Assessing A Business Course - Using an Online Course Management System to Prepare for AACSB Accreditation

Roy R. Twaddle

2008/2009 No. 2
Assessing a Business Course

Using an Online Course Management System to Prepare for AACSB Accreditation

Roy R. Twaddle
6/30/2008

Dr. Twaddle is experienced in curricula management, development, and assessment. He directed the assessments of 135 high-risk/life-threatening military curricula taught worldwide and provided oversight to assessment and instruction at several DOD organizations in the U. S, including the U. S. Naval Academy. His interests are in ethical assessments of the effects of technology on strategic development and business behavior.

I am grateful to Kathleen Ferris-Costa, Steven Wesley McCann, and Ralph Perfetto for assisting in the refinement, delivery, and AACSB assessment of this course.

ABSTRACT

This paper explores how a Web-based course management system effectively and efficiently leveraged our abilities to assess and deliver a business course to students. The context is preparing for AACSB accreditation by applying technology to harvest and process assessment data. Observations, discoveries, and recommendations further our understanding of the flexibility of these powerful technologies.
INTRODUCTION

Web-based Course Management Systems (CMS) have achieved functional levels substantially beyond integrated course design, online delivery, and automated student testing. During the Fall/Spring semesters of 2007-08 we piloted and corroborated the preliminary outcomes of a business course that simultaneously addressed mandatory AACSB assessments, improved student performance, trained graduate assistants, reduced demands on limited resources, and established a tentatively effective teaching and assessment platform for future semesters. We approached this study with caution owing to the sheer complexity of a project that at first seemed simplistic. The known variables turned out to be so numerous and illusively dependent that results, conclusions, and areas for further study should only be considered as modest contributions toward supporting good judgment.

THE COURSE

Marketing Principles (BUS 365) is a lower-level, enhanced, undergraduate core course offered in five or six sections by the College of Business Administration at the University of Rhode Island. Figure 1, the course timeline for four sections during Spring-08, reveals essential features. The development concept, based on motivation theory
applied to individual students, student teams, and faculty team-teaching, was proposed in July, 2007, as the college prepared for accreditation. The course is geared to repetitiously reinforce and measure key marketing management terms, concepts, processes and applications by means of random “pop” assignments, eight quizzes, “marketing plan week,” marketing plan projects, and a high-stakes final exam. Incorporated into the design were extra credit opportunities, visits by entrepreneurs, show-and-tell student demonstrations, and a variety of online homework exercises.

A senior lecturer in entrepreneurial management & marketing developed the course; and he and an experienced Ph.D. graduate assistant (GA) introduced the pilot edition to three sections via WebCT 4.1 in September, 2007. The same lecturer and two recently trained Ph.D. GAs who were not experienced in lecturing and CMS, delivered the updated course to four sections via the same venue during Spring-08. Formal training included weekly face-to-face seminars, Web-based demonstrations, and videoconferencing. Agendas, agreements, minutes, and related documentation were published online as a readily accessible, interactive record (i.e., a wiki) to remind participants of what we planned to do and what we did.

As an enhanced (i.e., hybrid) course, BUS 365 was intended for classroom and asynchronous delivery to exploit the flexibility of Web-based technology and the creative closeness of the traditional classroom. Using the content module and cross-listing features of WebCT, the course was also uniquely configured for multisectional delivery so that each semester a team of instructors simultaneously used the same CMS cartridge (i.e., ePack). This made collaboration easier and cooperation essential. The configuration required an auxiliary server to support frame-based software incompatible
with WebCT display protocols.\(^1\) The original cartridge was a product of South-Western Cengage Learning called *WebTutor*, packaged with Lamb, Hair, and McDaniel’s *MKTG 2007/2008*.

**AACSB ASSESSMENT**

BUS 365 was among seven courses assessed during Fall-07 for the first time using AACSB standards. Twelve courses were assessed during Spring-08. Our assignment was to harvest and submit individualized, quantitative, multiple-choice test data – not written or oral presentation data – to the college’s management team for the purpose of revealing college-wide areas (e.g., student advisement, course prerequisites) to improve student performance at the individual and aggregate levels. As teaching practitioners and curriculum developers, we focused on GA training and effectiveness, resource efficiency, and course design directly related to student outcomes. Figure 2 summarizes by semester the results of assessment instruments used to accomplish these complementary goals. Optimally, we needed two semesters of data using similar

\(^{1}\) For example, Microsoft’s Visio 2007 (used to develop Figure 1) provides hypertext links that are nonfunctional if the targets reside in the WebCT file structure.
course configurations. Four learning goals (i.e., dimensions) were identified by the marketing faculty, and ten multiple-choice questions per goal were selected from the *MKTG 2007/2008* test bank and reviewed by the teaching team. These forty questions were the candidates for data collection. After a validation process conducted during the pilot semester (including construction and implementation of twenty surrogate questions to protect the originals), five questions per goal = twenty questions were accepted as measures of student knowledge. Quiz questions were not validated; but all questions on the final exam were. The instruments for delivery of these measures to students were restricted to the pre-test and the final exam (post-test embedded). Assessment data for both semesters have been harvested and are under management review. During Fall-07, 98% of 128 students met expectations and 58% exceeded expectations in all four assessed dimensions. During Spring-08, 99% of 146 students met expectations and 68% exceeded expectations in all four assessed dimensions. The following sections of this study provide the bases for developing our responses to the management team and honing good judgment and pragmatic decision making.

**PRE-TEST/POST-TEST RESULTS**

This technique answers the question, “Did students learn anything?” Results in Figure 3 indicate that they did. The process is non-intrusive, easy to do, and is a reasonable

---

2 A marketing proficiency exam to screen transfer students was also derived from these same questions.

3 Test question validation uses a statistical range of acceptance applied to actual trials. The publisher’s test bank validated at about 65% using our standard, the reason we selected ten v. five questions.

4 If a student scored less than 3 (scaled 1-to-5) in any dimension, i.e., Strategy, Consumer Behavior, Segmentation, and Marketing Mix), expectations were not met. If all four scores for a student were each greater than 3, expectations were exceeded. This standard focused on distributions rather than sums of scores. For example, a student might have a summed score = 17 of 20 = exceeds expectations, but the distribution of scores is 5, 5, 5, 2 = below expectations. For course improvement, we opted for the more stringent standard.
indicator if something is grossly wrong. Clearly, pre/post-testing was critical during the Fall-07 pilot phase. The pre-test may also indicate that some students are not proficient with CMS testing procedures, while acquainting all students with online timed tests, syntax, and question types used during the semester.\(^5\) The pre-test is not part of a student’s score for the course and, consequently, a non-motivator to do well. These variations and characteristics contribute to why pre-test scores were depressed relative to post-test scores. By post-test time, i.e., exam day, students have learned pedagogical material, as well as test-taking skills secondarily associated with the pedagogical effort. The proctored pre-test, composed of surrogate questions in the spring semester, was administered at the start of each semester prior to classroom pedagogy; although students may have remembered some course content by previously browsing the textbook, online learning modules, or by conferring with students who had completed the pilot version. In both semesters the proctored post-test was composed of the original twenty assessment questions embedded in the final

\(^5\) All testing employed one-at-a-time question delivery where no question could be revisited.
exam, the same instruments used to measure student learning in each of the four AACSB assessment dimensions. The proctored, cumulative final exam was administered at the end of each semester in a computer lab.

In both semesters the pedagogical technique appears to have facilitated nearly 100% improvement (i.e., Fall-07: 43% improved to 85%; and Spring-08: 46% improved to 89%), indicating nothing grossly wrong with the overall process and design of the course. Modest improvements in delivering the pre-test contributed to the 4 percentage point differential between semesters. Substantial design improvements probably account for the same (coincidentally) 4 percentage point differential for the post-test. It is also possible that assessment questions may have been compromised, for example, because of use during the fall semester. Although proctored in a well-monitored computer lab, questions could have been remembered and passed along. Technology-based means (e.g., covert cell phone texting) cannot be dismissed. Best practices suggest that another set of assessments should be in place for Fall-08.

In Figure 4 we compare post-test and final exam results to provide an indication whether pedagogy favored assessment outcomes. In other words, did instructors teach to assessment questions? According to the data, probably not. However, the differential in the fall semester was 4 percentage points (85% to 89%). In the spring it was less than a point (89% to 89.7%). Post-test results never exceeded exam results. In fact the percentage point differential (90% - 87%) was significant at 3%. But if they do, more objective methods to administer AACSB assessments are suggested. Another

---

6 For example, computer lab limitations in the fall were alleviated during the spring semester.
7 A new course element named “Marketing Plan Week” reinforced learning objectives across all four assessment dimensions, and new instructors were more familiar with active content such as streaming video and audio-enhanced, online PowerPoint presentations.
consideration is that instructors may be more motivated to teach to the entire final exam owing to the imminent probability – supported by C-minus quiz results – that failure rates could be substantial.

**QUIZ AND FINAL EXAM RESULTS**

While random or scheduled written assignments and oral presentations were reinforcing course content through practical applications, eight timed quizzes (20-min/20-questions) served as proximal goals enroute to a high-stakes, distal goal, the proctored, cumulative final exam (40-min/40-questions). The first four quizzes in Figure 5, Quiz 1 thru Quiz 4, and the last, Quiz 8, were proctored and worth 3 or 4-pts each. The remaining three quizzes, Quiz 5 thru Quiz 7, worth 2-pts each, were not proctored and were taken from any location. The reduced point value acknowledges potential cheating. An exam score of 24-pts (60%) was required to pass the course regardless of performance in other areas. This high-stakes aspect served to motivate learning of course content, link cheating to course failure, and motivate instructors to investigate and respond to
evidence of unethical, weak or declining performance. As the course timeline illustrates at Figure 1, the first four proctored quizzes addressing eight chapters of textbook material occur in successive weeks. The intention was to encourage team interactions (e.g., study sessions, presentation planning) before or after quizzes. Proctored quiz score averages by semester were similar at 70.8% and 70.1% respectively (combined average = 70.5%). These C-minus grades probably resulted from the combined effects of non-validated test questions and a robust schedule of activities. For example, Quiz 8, addressing two more textbook chapters, was taken on Friday as some students prepared for marketing plan presentations (20% of final grade) the following Monday.\(^8\)

A motivational, 3-pt, non-proctored, extra credit opportunity was scheduled after Quiz 4 to reinforce learning goals and compensate for the score-suppression affects addressed previously. Since some students were inexperienced with remote testing, the opportunity also segued into the non-proctored quiz regime by anticipating and resolving technical impacts of remote quiz-taking locations and various computer

---

\(^8\) Preliminary indications are that scores on Student Evaluations of Teaching (SETs) were lower, perhaps owing to non-validated quiz questions and a challenging schedule of events.
configurations (e.g., broadband v. dial-up, PC v. Mac, Explorer v. Mozilla). Figure 6 shows rising scores for Quiz 5 thru Quiz 7 in both semesters. Quiz score averages by semester were comparable to proctored quizzes at 69% and 72.7% (combined average = 70.9%). Again, C-minus grades probably resulted from the same reasons associated with proctored quizzes. The small difference in combined averages, 70.9% (non-proctored) − 70.5% (proctored) = 0.4 percentage points, suggests an area for further study to see if the differential varies favorably (i.e., remains small) as remote test-taking skills and confidence levels rise. Demands on computer labs and other in-house services could be substantially reduced by doing more remote testing.

Also of interest is the consistent pattern established by the quizzing sequences in both semesters as shown in Figure 7. Trend lines suggest congruence. The low point at Quiz 5 (combined average = 58.1% = failing scores) seems counter-intuitive considering opportunities to cheat. But, at a rate of less than a minute/question, randomization and
no revisiting, little time remained to search for answers. Rather than study, students apparently cheated unsuccessfullly. It seems, however, that they adjusted to the negative reinforcement. Substantial improvement on Quizzes 6 and 7 (combined average = 77.2) suggests that students used better techniques, recruited others to take their tests, or studied. Final exam results (90%) support a conclusion that students studied.

**BEHAVIORAL BIOMETRICS**

An apparent biometric technique emerged from CMS data recorded for each student

---

9 All tests are also unique since questions are randomly presented and multiple choices are scrambled.
taking a quiz or other test. Student computer logs from WebCT contain second-by-second, mouse-click data.\textsuperscript{10} Although unintended as a surveillance methodology, these data recorded during the proctored quiz phase could be used to construct student mouse-click rhythms, i.e., uniquely individual behavioral biometrics, as conceptually represented in Figure 8. Recent studies address mouse-clicks as a sub-set of more complex behavioral mouse-movements used to develop these biometrics. Even so, simple mouse-click rhythms may be suitable for comparisons to quiz-by-quiz rhythms recorded during the non-proctored quiz phase.

During Spring-08 five incidences related to cheating were reported. Three involved students logging in from remote locations and taking the proctored Quiz 8. Security features resident in the CMS detected the breech, and the quizzes were monitored and analyzed. Formal reporting of the incidents was initiated, but not consummated. Students stated that they did not realize Quiz 8 was scheduled in the lab. Examination of CMS data corroborated their assertions. It is reasonable that behavioral conditioning played a role owing to the three consecutive, remotely taken quizzes preceding Quiz 8. A similar incident had previously occurred – several students showed up in the lab to take the non-proctored extra credit opportunity the week after the proctored, Quiz 1-thru-4 sequence. Course policy published in the syllabus and in each online test question warns students that online quizzes are closed-book and “are no less regarded in the context of academic honesty, integrity, and professionalism than proctored quizzes and the proctored final exam.” We believe that there is an ethical imperative to inform students about this biometric and its potential associations to online cheating.

\textsuperscript{10} For example: quiz start, save question, next question, redisplay, and finish.
AREAS FOR FURTHER STUDY

For reasons probably associated with recent state-wide budget cuts, the team-teaching aspects of this program seem to have been suspended for Fall-08 or discontinued altogether. However, the following areas offer opportunities for further study and consideration.

1. **Longitudinal study to verify trends.**
   a. Will the overwhelming majority of our students continue to meet or exceed AACSB standards? Initial indications are favorable. Even if the program is abandoned, their academic progress can still be analyzed by means of AACSB assessments relative to cohorts. Future outcomes of these comparative assessments may support resurrection of the program.
   b. Will post-test results eventually exceed exam results; and if so, what are the most feasible alternatives to objectively oversee AACSB assessment instruments? Ancillary to this question is the final exam issue. Teaching-to-the-test can be a temptation hard to resist, especially where high profile assessment results are tied to consequences or high failure rates loom. We may need the services of an Assessment Keeper.
   c. Will differentials between combined averages for proctored v. remote quizzes remain small over time? The impact on in-house resources and other economic implications, e.g., commuting to school v. remote test-taking or other distance learning processes, are substantial.
   d. Will SET scores continue below historical instructor norms? If suppressed SET scores are expressions of student dissatisfaction, this could motivate
students to avoid these particular sections of BUS 365. More broadly, if
the same phenomenon occurs among the other business courses
undergoing assessment, enrollment may decline across the board.

2. **Specific action or methods to address pending problems.**

   a. Can validated assessment instruments (i.e., the twenty questions) be
      produced in time to accommodate annual revisions or can the frequency
      be decreased? Ancillary to this question is the issue of quiz-question
      validation. Validation in general is labor-intensive and time-consuming.
      Eliminating the effort is inviting, but, consequences of inaccurate student
      assessment can be substantial as illustrated by the SET issue. A
      compromise for the first question may be to develop a statistical process
      of interim evaluation to reduce the rate of semester assessments.

   b. Did course design actually produce team interaction on the eight quiz
      days? Students may have used the extra time to study or perhaps they
      merely took a break before or after each quiz. In either case the answer to
      this question should be geared toward the most effective use of
      instructional time – eight quiz days translate to about four hours.

   c. Is the SET instrument valid in the context of AACSB assessments? The
      current instrument has been used for decades. On its face the instrument
      seems inappropriate for student evaluation of enhanced courses like BUS
      365. Similarly, AACSB assessments introduce considerations and new
      variables not accommodated by the SETs.
d. Can behavioral biometrics reliably mitigate the affects of cheating, whether as a counseling asset or a detection tool? Cheating in an unproctored environment compromises pedagogical outcomes and may undermine confidence in the technique and the technology. Creating a virtually proctored assessment environment may seem pie-in-the-sky, but it is an important area for further study and potential discovery.

3. **Ethical assessments associated with technological imperatives.**

   a. Is it ethical to deliberately employ competing demands to motivate students and instructors during the conduct of a business course? Because course design probably suppresses quiz scores (beyond the validation issue), student tension and motivation to succeed on the high-stakes final exam is high. Instructor anxiety is also high regarding the perceived likelihood that a comparatively high number of students may fail the course. In fact, GAs were inclined to make unempirical curriculum adjustments as the course proceeded (especially toward semester’s end), which reduced incentives for students to exceed expectations. The philosophical origens of appropriate ethical standards may stem from values associated with free market economic systems.

   b. Is it ethical to relax standards of student accountability for policy violations associated with cheating during the conduct of a business course? Behavioral conditioning, apparently coupled with quiz-taking regimes and transitions, suggest that course design contributed to student violations of important course policies. The philosophical origens of appropriate
ethical standards may stem from value systems associated with individual and organizational integrity, truth-telling, and responsibility.

c. Is it ethical to employ surveillance biometrics in a business course? If, as we believe, there is an ethical imperative to inform students about the use of biometrics, an ethical assessment regarding university-wide policy should reveal important ramifications of this emerging technology. The philosophical origins of appropriate ethical standards would likely stem from human rights associated with freedom of choice, privacy, and self-determination.

REFERENCES

None. I will include a complete reference list.
Our responsibility is to provide strong academic programs that instill excellence, confidence and strong leadership skills in our graduates. Our aim is to (1) promote critical and independent thinking, (2) foster personal responsibility and (3) develop students whose performance and commitment mark them as leaders contributing to the business community and society. The College will serve as a center for business scholarship, creative research and outreach activities to the citizens and institutions of the State of Rhode Island as well as the regional, national and international communities.

The creation of this working paper series has been funded by an endowment established by William A. Orme, URI College of Business Administration, Class of 1949 and former head of the General Electric Foundation. This working paper series is intended to permit faculty members to obtain feedback on research activities before the research is submitted to academic and professional journals and professional associations for presentations.

An award is presented annually for the most outstanding paper submitted.

Mission

Our responsibility is to provide strong academic programs that instill excellence, confidence and strong leadership skills in our graduates. Our aim is to (1) promote critical and independent thinking, (2) foster personal responsibility and (3) develop students whose performance and commitment mark them as leaders contributing to the business community and society. The College will serve as a center for business scholarship, creative research and outreach activities to the citizens and institutions of the State of Rhode Island as well as the regional, national and international communities.

The University of Rhode Island started to offer undergraduate business administration courses in 1923. In 1962, the MBA program was introduced and the PhD program began in the mid 1980s. The College of Business Administration is accredited by The AACSB International - The Association to Advance Collegiate Schools of Business in 1969. The College of Business enrolls over 1400 undergraduate students and more than 300 graduate students.

Founded in 1892, the University of Rhode Island is one of eight land, urban, and sea grant universities in the United States. The 1,200-acre rural campus is less than ten miles from Narragansett Bay and highlights its traditions of natural resource, marine and urban related research. There are over 14,000 undergraduate and graduate students enrolled in seven degree-granting colleges representing 48 states and the District of Columbia. More than 500 international students represent 59 different countries. Eighteen percent of the freshman class graduated in the top ten percent of their high school classes. The teaching and research faculty numbers over 600 and the University offers 101 undergraduate programs and 86 advanced degree programs. URI students have received Rhodes, Fulbright, Truman, Goldwater, and Udall scholarships. There are over 80,000 active alumnae.