.
Understanding	Explaining Fig. 1 and each phase represented in it.
Applying	Implementing the process depicted in Fig. 1.
Analyzing	A higher level of understanding and application where the purpose of each step is clearly understood and only used when necessary.
Evaluating	Comparing the process in Fig. 1 with other design processes and explaining the strengths and weaknesses of each process.
Creating	Forming an entirely new design process.

design processes (i.e., different specific manifestations of the process in Fig. 1).

The assessment approach being developed aims at assessing if students can explain (Understand) and Analyze an engineering design process by having them critique (Evaluate) a proposed process in the context of what they learned in class. Students are asked to compare the process in Fig. 1 to a very poor proposed process. An analytic scoring rubric is used to assess the design knowledge embodied by their responses.

ASSESSMENT OF ENGINEERING DESIGN PROCESS KNOWLEDGE

In creating an assessment strategy for engineering design process knowledge, the following are key criteria. The strategy must be:

- At the individual, not team, level
- Process-focused (not only focused on quality of end result)
- Not too time-intensive (not requiring significant class time or unreasonable amounts of time to prepare and score)
- Reliable from student to student, project to project, and year to year
- Linked to more than just one level of Bloom's taxonomy (this is important because engineering design process knowledge can span many levels of Bloom's taxonomy).

A trade study of several basic assessment strategy options based on these criteria is shown in Table 2.

It is clear that each approach in Table 2 has strengths and weaknesses. Not being at the individual level is a big weakness of both *design reports* and *final designs*, with final designs also being hampered because they are not process-focused. These two approaches are attractive to many,

Table 2. Options for Assessing Engineering Design Process Knowledge

Assessment option	Positives	Negatives	Scoring method
Base assessment on *Design Reports* that students already turn in as part of class	Process-focused Already a part of class, so not too time intensive	Not at individual level Potential reliability problems with raters (both intra and inter†) Only linked to application level of Bloom's taxonomy	Checklist rating scale
Base assessment on the *Performance of Final Designs*	Already a part of class, so not too time intensive Customer is ultimately most interested in final design performance	Not process focused Not at individual level Potential reliability problems with raters (both intra and inter) Only linked to application level of Bloom's taxonomy. Too specific to a single project	Rubric-based assessment of final designs
Base assessment on responses to *Close-Ended Questions (e.g., multiple choice)*	At individual level Can be process-focused No rater reliability issues Easy to score/grade Particularly good at lower levels‡ of Bloom's taxonomy	Difficult to link to higher levels§ of Bloom's Taxonomy Can close-ended questions really assess design skills?	Answers are either right or wrong
Base assessment on responses to *Open-Ended Questions (e.g., short answer, essay, concept map)*	At individual level Can be process-focused Can be linked to multiple levels of Bloom's taxonomy	Potential reliability problems with raters (both intra and inter) Can be time intensive to score/grade Difficult to formulate questions to assess design skills effectively	Analytic or holistic rubric-based assessment of responses
Base assessment on *Video* of design teams 'in action' or reflecting on process	Process-focused Can be at individual and/or team levels Can be linked to multiple levels of Bloom's taxonomy	Extremely time intensive to transcribe and score videos. Potential reliability problems with raters (both intra and inter)	Performance assessment.
Base assessment on *Portfolios* of student work	At individual level Process-focused Can be linked to multiple levels of Bloom's taxonomy Based on coursework, not additional assignments Can assess learning over time	Potential reliability problems with raters (both intra and inter) Very time intensive to score/grade	Rubric-based assessment of portfolios.

† Intrarater reliability concerns the consistency with which a single rater scores projects (if one person scores the same project twice at different times and the scores are significantly different, then there are intrarater reliability problems). Interrater reliability concerns the consistency with which multiple raters score projects (if two raters score the same project and their ratings are significantly different, then there are interrater reliability problems).
‡ Refers to Remembering, Understanding, and Applying levels.
§ Refers to Analyzing, Evaluating, and Creating levels.

however, because they are already integrated into most design courses

Asking students close-ended and open-ended questions are also possible ways of assessing design process knowledge. *Close-ended questions* (i.e., questions with right and wrong answers) have a long list of positives. They are easy to score, avoid reliability problems with raters, and can be at the individual level. Carefully worded close-ended questions can even tap into process-related knowledge of a student. These questions, however, are typically limited to the lower three levels of Bloom's taxonomy. That is, students can score well on such questions by memorizing parts of a design process and why those parts are necessary. Scenarios can be posed where students must show some application-level knowledge from Bloom's, but it is very hard to assess knowledge from the analysis, evaluating, or creating levels.

A team of researchers at the University of Massachusetts at Dartmouth has developed the Design Process Knowledge Test (DPKT), which is a set of close-ended questions used to assess the declarative knowledge of design students [4]. The focus of the DPKT is assessing on the two lower levels of Bloom's Taxonomy—remembering and understanding—via multiple-choice questions. The researchers at UMASS Dartmouth are developing other tools to link more strongly to process (or, procedural) skills and higher levels of Bloom's.

One such additional tool being developed at UMASS Dartmouth involves the use of *open-ended questions*. In their case, students are asked to reflect on a design task via a survey or essay [4]. Others have explored the use of concept maps as a means to measure a student's design skills [5]. With concept mapping, students are asked to draw a set of ideas and their interrelationships. At Georgia Tech, for instance, students have been asked on the first day of a design class to map 'all the words they associate with the concept design' [5, p. 678]. In each case, students are given the freedom to construct responses that show knowledge from multiple levels of Bloom's. With both of these examples, however, the scoring of student responses is a challenge that must be addressed. This challenge, however, is not insurmountable and can be addressed with a well-designed and validated tool.

Using *video* requires a huge amount of time to watch and reliably score the tapes. The researchers at UMASS Dartmouth have an innovative approach to circumventing this drawback. They plan to construct a computer simulation that puts students in design scenarios [4]. Then, the simulation will ask questions at key points to gain insight into how the student would proceed. If successful, this approach, a hybrid between videoing students in actual design scenarios and asking them close-ended questions, could capitalize on the advantages of these two types of assessment.

Finally, the use of *portfolios* for assessment has increased greatly in the last several years. Their ability to provide insight into a student's knowledge throughout a given time period is one of their greatest strengths. Their drawbacks are similar to the drawbacks for open-ended questions: problems with rater reliability and the time intensiveness of scoring. Scoring rubrics are the most common answer for improving rater reliability. A scoring rubric with six criteria for design process portfolios showed promise based on preliminary results at Georgia Tech [5]. A rubric with seven criteria for design developed at Rose-Hulman is under development [6]. One issue being addressed at Rose-Hulman is that if the students select to include poor examples of their work in their portfolio, then the results will show lower knowledge than actually exists.

In the work presented here, a new method of using open-ended questions is presented along with results from the first round of its implementation. Its purpose in a design course is to address the areas of assessing design process knowledge not addressed by final reports and final demonstrations. That is, it is process-focused and at the individual level. The primary research performed has been in creating and evaluating a rubric whose application is not too time-intensive; such a rubric will provide more reliable assessment from across different projects and different semesters in less time than videoing and scoring each design team would. Open-ended instead of close-ended questions are used in the strategy presented here in an effort to capture the complexities of a design process and to reach toward higher levels of Bloom's taxonomy.

The fourth column in Table 2 indicates the scoring method associated with each assessment option. For the open-ended questions approach used here, responses are assessed with a rubric. In the following section, a review of rubrics is presented.

Assessment of student learning

Authentic assessment (performance assessment) is used when the goal of assessment is to observe the process of thinking or the actual behavior of a performance task. The authenticity of a test is measured by how closely the test assesses the knowledge and skills required in real life [7]. The essence of authentic assessment is performance. Performance assessment requires students to use their knowledge and produce something (e.g., a group project, a demonstration, to build something, produce a report). As defined by Nitko [8], performance assessment requires students to apply their knowledge and skills and presents them with a hands-on task that requires them to do an activity.

In order to evaluate how well the students have achieved the task, clearly defined criteria are used. When performance assessment is used, students are required to demonstrate their achievement by producing a developed written or spoken answer that will demonstrate their achievement of a learning target. The performance task can be used to assess the process, the product, or both. To assess the performance task, scoring rubrics or rating scales are used.

The use of essay responses in student assessment [8], on the other hand, permits the measurement of students' ability to describe relationships, apply 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. The use of essay assessment is beneficial in assessing higher-order thinking skills. For this reason, short essay responses were used in assessing engineering students' learning in project-based courses. When using essay assessment, it is very important to set well-defined criteria describing how the essays will be graded. Nitko [8] 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. 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. Students who respond correctly to that element get the full credit, as compared to those who responded incorrectly and receive no points for that element.

The holistic rubric [8] assesses 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. In the bottom-up method, the teacher begins using actual student responses of different qualities and sorting the responses in categories that would help identify the different levels of students' 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 is that it provides a more objective way of assessing students' strengths and weaknesses. Also the analytic rubric can give teachers a clearer look over the elements where students have difficulties in answering, and might need to be retaught. The disadvantage of using analytic rubrics for assessment compared with holistic rubric is that student performances are compared with ready-set standards developed in accordance to teachers' expectations about what students are supposed to know. The ideal answer might not always reflect what students really would be able to answer based on what was taught. Hence, analytic rubrics usually undergo many revisions as the top-down expectations are adjusted to better match actual student responses.

The scoring in both holistic and analytic rubrics is slower than when objective items are used (e.g., true–false, multiple choice, matching). In the assessment presented in this paper, the analytic rubric with a top-down crafting method (and many revisions) was selected due to its increased objectivity and ability to target specific elements where students are excelling or having trouble.

Strategy for validation

The purpose of the work presented in this paper is to validate that two questions and associated analytic rubrics reliably measure students' design process knowledge. There is one question in which students critique a process to design a shopping cart and another where they critique a process to design an egg counter for eggs traveling down a conveyor belt. One master rubric was developed and then particularized to the two versions of the question.

The two questions were given as a pre- and posttest pair to students in the introduction to engineering design class in both Fall 2003 and Spring 2004 (total population size of approximately 300 students) [9]. Additionally, seniors in two different capstone classes completed both the shopping cart and egg counter tests back-to-back near the end of their two-semester classes (total population of 104 students). Some of the seniors took the shopping cart test first and others the egg counter first. Between these two sets of data—the

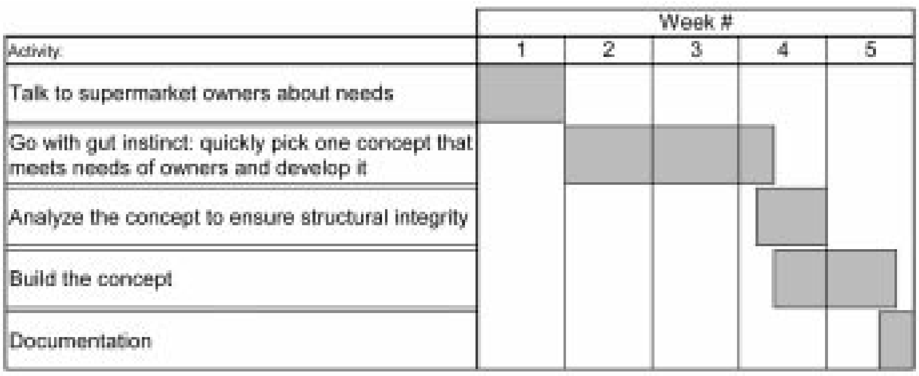

Fig. 2. Shopping cart question.

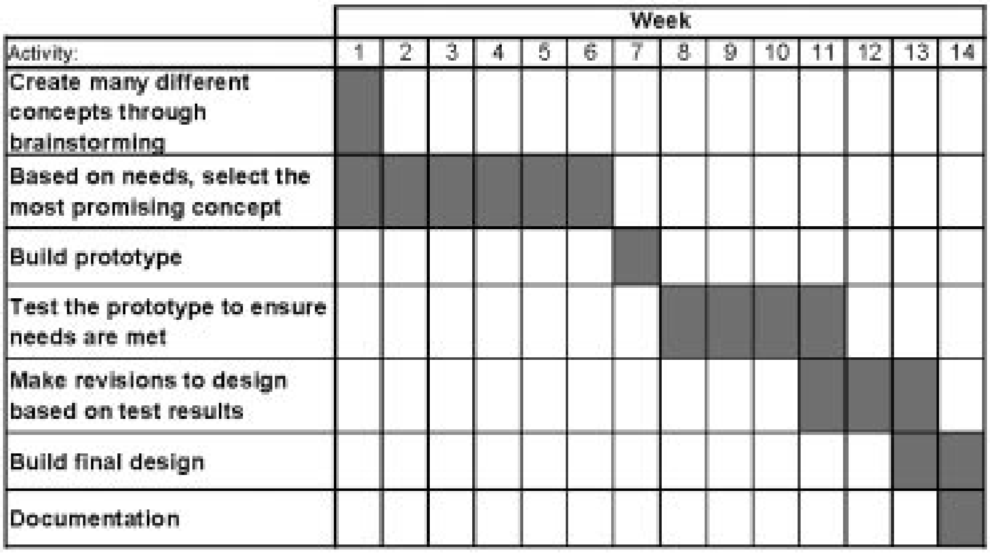

Fig. 3. Egg counter question.

introduction class and the capstone class—the validity of each question and of the analytic rubric is investigated.

Pre- and Post-tests

The shopping cart and egg counter tests are shown in Figs. 2 and 3. In each case, students are instructed to identify the strengths and weaknesses of the proposed process and to explain why something is a strength or weakness.

The two tests are purposefully developed to not have the same strengths and weaknesses. For instance, the process proposed for the shopping cart does not do a good job of generating several alternatives before selecting one to develop whereas the egg counter does do this well. As will be discussed in a later section, however, such differences led to difficulties in validating the questions.

Rubric

The rubric for each question is derived from a master rubric that is based on the common elements of design shown in Fig. 1. The rubrics are split into seventeen different levels, each focused on a different aspect of engineering design and a specific instructional objective. The shopping cart rubric is in the appendix of this paper.

Within the overall objective of 'students should be able to explain and analyze a design process in the context of the three phases shown in Fig. 1,' there are several more specific sub-objectives measured by the rubric. These objectives are as follows—students should be able to:

- *Explain* why needs must be gathered and *analyze* the effectiveness of techniques for gathering needs.
- *Explain* why multiple alternatives should be generated before developing a single alternative in depth.
- *Explain* that a combination of analysis and decision-making (based on the needs of the project) is required to eliminate ideas before building them and *analyze* the completeness of approaches used in analysis and decision-making.
- *Explain* that built designs should be tested to determine if they meet the needs.
- *Explain* how the three phases of design fit together and involve iteration, and *analyze* how much time is necessary for each step.
- *Explain* that documentation must occur throughout a design process.

Note how each instructional objective involves *explaining* and/or *analysis*—clearly tying these objectives to the second and fourth levels of Bloom's taxonomy. The nature of the questions is more strongly tied to the evaluation level on Bloom's taxonomy (asking students to critique a proposed design process). In this assessment, however, evaluation is used primarily as a means to elicit *explanations* and *analysis* of an engineering design process taught in class (which includes the common elements in the process shown in Fig. 1). By assessing at multiple levels of Bloom's, the tool will be able to measure students' progression along Bloom's levels as they progress from high school graduates to first year students completing an introduction to engineering class to seniors in a capstone class and finally to graduating seniors To demonstrate how the rubric assesses these instructional objectives, a level from the rubric is presented in Fig. 4.

In Fig. 4, the master rubric is in the column labeled 'Description' while the shopping cart scoring is in the column labeled 'Shopping Cart.' A response that indicates that doing analysis before building the design is good receives one point. If

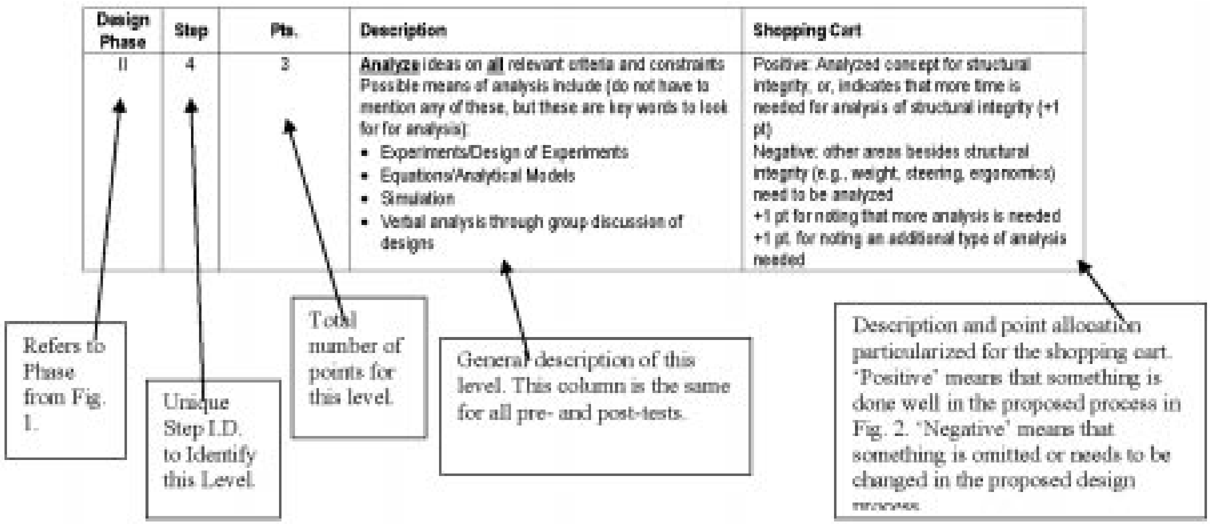

Fig. 4. Analysis level of rubric [9].

the response also indicates that other analyses, such as mass or ergonomic analyses, are needed, then they would receive all three points for this level of the rubric. This directly relates to the third instructional objective in the preceding list. Additional examples of how to score student responses are shown by Bailey *et al* [9].

RESULTS AND DISCUSSION

Statistical analysis of the data

Results from the statistical analysis of data collected in Fall 2003 and Spring 2004 seem to show that students in the introduction to engineering class learn statistically significant design process content across the semester (Fall 2003: $t = 5.14$, $df = 178$, $p < 0.0001$; Spring 2004: $t = 4.77$, $df = 125$, $p < 0.0001$). With these students, the shopping cart was used as a pre-test and the egg counter was used as a post-test. When both tests are given to seniors in one sitting, however, the average score for the egg counter was higher than that for the shopping cart. Furthermore, when these two tests were given in different orders to senior students, the results show that senior students obtained statistically significant different scores for the shopping cart ($t = 2.38$; $df = 104$, $p = 0.019$) but not statistically significant different scores for the egg counter ($t = 0.17$, $df = 104$, $p = 0.862$). This means that the order in which the tests are taken is important for the shopping cart but not for the egg counter. The results indicate that the two tests are not as parallel as intended and, consequently, assess different objectives.

Item correlation analysis was conducted and the results show that in both tests (shopping cart and egg counter) there are some items that do not work well in the rubric. There were five common items that did not work in either rubric. These five items are as follows:

Item 1: State that a team is needed to work on the project.
Item A: Indicate that the three phases of design are addressed in an appropriate order.
Item C: State that iteration should be planned into a design process.
Item H: Extra points for indicating a strength of the proposed process not listed elsewhere in the rubric.
Item I: Negative points for answers that directly oppose to a correct answer.

The split-half correlation coefficient (between the two tests) for students in the introduction to engineering class was 0.255, and for senior students the split-half correlation coefficient was 0.497. The results show that the two tests work better for the senior students (this is expected since senior students are supposed to know more content about engineering design). The common items that are not reliable show that the rubric is not well designed and that some items would work better if they were collapsed with similar items. The rubric with seventeen items is too detailed for grading students' short answer responses (10–15 minutes response time). Even though it is possible to train raters that are not familiar with the engineering field to use the analytic rubric [10], the training is longer and needs many exercises in order to improve interrater reliability. A shorter rubric would be more advantageous in training the raters and would also decrease grading time.

In conclusion, there are two main problems identified with the statistical analysis. First, the two tests—the egg counter and the shopping cart—measure different things. This is shown with the data from seniors who took the two tests back-to-back. Second, the rubric needs to be redesigned and the rubric items improved. This is indicated with data from both sets of students using item correlation analysis.

	Week													
Activity:	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Create many different concepts through brainstorming	■													
Based on needs, select the most promising concept	■	■	■	■	■	■								
Build prototype shopping cart							■							
Test the prototype cart to ensure needs are met								■	■	■	■			
Make revisions to design based on test results											■	■	■	
Build final shopping cart													■	■
Documentation														■

Fig. 5. Revised shopping cart question.

Next Phase: Correcting the identified problems

In the next phase of this research, problems identified in the first phase will be addressed. To address the problem between the two tests, the shopping cart test has been changed to be more parallel to the egg counter test. Whereas the current tests show processes with different strengths and weaknesses, the new shopping cart test will have the same strengths and weaknesses as the egg counter test. The new shopping cart test is shown in Fig. 5.

The process shown in Fig. 5 is nearly identical to that of the egg counter design shown in Fig. 3, with the only differences being that 'shopping cart' is added to a few of the task names.

To address problems with the rubric, the seventeen levels of the rubric have been collapsed into seven levels. This new rubric will be sufficient to measure the instructional objectives for the short answers provided by students and will be more efficient in training raters and in grading. Several

Table 3. Revised scoring rubric

	0	1	2	3	4
1 NEEDS	No mention of lack of gathering needs 0 pts	State that gathering needs is important 1 pt	State that needs should be gathered before brainstorming 2 pts		Gives a suggestion as to how to find needs before brainstorming 4 pts
2 IDEA GENERATION	No mention of brainstorming/idea generation 0 pts		States that it is good that multiple concepts are created or that brainstorming is good/needs more time 2 pts		
3 ANALYSIS AND DECISIONS	No mention of analysis or decision-making/selection 0 pts		States that device should be analyzed before making a decision - OR - States that it is good to base decision on needs 2 pts		States that device should be analyzed before making a decision - AND - States that it is good to base decision on needs 4 pts
4 BUILDING AND TESTING	No mention of building or testing 0 pts	States that building (either the prototype or final design) is good or that more time is needed for building (no mention of testing) 1 pt	States that it is good to test the prototype, that enough time is allotted for testing, or that it is good to be tests to the needs.* - OR - States that a test of the final design is needed 2 pts	Satisfies both the "1 point" and "2 point" levels 3 pts	States that it is good to test the prototype and that the final design needs to be tested 4 pts
5 LAYOUT AND ITERATION	No mention of overall layout or iteration or simply says that the plan is "organized." 0 pts		States that the overall plan is LOGICAL or follows the 3 phases of design - OR - States that it is good to include the iteration/revisions (or that more time is needed for iteration/revisions) 2 pts		States that the overall plan is LOGICAL or follows the 3 phases of design - AND - States that it is good to include the iteration/revisions (or that more time is needed for iteration/revisions) 4 pts
6 TIME	No mention of time allotments 0 pts	States that time is not allocated well among the steps of the process 1 pt	Explains where time is not allotted well (e.g., any of the following would count): * too much for concept selection * not enough for building * not enough brainstorming time * too much testing time 2 pts		
7 DOCUMENTATION	No mention of documentation 0 pts	States that more time is needed for documentation 1 pt	States that documentation needs to occur throughout the process, not just at the end 2 pts		

of the problem areas on the rubric (Items 1, A, H, and I) have been removed, whereas other parts of the rubric are joined to reduce the total number of levels. The new rubric is shown in Table 3.

Several features of the revised rubric are noteworthy. First, it directly addresses the instructional objectives and simplifies the scoring. For example, instead of two levels for testing and building as in the old rubric, now there is one. All but one instructional objective is measured with a single level of the rubric (with the only exception being that levels five and six are both used to measure a single instructional objective). Additionally, the scoring within each level has been simplified to a small number of discrete levels.

The new rubric, with its smaller set of clearly defined levels, is more commensurate in detail with the assessment question. Furthermore, its simplicity should increase the scoring reliability both within and between raters.

CLOSURE

Because design process knowledge is less concrete than most of engineering, assessing if students are learning it is very difficult. The first phase of development of an assessment strategy has been completed and statistical analysis indicates that changes are necessary to increase the validity of the tool. The analysis gives clear direction with respect to areas that need adjustment. The questions asked to the students need to be more parallel in structure in the pre- and post-tests and the rubric used to score responses needs to be simpler. These changes will be implemented in phase two of the research.

Acknowledgements—The authors would like to thank the Dr. Darrell Sabers and the students participating in EdP 358 in Fall 2003 and Spring 2004. We would also like to thank the National Science Foundation who supported this work through grant EEC-0338634.

REFERENCES

1. S. B. Bloom, D. R. Krathwohl, and B. B. Masia, *Taxonomy of Educational Objectives; The Classification of Educational Goals,* Longmans, Green, New York, (1956).
2. L. W. Anderson, D. R. Krathwohl, P. W. Airasian, K. A. Cruikshank, R. W. Mayer, P. R. Pintrich, *et al. A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives,* Longman, New York, (2001).
3. A. Woolfolk, *Educational Psychology* (9th edn), Allyn and Bacon, Boston, MA, (2003).
4. J. Sims-Knight, R. Upchurch, N. Pendergrass, T. Meressi, and P. Fortier, *Assessing Design by Design: Progress Report 1.* Paper presented at the Frontiers in Education, Boulder, CO, (2003).
5. Newstetter, W. C. and Kahn, S. *A Developmental Approach to Assessing Design Skills and Knowledge,* Paper presented at the Frontiers in Education, Pittsburgh, PA, (1997).
6. M. P. Brackin, and J. D. Gibson, *Methods of Assessing Student Learning In Capstone Design Projects With Industry: A Five Year Review,* Paper presented at the American Society for Engineering Education Annual Conference, Montreal, Canada, (2002).
7. R. J. Gentile, *Educational Psychology* (2nd edn), Kendall, Dubuque, IA, (1997).
8. Nitko, A. J. *Educational Assessment of Students* (4th edn), Upper Saddle River, Prentice Hall, New Jersey, (2004).
9. R. Bailey, Z. Szabo, and D. Sabers, *Assessing Student Learning about Engineering Design in Project-Based Courses,* Paper presented at the American Society for Engineering Education Annual Conference, Salt Lake City, Utah, June 20–23 2004.
10. R. Bailey, Z. Szabo, and D. Sabers, *Integrating Education Students in the Assessment of Engineering Courses,* Paper presented at the American Society for Engineering Education Annual Conference, Salt Lake City, Utah, 2004, June 20–23.

APPENDIX

The shopping cart rubric is presented in Tables A.1 and A.2.

Table A.1 First eight levels of shopping cart rubric

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 A.2 Final nine levels of the shopping cart rubric

	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)	

Reid Bailey is an Assistant Professor in the College of Engineering at the University of Arizona. His research interests include engineering design, environmental issues affecting design, and engineering education. He received his BS from Duke University and both his MS and PhD from the Georgia Institute of Technology.

Zsuzsanna Szabo is Assistant Professor in the Educational Leadership Department at the Southern Illinois University at Edwardsville. Her research interests include team learning, assessment and gender issues in education. She received her BS in Civil Engineering from Technical University Cluj, and BS in Psychology from University Babes-Bolyai, both in Romania, her MEd from SUNY at Buffalo, NY, and PhD from the University of Arizona. She has also worked for nine years as a civil engineer.