Capstone Design Outcome Assessment: Instruments for Quantitative Evaluation

David G. Meyer
Purdue University, School of Electrical and Computer Engineering
West Lafayette, IN 47907 meyer@purdue.edu

***Abstract* - For capstone design experiences, the course outcomes, assessment strategies, and outcome remediation strategies are significantly different than those that might typically be utilized in lower-division "content" courses. This paper describes the instruments developed for a *Digital Systems Senior Project* course that provide a mechanism for systematic, quantitative evaluation of outcomes appropriate for capstone design experiences. Data tracking the performance of these instruments over several trials are presented.**

Index Terms - ABET 2000, capstone design, evaluation instruments, outcome assessment, remediation strategies.

INTRODUCTION

Virtually all engineering degree programs feature some form of capstone design experience. Many involve teamwork, require demonstration of communication skills, develop an awareness of professional responsibility, and typically require the design, construction, debugging, etc., of a device or system. For the purpose of accreditation, engineering programs must demonstrate that their graduates have [1]:

- (a) an ability to apply knowledge of mathematics, science, and engineering
- (b) an ability to design and conduct experiments, as well as to analyze and interpret data
- (c) an ability to design a system, component, or process to meet desired needs *within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability*[1]
- (d) an ability to function on multi-disciplinary teams
- (e) an ability to identify, formulate, and solve engineering problems
- (f) an understanding of professional and ethical responsibility
- (g) an ability to communicate effectively
- (h) the broad education necessary to understand the impact of engineering solutions in a global, *economic, environmental*[*], and societal context
- (i) a recognition of the need for, and an ability to engage in life-long learning
- (j) a knowledge of contemporary issues
- (k) an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice.

In this context, a capstone design course must provide students with "…a major design experience based on the knowledge and skills acquired in earlier course work and incorporating appropriate engineering standards and multiple realistic constraints." [1] Although not explicitly required, most engineering capstone design experiences are team-based (based on Criterion 3d). The experience should also reinforce the students' understanding of ethical and professional responsibility (based on Criterion 3f) and their ability to communicate effectively (based on Criterion 3g). Realistic constraints (e.g., economic, environmental, social, political, ethical, health and safety, manufacturability, sustainability) should be employed as well (based on Criterion 3c).

CAPSTONE DESIGN OUTCOME ASSESSMENT

According to [2], "…engineering programs must have in place an appropriate assessment process that produces documented results that demonstrate that students have achieved each and every item listed in (a) through (k). It is expected that all students will demonstrate achievement of every item listed in (a) through (k). Programs must show, even by appropriate sampling, that there is convincing evidence to assume that all students by the time they have graduated have demonstrated achievement, to a level acceptable to the program, of every item listed in (a) through (k). However, it is not necessary for evidence to be provided for each and every student." Further, "student self-assessment, opinion surveys, and course grades are not, by themselves or collectively, acceptable methods for documenting achievement of outcomes."

In [3], the authors provide an overview of the methodologies that are available for use in assessing undergraduate engineering programs, along with research questions associated with these methodologies that are currently outstanding. To measure how well engineering students can apply classroom knowledge and skills to realistic design problems, *authentic assessment* and *performance-based assessment* methods can be used. The key to authentic assessment, according to [3], is to "create a context in which the student can individually or collaboratively demonstrate an ability to apply a well-developed problem-solving strategy" which might involve "problem definition, gathering relevant information, generating solution alternatives, choosing the optimum solution given implicit and explicit constraints, assessing and improving the proposed solution, and effectively reporting results…"

[1]Italicized wording added for the 2005-2006 accreditation cycle.

Some of the outstanding research issues associated with authentic assessment, cited in [3], include "development of well-designed scoring rubrics and methods for ensuring inter-rater reliability." The authors go on to state that "guidelines also need to be developed which help faculty choose tasks that are good candidates for collecting authentic assessment data in engineering courses." These are precisely the areas to which the work reported here contributes.

A national survey of design courses and assessment is reported in [4]. The purpose of the study was to obtain a better understanding of the nature and scope of assessment practices within capstone design courses across engineering disciplines, and in particular, the extent to which current practices align with ABET EC 2000 expectations. Findings suggest uncertainty on the part of many faculty members concerning sound assessment practices, including using appropriate assessment strategies. Faculty identified several ways they wanted to improve the quality of their capstone assessments. About one-half of the respondents felt the measures should be more objective, wanting to develop more detailed scoring guidelines/rubrics and desiring clearer performance criteria. A significant number of respondents also wanted to increase the variety of assessment instruments. These are issues the work reported here attempts to address.

Related work includes [5], where the authors describe an ECE capstone design experience and detail how project proposals are formulated and grades for the two-course sequence are determined. The authors do not, however, address the issue of design course outcome assessment. Another capstone design course evaluation strategy is described in [6]. Similar to [5], the authors provide a detailed strategy for grading a capstone design course (here, in Civil Engineering), but again do not address the issue of outcome assessment *per se*. Methods of assessing student learning in capstone design projects sponsored by industry are outlined in [7]. Here the emphasis is on expanding the project evaluation beyond the design report to include teaming skills as well as technical competence. Instruments used include company evaluations, status reports, student self-assessments, peer reviews, and oral reports (in addition to the traditional design report) to quantify student performance both as team members and design engineers. Unlike [5] and [6], no mention of design course grade determination is made in [7]; instead, the focus is on assessing teaming skills and technical competence over multiple trials.

A different means of measuring capstone design project outcomes is described in [8]. Here, the focus is on *quality* measurements (in the context of an industrial sponsor), specifically the development of two distinct instruments intended to measure the quality of a design outcome: the Client Satisfaction Questionnaire (CSQ) and the Design Quality Rubric (DQR). While these are valuable tools for measuring the *quality* of the end-product (project deliverables), they do not specifically address the *assessment* of specific *design course outcomes*. In [9], the authors describe a method for assessing design *knowledge* (i.e., students' knowledge of what constitutes the *design process*) which might be construed as the complete antithesis of the methodology described in [8]. While it is interesting to understand how a given student's knowledge about design changes from his/her freshman year to senior year, this work does not specifically address the issue of capstone design course outcome assessment.

An illustration of how ABET assessment can be conducted within the framework of a capstone design course sequence (rather than on specific project topics themselves) is given in [10]. A variety of assessment techniques are detailed that provide both quantitative measurements and qualitative indicators that can be used to demonstrate achievement of outcomes (as well as to improve the design course sequence and the curriculum as a whole). While closely related to the work reported here, there are some important differences. First, in the work reported here, specific learning outcomes germane to capstone design are defined; in [10], the ABET Criterion 3 Program Outcomes are used directly. A systematic, quantitative strategy for assessing each of the design course learning outcomes is described in the work reported here; in [10], a "mapping" of instruments and criteria is provided, along with an algorithm for course grade determination.

Assessment of outcomes in capstone design courses requires different strategies than in lower-division "content" courses. A good discussion of "what works and what doesn't" can be found in [11]. As detailed in [12], the first decision toward effective outcome assessment in content courses is choice of evaluation instrument(s). Possibilities include exams (whole or question subsets), quizzes (written/oral), homework assignments, labs, and papers/presentations. Here, proctored exams/quizzes have proven to be most effective. For capstone design courses, however, there typically are no "exams" *per se*. Here, project deliverables such as papers, presentations, lab notebooks, and device functionality ("project success criteria satisfaction") are generally the operable evaluation instruments.

Another important issue is determination of outcome demonstration "passing thresholds". Choices include *static* thresholds (plus what absolute value to choose) or *dynamic* thresholds (plus what algorithm should be used to "adjust" them). For content courses there is a fairly delicate balance between establishing reasonable, meaningful thresholds for outcome demonstration success that are decoupled to the extent possible from the "exam difficulty" (and other factors beyond the instructor's control, such as which other courses have exams during the same time period). Here, a proven strategy is to apply dynamic thresholds on exams [12]. Use of EXAM MEAN – STANDARD DEVIATION as the dynamic threshold (limited to the range of 40% to 60%) has been shown to produce meaningful, predictable results. In the study documented in [12], approximately 80-90% of the students were able to successful demonstrate a given outcome on the primary outcome assessment exam based on use of such a dynamic threshold, while typically 90-95% were able to successfully demonstrate that outcome given a second opportunity (referred to as the final assessment).

Noting that the evaluation instruments typically used in capstone design courses are significantly different than those used in lower-division content courses, it makes sense that a different threshold strategy would be more appropriate as well. If the evaluation instruments involve "standardized" grading rubrics (i.e., that utilize a pre-determined set of evaluation criteria, scores, and weights), then application of fixed thresholds is appropriate. Note that different static thresholds might be applicable depending on the evaluation instrument utilized.

A quantitative strategy for outcome assessment in a computer engineering capstone design course is presented in this paper that exemplifies these requirements. First, the course is described along with the stated learning outcomes. Next, the evaluation instruments used to quantitatively assess each outcome are illustrated. Finally, data tracking the performance of these instruments over several trials are presented. Observations concerning the results and implications for possible improvements conclude this paper.

CAPSTONE DESIGN COURSE SPECIFICS

In ECE at Purdue, there are currently three senior design options: (a) the *EE Design Project* course, in which all teams design and construct the same device during a given semester (targeted primarily at Electrical Engineering degree option students); (b) the *Digital Systems Design Project* course [13], where each team can design an embedded microcontroller based system of their choice (subject to instructor approval); and (c) the *Engineering Projects in Community Service* (EPICS) [14] course sequence, where teams of students work on service learning projects (of their choice, from the available set) that span multiple semesters.

All ECE capstone design options at Purdue have the following learning outcomes in common (tied to the overall program outcomes indicated in parenthesis):

1. an ability to apply knowledge obtained in earlier coursework and to obtain new knowledge necessary to design and test a system, component, or process to meet desired needs (a,b,c,e,i,j,k)
2. an understanding of the engineering design process (b,c,e,f,h)
3. an ability to function on a multidisciplinary team (d,h,j)
4. an awareness of professional and ethical responsibility (f,h,j)
5. an ability to communicate effectively, in both oral and written form (g).

The discussion that follows will focus primarily on the *Digital Systems Design Course* option. This course is advertised as "a structured approach to the development and integration of embedded microcontroller hardware and software that provides senior-level students with significant design experience applying microcontrollers to a wide range of embedded systems (e.g., instrumentation, process control, telecommunication, intelligent devices, etc.)." The fundamental course objective is to provide practical experience developing integrated hardware and software for an embedded microcontroller system in an environment that models one which students will most likely encounter in industry.

A unique feature of this course, compared to the other two ECE senior design options at Purdue, is that students are able to *choose* the embedded, microcontroller-based system they ultimately design (subject to instructor approval). The basic constraints imposed are that the system design must utilize an "approved" microcontroller (typically a PIC, Rabbit, Atmel, HCS12, or ARM variant), meaningfully incorporate several "standard" interfaces (e.g., I^2C, SPI, TCP/IP, RF, IR, Bluetooth, X10, etc.), be implemented on a custom-designed printed circuit board, be neatly (and appropriately) packaged, be of personal interest to at least one team member, and (last but not least) be *tractable*. To this end, each team of four students prepares and submits a project proposal (in "draft" and then "final" form, following an initial review). Included in this proposal are five, student-specified project success criteria by which the system functionality will be judged (there are five additional success criteria that are common to all projects, covering deliverables such as the schematic, the bill of materials, the printed circuit board layout, the packaging, and system integration).

Quantifying the assessment of the inherently qualitative course outcomes (listed previously) and determining appropriate thresholds to apply has been a major challenge. The "breakthrough" in quantifying the assessments of Outcomes 1 and 4, respectively, was creation of a series of four *design component* and four *professional component* "homework assignments" (in actuality, written reports that serve as the precursor of corresponding sections in the final written report). Implementation of this strategy requires a "fixed" team size of four members and a corresponding class enrollment that is an integer multiple of four. Here, each team member is required to pick one topic from each set to individually research and produce a formal written report, complete with references. Together, the two reports constitute a significant portion of each student's grade (20%). Grading rubrics, such as the one illustrated in Figure 1, are used to evaluate the reports. A fixed threshold of 60% is the minimum requirement for successful outcome demonstration.

The design component reports are as follows:

1. Packaging Specifications and Design
2. Schematic and Hardware Design Narrative/Theory of Operation
3. Printed Circuit Board Layout
4. Firmware Listing and Software Narrative

The professional component reports are as follows:

1. Design Constraint Analysis and Component Selection Rationale
2. Patent Liability Analysis
3. Reliability and Safety Analysis
4. Social/Political/Environmental Product Lifecycle Impact Analysis

Component/Criterion	Score (0-10)	Wgt	Pts
Introduction		X 1	
Results of Patent Search		X 3	
Analysis of Patent Liability		X 3	
Action Recommended		X 1	
List of References		X 1	
Technical Writing Style		X 1	

FIGURE 1. GRADING RUBRIC FOR A PROFESSIONAL COMPONENT REPORT.

For Outcome 2 ("understanding of the engineering design process"), multiple evaluations of the individual lab notebooks provide a meaningful quantitative measure of successful demonstration. The breakthrough here was to create group accounts and team websites that hosted each member's on-line laboratory notebook. Adoption of this approach allowed the course staff to conveniently check on team progress as well as individual contributions. Further, the web-based approach allowed students to include hyperlinks in their notebook entries to photos of prototyping setups, source code for testing various interfaces, video demos of project specific success criteria fulfillment, PDFs of data sheets used in the design, etc. The relatively simple grading rubric shown in Figure 2 is used to evaluate the notebooks (at midterm, following the formal design review; and at the end of the semester, following the final presentation). Together, the lab notebook evaluations count for 10% of the course grade. A minimum average score of 60% is required for successful demonstration of Outcome 2.

Component/Criterion	Score (0-10)	Wgt	Pts
Technical content		X 3	
Update record/completeness		X 2	
Professionalism		X 3	
Clarity/organization		X 2	

FIGURE 2. GRADING RUBRIC FOR THE LABORATORY NOTEBOOK.

For Outcome 3 ("ability to function on a multidisciplinary team"), the project success criteria provide a meaningful quantitative measure of successful demonstration. As part of the final presentation, each team prepares video clips illustrating success criteria satisfaction (these videos are also posted on the team web sites). Upon review, the course staff assigns a score (worth 10% of the course grade) based on an analysis of how completely the success criteria have been met. Here, a minimum threshold of 80% is required to establish basic competency (i.e., the project must be reasonably functional and capable of producing the specified behavior).

Finally, demonstration of Outcome 5 ("ability to communicate effectively, in both oral and written form") is based on the Design Review, the Final Presentation, and the Final Report. A minimum score of 60% on the Design Review <u>and</u> a minimum score of 60% on the Final Report <u>and</u> a minimum score of 60% on the Final Presentation is required to establish basic competency for this outcome.

In summary, evaluation instruments that have been chosen to quantitatively evaluate the five capstone design learning outcomes include:

1. a *design component homework* (to evaluate "an ability to apply knowledge obtained in earlier coursework and to obtain new knowledge necessary to design and test a system, component, or process to meet desired needs").
2. the *individual lab notebook* (to evaluate "an understanding of the engineering design process").
3. the *project success criteria* (to evaluate "an ability to function on a multidisciplinary team").
4. a professional component homework (to evaluate "an awareness of professional and ethical responsibility").
5. the *formal design review, final presentation,* and *final written report* (to evaluate "an ability to communicate effectively, in both oral and written form").

This quantitative assessment strategy has been used for five consecutive offerings of the *Digital Systems Senior Design* course described in this paper; the average scores produced by the evaluation instruments for each cohort group are listed in Table I. The typical cohort size for each trial is 48 (12 teams).

TABLE I. COHORT AVERAGES FOR EACH COURSE OUTCOME.

Outcome	Spr-03	Fall-03	Spr-04	Fall-04	Spr-05
1	85.5%	79.0%	81.7%	85.9%	80.8%
2	72.0%	81.3%	74.9%	84.7%	77.1%
3	93.3%	87.5%	85.0%	91.7%	91.4%
4	82.1%	81.5%	80.2%	84.6%	77.4%
5	85.7%	87.3%	85.9%	87.7%	85.3%

Several observations can be made based on the cohort averages recorded. First, it appears the evaluation instruments used for the formal presentations and final report (Outcome 5) have produced the most consistent results. Further, while there is some notable variation semester-to-semester, within a given offering the assessments of Outcomes 1 and 4 (the professional and design component reports) appear to be fairly coherent. Because the course staff (typically consisting of two professors and two teaching assistants) varies from semester-to-semester, some variation in average score for these outcomes is expected. Some "natural variation" in the average score for Outcome 3 (the project success criteria) is also expected, due to the wide variety of designs attempted for this course. The assessment and/or "quality" of Outcome 2 (the individual lab notebooks), however, could most likely be improved. In part, the large variation in average score is due to the widely varying quality of the lab notebooks – some lack sufficient technical details, while others look more like "blogs". Use of tablet computers has the potential for improving the timeliness of the updates and the quality of the entries. These are areas of concern for which solutions are actively being pursued.

OUTCOME REMEDIATION AND GRADE DETERMINATION

Despite being seniors about to be unleashed on the "real world" with degrees in hand, their "first attempt" at writing a formal, technical report may not always meet the minimum standards (and if they did, the application of thresholds would be meaningless). Opportunities for outcome remediation therefore need to be provided. Unlike the "content" courses referred to earlier in this paper (where outcome remediation is typically accomplished by providing a second exam over the same material), students who initially fail to demonstrate an outcome (e.g. receive a score on a design or professional homework below the prescribed passing threshold) must be given an opportunity for prompt remediation (e.g. rewriting the deficient paper, correcting the printed circuit board layout, etc.). One complication that arises is how to "count" the updated score toward the course grade. To prevent "abuse of the system", it is probably wise to either average the original score with the revised score or award a nominally passing grade for the repeat submission. Experience has shown that if the higher score simply replaces the former score, students will quickly sense this "loophole" and exploit it to "buy time".

Another issue related to remediation is the question of how to handle cases in which all the course outcomes have not been successfully demonstrated, yet the student has otherwise earned a passing grade. Unlike content courses, where repeating a course is a viable (and often the best) option, students in capstone design courses who are still deficient in a particular outcome (e.g., Outcome 3, due to a hardware or software "bug" that the team has been unable to resolve) should initially be awarded a grade of "I" (incomplete) and provided with a reasonable timetable to resolve the deficiency. If a team or individual is deficient in more than one outcome, however, then the best option might be requiring those individuals to repeat capstone design (perhaps electing an alternate option).

Another challenge for a team-oriented project course is ensuring equitable distribution of workload and grade determination based on both individual as well as corporate contributions. To this end, 50% of the course grade is based on team components, while the remaining 50% is based on individual components. The complete grading breakdown for the *Digital Systems Design Project* course is given in Table II.

TABLE II. WEIGHTS OF GRADING COMPONENTS.

TEAM COMPONENTS		INDIVIDUAL COMPONENTS	
Design Review	10%	Individual Contribution	10%
Final Video Presentation	10%	Lab Notebook Evaluations	10%
Final Report & Archive CD	10%	Design Component Rpt	10%
Project Success Criteria	10%	Professional Component Rpt	10%
Project Proposal	2%	OrCAD Exercise	2%
User Manual	3%	Presentation Peer Review	4%
Senior Design Report	2%	Confidential Peer Review	2%
Poster	3%	Weekly Progress Briefings	2%

SUMMARY AND CONCLUSIONS

Different kinds of courses (in particular, "content" vs. "capstone design") require different outcome assessment strategies, and finding the "best practices" for each case is non-trivial. As documented in [4], many assessment strategies have been employed in capstone design courses, yet uncertainty persists concerning sound practices. This paper has presented a systematic, quantitative strategy for assessing capstone design course outcomes and integrating the outcome assessment with course grade determination. Data from five consecutive trials show that meaningful results can be obtained despite inter-rater differences. Effective application of outcome assessment (using appropriate evaluation instruments, outcome demonstration success thresholds, and grading strategies) can truly *promote and help learning.*

REFERENCES

[1] http://www.abet.org/Linked%20Documents-UPDATE/Criteria%20and%20PP/05-06-EAC%20Criteria.pdf

[2] http://www.abet.org/Linked%20Documents-UPDATE/Program%20Docs/EAC%20Guidelines%20for%20Criterion3.pdf

[3] Atman, C. J., *et al.*, "Matching Assessment Methods to Outcomes: Definitions and Research Questions," 2000 *American Society for Engineering Education Conference Proceedings.*

[4] McKenzie, L. J., Trevisan, M. S., Davis, D. C., and Beyerlein, S. W., "Capstone Design Courses and Assessment: A National Study," 2004 *American Society for Engineering Education Conference Proceedings.*

[5] Gesink, J., and Mousavinezhad, S. H., "An ECE Capstone Design Experience," 2003 *American Society for Engineering Education Conference Proceedings*.

[6] Quadrato, C., and Welch, R. W., "Grading Capstone Design: On Time and On Target," 2003 *American Society for Engineering Education Conference Proceedings*.

[7] Brackin, M. P., and Gibson, J. D., "Methods of Assessing Student Learning in Capstone Design Projects with Industry: A Five Year Review," 2002 *American Society for Engineering Education Conference Proceedings*.

[8] Sobek, D. K., and Jain, V. K., "Two Instruments for Assessing Design Outcomes of Capstone Projects," 2004 *American Society for Engineering Education Conference Proceedings.*

[9] Caso, R., Lee, J. H., Froyd, J., and Kohli, R., "Development of Design Assessment Instruments and Discussion of Freshman and Senior Design Assessment Results," 2002 *Frontiers in Education Conference Proceedings*.

[10] Davis, K. C., "Assessment Opportunities in A Capstone Design Course," 2004 *American Society for Engineering Education Conference Proceedings*.

[11] Jenkins, M. G., and Kramlich, J. C., "Assessment Methods under ABET EC2000 at the University of Washington – Lessons Learned: What Works and What Doesn't," 2002 *American Society for Engineering Education Conference Proceedings.*

[12] Meyer, D. G., "Outcome Assessment: Practical Realities and Lessons Learned", 2004 *Frontiers in Education Conference Proceedings*.

[13] http://shay.ecn.purdue.edu/~dsml/ece477

[14] Oakes, W. C., Jamieson, L. H., and Coyle, E. J., "EPICS: Meeting EC 2000 through Service-Learning," 2001 *American Society for Engineering Education Conference Proceedings.*