2006-1249: LESSONS LEARNED IN ASSESSING SENIOR ENGINEERING CAPSTONE DESIGN COURSE LEARNING WITH A VARIATION ON THE TIDEE DESIGN TEAM READINESS ASSESSMENT (DTRA) I AND II

Martha Green, Texas A&M University

Martha Green is an Educational Psychology PhD student at Texas A&M University,with an M.Ed. in Educational Technology and BA/BS degrees in History and Business/Accounting. Formerly a K-12 teacher, she is a Graduate Assistant in the Educational Assessment Resources unit in the College of Engineering at TAMU, was assessment consultant to the NASA Reduced Gravity Student Flight Opportunities Program and is managing assessment of the BOEING Project

Parag Ravindran, Texas A&M University

Parag Ravindran is a PhD student in Mechanical Engineering at Texas A&M University with both M.S. and B.E. degrees in Mechanical Engineering. He is currently a Teaching Assistant and was formerly a Graduate Assistant Analyst in the Educational Assessment Resources Unit of the College of Engineering at TAMU.

Rita Caso, Texas A&M University

Rita Caso is the Director of Institutional Research and Effectiveness for Sam Houston State University and was previously Deputy Director of the Division of Academic Development in the College of Engineering at Texas A&M University, directing its Educational Assessment Resources unit. She has been a principal researcher and evaluator for several NSF initiatives in the areas of engineering and science education.

Jefferey Froyd, Texas A&M University

Jeffrey Froyd is a Research Professor in the Center for Teaching Excellence and Director of Academic Development at Texas A&M University. He was Project Director for the Foundation Coalition, one of the NSF Engineering Education Coalitions and Project Director for "Changing Faculty through Learning Communities," a project sponsored by the NSF Research on Gender in Science and Engineering Program.

Xiafeng Li, Texas A&M University

Xiafeng Li is a PhD student of computer science at Texas A&M University. He got his B.S. from Jingdezhen Ceramic Institute in mechanical engineering in 1997, and M.S. from Shanghai Jiaotong University in computer engineering in 2001. His research areas include computer network, bin packing, and statistical data analysis.

L. Alan Minnick, Texas A&M University

L Alan Minnick (retired) was an educational research technician for the Educational Assessment Resources unit in the College of Engineering at Texas A&M University with previous experience in testing and measurement at the State University of California, Hayward. He participated in the scoring of engineering design assessments at TAMU which used several versions of the TIDEE assessment instruments.

Ram Shukla, Texas A&M University

Ram Shukla is currently pursuing PhD in the Department of Industrial and Systems Engineering, Texas A & M University. He has a Master of Industrial Engineering from this department. He also holds an MBA with specialization in Finance from the University of Lucknow and a Bachelors of Mechanical Engineering-Honors from University of Gorakhpur in India. He has over 7 years of combined teaching and industry work experience.

Lessons Learned in Assessing Senior Engineering Capstone Design Course Learning with a Variation on the TIDEE Design Team Readiness Assessment I and II

Abstract:

Efficacy of engineering design education in an interdisciplinary team-based course setting, with exposure to the broader concerns of business, finance and management, interests many educators. This paper reports on the use of design knowledge assessment patterned after the Transferable Integrated Design Engineering Education (TIDEE) Design Team Readiness Assessment (DTRA) I and II [1,2,3,4] to evaluate what aerospace, mechanical, electrical and computer science students learned about engineering design as a result of participating in a Boeing-supported, two-semester, project-based senior capstone design course at Texas A&M University. The Boeing course sequence was designed to provide students with the experience of working in multi-disciplinary teams on a design and implementation project while receiving instruction in elements of project management from a business perspective. Students were assessed using rubric-based instruments and customized team design simulation procedures similar to Parts I & II of the TIDEE Design Knowledge Assessment at the beginning and at the end of their senior capstone design project. This paper will discuss the lessons learned about the design knowledge assessment process to measure the preparedness of students at the beginning of the capstone design class and the evolution of their understanding of the design process at the end of the two-semesters of course work. Pre-to-post treatment group comparisons and control group comparisons were used in the assessment design, which also compared the rubric-based assessment with content analysis of students' written responses to assessment questions. Results indicated that the content analysis was more effective in capturing students' increased understanding of the design process, and the reasons for this are explored. Specific changes for improving the assessment process have been implemented for the project in the 2005-06 academic year.

I. Introduction

Many engineering programs incorporate a design course in the senior year as a culmination and refining process for the engineers they graduate [5-7]. Driven by industry demands, there have been attempts to improve the quality of the design education of engineers by incorporating increasingly more authentic, professional design conditions and incorporating elements of business, finance and management in the capstone design class [8-9]. The aim is to produce engineers who are more productive earlier in their careers. But attempts to ascertain the efficacy of these efforts require more tools and wider application of these tools [10–11]. For example, McKenzie et al. wrote that "…faculty members suggested that they lacked information and know-how to develop assessments for all users, write clear and appropriate course objectives, and determine whether assessments used in courses are as fair as desired" [9]. This paper examines the adaptation and development of a specific tool for assessing changes in student performance and cognitive growth in interdisciplinary and single-discipline capstone design courses in the Dwight Look College of Engineering at Texas A&M University.

II. Methodology

An instrument, based on Transferable Integrated Design Engineering Education (TIDEE) and the Design Team Readiness Assessment (DTRA) I&II, was developed and administered to students during the 2004-05 academic year. It was administered at the beginning (pre-test) and end (post-test) of the two-semester capstone design sequence. The students were split in two groups: an experimental group and a comparison group. The experimental group consisted of students from four majors, Aerospace, Mechanical, Electrical Engineering and Computer Science, who were enrolled in an interdisciplinary capstone sequence in which students from different disciplines participated on the same project team. The students were provided additional instruction in aspects of project management and finance by a professor and graduate teaching assistants from the Mays Business School at Texas A&M University. The comparison group consisted of students from Aerospace and Mechanical Engineering who were in conventional, single-discipline capstone design courses. Students in the comparison group did not receive any instruction in business.

Differences in experimental and comparison group sizes, team composition and stability posed major limitations in the study. These differences were due to smaller class size for the interdisciplinary experimental class, attrition and students changing classes from fall to spring in both the and the conventional courses. For example, the experimental group consisted of fourteen students (10 males, 4 females; 9 aerospace engineering majors, 5 mechanical engineering majors) who completed both the pre and post tests. In the fall semester of the 2004-05 academic year, faculty members assigned these students to five different teams. However, due to attrition and enrollment changes from fall to spring, some students in the experimental group did not take the pre and post tests as members of the same team. One of the five teams had to be eliminated from the analysis since it had only one student who had completed both the pre and post tests. On the other hand, the comparison group was considerably larger since enrollment in the single-discipline capstone courses was typically larger. The comparison group consisted of 43 students (38 males, 5 females; 23 aerospace engineering majors, 20 mechanical engineering majors) who completed both the pre and post tests. Since enrollment in the single-discipline course was larger, team sizes were larger. Faculty members assigned these 43 students to seven different teams. Differences in team size and stability and the built-in differences in student majors between experimental and conventional comparison group might have contributed to differences in results, but these factors were beyond the control of the researchers on the program evaluation team.

The DTRA consists of two parts. In part I, each student was individually tested for his/her knowledge of engineering design through a 15-minute written test consisting of three open-ended questions. The three questions were on: engineering design process (henceforth referred to as subscale 1A), effective teamwork (subscale 1B), and effective communication (subscale 1C). Part II consisted of questions to test how students applied knowledge about team design in a simulated design assignment. It was administered in teams, with a time limit of 45 minutes. The four questions on part II were: team organization (henceforth subscale 2A), customer expectations of the tool (subscale 2B), resources required (subscale 2C), and design process (subscale 2D). A standard rubric also derived from TIDEE DTRA I& II was used to score responses to the open-ended questions. In addition, a content analysis was carried out to test the hypothesis that the standardized rubric might have missed nuances in the results.

III. Results

The experimental group had 14 individuals distributed over 5 teams. The Control group had 43 individuals distributed over 7 teams. Performances on part I are compared in Tables 1 and 2.

IIIa. Scoring using the standard rubric, Results of part-I:

The results presented in this section are based on scores obtained by using the standardized rubric.

Table 1: Pre and post-test scores for students in the experimental group (Part I)

Subscales	1A	1B	1C	Total
Pre-test Mean (N=14)	4.64	3.36	2.79	10.79
Post-test Mean (N=14)	4.5	3.21	3.57	11.29

Table 2: Pre and post-test scores for students in the comparison group (Part I)

Subscales	A	B	C	Total
Pre-test Mean (N=43)	4.53	2.84	3.51	10.88
Post-test Mean (N=43)	3.98	3.16	3.53	10.67

For students in the experimental group, there was no statistically significant difference between pre and post-test means on subscales 1A, 1B, 1C or on the Total score. For students in the control group, the only statistically significant change was a decrease in group mean on the subscale that accounted for understanding of 'Engineering Design Process' (1A), between pre and post testing. Analysis was done using both non-parametric repeated measures test (Wilcoxon) and parametric repeated measures test (paired samples t-test). Significance was assumed at 0.05 or lower. (See tables 1 and 2)

IIIb. Scoring using the standard rubric, Results of part-II:

Score on part II of the instrument are shown in Table 3 for the experimental group and in Table 4 for the comparison group. Compared to the experimental teams, the comparison group teams, on an average, scored higher in the post-test.

Table 3: Pre and post-test scores for students in the experimental group (Part II)

	Team-1	Team-2	Team-3	Team-4	Total
Pre Test 2A	1	2	2	3	8
Pre Test 2B	4	1	4	3	12
Pre Test 2C	5	4	5	3	17
Pre Test 2D	6	7	5	5	23
Pre Test Total	16	14	16	14	60
Post Test 2A	2	2	2	3	9
Post Test 2B	4	1	1	1	7
Post Test 2C	5	5	5	4	19
Post Test 2D	6	4	4	2	16
Post Test Total	17	12	12	10	51

Table 4: Pre and post-test scores for students in the comparison group (Part II)

	Team 6	Team 7	Team 8	Team 9	Team 10	Team 11	Team12	Total
Post test 2A	3	4	4	3	4	1	4	23
Post test 2B	2	4	4	4	2	4	3	23
Post test 2C	3	5	5	4	4	5	3	29
Post test 2D	7	4	4	5	3	3	6	32
Post test Total	15	17	17	16	13	13	16	107

Two of the four experimental teams suffered a statistically significant (significance was assumed at 0.05 or lower) decrease in their pre to post-test scores. These somewhat perplexing results prompted a deeper look at the scoring rubric and a content analysis was carried out to see if additional information could be ferreted out.

IIIc. Scoring using content analysis, results of part-II

The results obtained from content analysis for question 2A (Team Organization and Member Responsibilities) showed that in the post-test, all teams in the experimental group had assigned a person to document the design process of the test's hypothetical design project and 75% of the teams had a person assigned to take care of financial aspects. In contrast, 71% of the control group teams assigned a person for documentation and only 28% of the teams had assigned a person to take care of financial aspects of the test's hypothetical design project. The rubric awarded points for each listing of a set of unique points about team organization.

On question 2B (Customer Expectations and Their Explanation), the content analysis revealed that students were more concerned with the 'design requirements statement' on the post-test surveys as opposed to the pre-test surveys. This was not revealed in the test scores, because the standardized rubric did not award credit for repeating points which were noted in the 'design requirements statement' as part of the test question.

On question 2C (Resources Required), the standard rubric and the content analysis were comparable in performance. In this case, the content analysis did not pick up additional student responses over and above those that were part of the rubric.

On question 2D (Design Procedure), content analysis of responses by all teams in the experimental group either confirmed the knowledge and performance level indicated by the post-test rubric's scores or revealed that responses were richer and more sophisticated than the post-test rubric scores had indicated. Moreover, content analysis revealed that different teams chose to highlight different issues based on their interests. It was also interesting that for the overall score in part-II, the same team, in the experimental group, was rated best by both content analysis and by the test's rubric-scoring. This group was characterized by certain interesting traits that merit future investigation. It was a mixed gender group (2 boys and 2 girls). It had the lowest mean GPA among all groups both at the beginning of Fall 04 and end of Spring 05. It also had the highest relative difference in GPA between two members in same group.

IIId. Investigation on the relation between scores in part-I and part-II

It is reasonable to suppose that correlations may exist between the scores of part I and part II. One purpose for examining correlations between parts I and II was to check if the students' knowledge about team design contributed to their applied knowledge and performance as part of a team in a simulated environment. The other purpose was to check the validity of test questions which made up the subscales. Even within the same section of the test, checking for question response correlations is useful for gauging the independence or inter-relatedness of different domains of knowledge and performance which are frequently addressed as separate areas in instruction and assessment. (The discussion below focuses mainly upon the correlations found, and only some findings of non-correlation have been explicitly discussed.)

First, the inter-subscale relationships among questions were examined. In the case of teams in the experimental group, no correlations were found between any of the subscales on part I and part II. For the control group, two correlations were observed. First, the post-test results showed a positive correlation between 1A (Engineering Design Process) and 2D (Design Procedure). This is reasonable: students with good individual understanding of design demonstrating good understanding of design as a team. Second, a negative correlation was also observed between 1A (Engineering Design Process) and 2C (Resources Required). Since this observation was made on the control group, it maybe that students as a team have not understood the management challenges inherent in any large design project despite of good understanding of the design process as individuals.

Next, the intra-subscale relationships among questions were examined. Within Part I, among subscales, the experimental group scores were not correlated. However, in the control group, in the pre-test, 1B (Effective Teamwork) was correlated with 1C (Effective Communication). Also, both pre and post-test, subscale 1A (Engineering Design Process) and 1C (Effective Communication) were correlated. These relationships are intuitively understandable, and, if nothing else, serve to illustrate that this assessment tool was not rigorous in separating the measurement of these areas of competence.

Among subscales of part-II, in post-test results for the experimental group, 2B (Customer Expectations and Their Explanation) and 2D (Design Procedure) are positively correlated, contrasting with the post-test results for the control group, in which a negative correlation was observed between the same two subscales. This may be reasonable, as the students in the experimental group may have gained a heightened appreciation of customer requirements in the special project management and finance curriculum of their design course.

It is usually a point of interest for educators to examine correlations between overall student GPR, a conventionally used measure of academic ability, and various specific measures of student knowledge and skill. In this study, correlations were sought between the averaged cumulative GPRs of the teams and the scores on the various subscales. The resulting correlations were not illuminating and have been merely stated for the sake of completeness. For teams in the experimental group, on the pre-test, subscale 1B (effective teamwork) was negatively correlated to cumulative GPR, but GPR correlated positively with post-test subscale 2A (Team Organization).Teams in the control group showed a positive correlation between 1B (Effective Teamwork) and cumulative GPR.

IV. Conclusions

The results suggest that the standard rubric needs to be modified to better judge the improvements attained over the course of the two semesters of design education at Texas A&M University. It also appears, per the content analysis, that the students in the experimental group were more aware of business aspects and documentation needs .One interesting observation made in the course of the study was the team that performed best was a mixed gender group with the largest relative GPA difference between two members. Although this data is too thin to offer any conclusions, it certainly piques curiosity sufficiently to warrant further investigation.

On the basis of this study, one could recommend upgrading part I of the test to a level of complexity and sophistication consistent with the maturity of the students. More time could also be allocated to completing Part I in order to encourage more complete and richer responses. Part II might be improved by video-taping team activity. Also, scoring with a combination of a standards-based rubric and a content or observations analysis seems the best way to assess the design preparedness of the students. In addition, more questions could be included to reveal the group dynamics. A reflective interview with all the students may also be an appropriate tool to consider for capturing the attitudinal and conceptual learning, and particularly the perceived change or growth that has occurred in the students over two semesters of learning. Some of these ideas are already being implemented in the 2005-06 academic year.

Acknowledgement

1) The authors wish to thank the Boeing Company for its financial support of the project and its evaluation, upon which this study was based. Results and conclusions expressed in this article are those of the author(s) and do not necessarily reflect the views of the Boeing Corporation.
2) Assessment instruments used were significantly adapted from TIDEE Design assessment tools. The authors are indebted to the inspiration and models provided by Denny C. Davis, TIDEE project director, Washington State University, PO Box 646120, Pullman, WA 99164-6120. Phone: 509-335-7993.Email: davis@wsu.edu
(For those who are interested in the actual assessment tools used in this study please contact authors.)

Bibliography

1. Davis, D., Gentili, K.L., Trevisan, M., Calkins, D.(2002). Engineering Design Assessment Processes and Scoring Scales for Program Imporvement and Accountability. *Journal of Engineering Education91(2),*211-221.
2. Davis, D., Trevisan, M., McKenzie, L., Beyerlein, S., Daniels, P., Rutar, T., Thompson, P., and Gentili, K. (2002). Practices for quality implementation of the TIDEE Design Team Readiness Assessment. *Proceedings of the 2002 American Society for Engineering Educational Annual Conference & Exposition*,
3. Davis,D., Trevisan, M., McKenzie, L,, Beyerlein, S. (2001). *Proceedings of the 2001 American Society for Engineering Educational Annual Conference & Exposition*
4.TIDEE. (2002). Transferable Integrated Design Engineering Education. www.tidee.cea.wsu.edu.
5. Dutson, A. J., Todd, R. H., Magleby, S. P., and Sorensen, C. D. (1997). A Review of Literature on Teaching Engineering Design Through Project-Oriented Capstone Courses. *Journal of Engineering Education, 86(1),* 17–28.
6. Pike, M. (2000). Capstone Design Courses: A Comparison of Course Formats. *Proceedings, ASEE Annual Conference & Exposition*, Retrieved from http://www.asee.org/acPapers/code/getPaper.cfm?paperID=2683, 13 January 2006.
7. Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., and Leifer, L. J. (2005). Engineering Design Thinking, Teaching, and Learning. *Journal of Engineering Education, 94(1),* 103–120.
8. Davis, K. C. (2004). Assessment Opportunities in a Capstone Design Course. *Proceedings, ASEE Annual Conference & Exposition*, Retrieved from http://www.asee.org/acPapers/code/getPaper.cfm?paperID=7685, 13 January 2006.

9. Ohland, M. W., Frillman, S. A., Zhang, G., and Miller, III, T. K. (2004). NC State's Engineering Entrepreneurs Program in the Context of US Entrepreneurship Programs. *Proceedings, NCIIA Annual Meeting,* Retrieved from http://www.nciia.org/conf_04/proceedings_04/htmldocs/papers/ohland.pdf, 13 January 2006.
10. Bruhn, R. E., and Camp, J. (2004). Capstone Course Creates Useful Business Products and Corporate-Ready Students. *Inroads – The SIGCSE Bulletin, 30(2),* 87–92.
11. McKenzie, L. J., Trevisan, M. S., Davis, D. C., Beyerlein, S. W. (2004). Capstone Design Courses and Assessment: A National Study. *Proceedings, ASEE Annual Conference & Exposition*, Retrieved from http://www.asee.org/acPapers/code/getPaper.cfm?paperID=7609, 13 January 2006.
12. Pavelich, M. J., and Moore, W.S. (1996). Measuring the Effect of Experiential Education Using the Perry Model. *Journal of Engineering Education, 85(2),* 287–292.
13. Adams, R. S., Turns, J., and Atman, C. J. (2003). Educating Effective Engineering Designers: The Role of Reflective Practice," *Design Studies, 24(3),* 275–294.