A Comprehensive System for the Evaluation of Innovative Online Instruction at a Research University: Foundations, Components and Effectiveness

Jeffrey D. Kromrey (E-mail: kromrey@tempest.coedu.usf.edu)
Kristine Y. Hogarty
Melinda R. Hess
Gianna Rendina-Gobioff
Amy Hilbelink
Tom R. Lang

College of Education
University of South Florida

Abstract

The delivery of post-secondary coursework via the Internet continues to gain momentum. As a result, investigations into effective and appropriate methods of evaluating the effectiveness of these courses are required. In an effort to meet this challenge, this study describes the development and implementation of an evaluation system applied to new online programs at a major research university. A systematic approach to evaluation provided formative feedback on the processes and products of course development using diverse data sources including course documents, interviews and web-based surveys. Results of both quantitative and qualitative analyses support the integrity of the evaluation system and provide preliminary indications of course effectiveness based on student satisfaction.

Introduction and Purpose

Online courses are proliferating at post-secondary institutions at an ever-increasing pace, bringing new and unique challenges to course assessment and evaluation. The National Center for Education Statistics reported “in 2000–2001, 90 percent of public 2-year and 89 percent of public 4-year institutions offered distance education courses” (Tabs, 2003, p. iii), a marked increase from the previous report (78% for 4-year and 62% for 2-year institutions). As online courses flourish, it is critical to conduct meaningful assessments. As part of a series of studies designed to facilitate and inform a five-year technology project, the primary purpose of this study was to develop and validate an evaluation system to examine the development, delivery, and effectiveness of web-based courses.

Rationale and Background

Large scale evaluation planning and implementation should be grounded in formal models of program evaluation. A consideration of the variety of standards, theories and models for evaluation, within the context and nature of the specific program to be evaluated often suggests an amalgam model or framework that will provide
appropriate direction for evaluation planning. Evaluation standards include utility, feasibility, propriety and accuracy (Joint Committee Standards for Educational Program Evaluation, 1994). These standards focus evaluators on critical criteria for the design, conduct and dissemination of evaluation efforts. Although a variety of formal evaluation models have been proposed, most such models trace their roots to Stufflebeam’s CIPP model, providing a framework for the evaluation of a program’s context, input, processes and products (Stufflebeam, 2002). Similarly, instructional design models such as the systems-based Dick and Carey model (Dick, Carey & Carey, 2004) focus evaluation efforts on applications of learning theories and principles that guide the design and development of instruction.

The instrumentation and methods used in this study were developed with special attention to Eaton’s (2002) recommendations to accreditation agencies for evaluating distance learning. These strategies, entitled Assuring Quality in Distance-learning, were gleaned from a publication produced by the Council for Higher Education Accreditation (Phipps, Wellman, & Merisotis, 1998). In addition to Eaton’s (2002) recommendations, the recently adopted standards of the International Society for Technology in Education (ISTE, 2002) were consulted, with careful attention to essential conditions required to create learning environments that are conducive to powerful uses of technology. Although the aforementioned standards were developed specifically for teacher preparation programs, a subset of the essential conditions was deemed appropriate during this stage of our research. For example, the essential conditions of an effective system that coheres around a shared vision for the integration of technology and its use to support new modes of teaching and learning are critical contextual factors in the project. Of course, educators must be skilled in the use of technology for learning, and must have consistent access to professional development in the support of technology use in teaching and learning.

Research Methods

The courses examined in this study were initial components of a comprehensive series of online courses that are anticipated to fulfill the requirements of master’s degree programs in three distinct disciplines. Data were gathered over three semesters from instructors and students in 16 courses. Four instructional designers provided technical support for the instructors during course development and initial delivery.

The process undertaken reflects a triangulated approach to evaluation, consisting of multiple methods of gathering data to both develop the evaluation system and begin the process of course evaluation. Initially, our 11-member research team, comprised of measurement, technology, and evaluation specialists, looked to the literature to glean information regarding accreditation standards, the use of technology in higher education, and effective course evaluation (see, for example, ISTE, 2002; NCATE Standards, 2000; Barron, 1998; Phipps, Wellman, & Merisotis, 1998; Moore & Kearsley, 1996). Next, we divided into work groups to develop a variety of instruments and processes to gather data on the development and delivery of online courses. Once instruments were developed, they underwent initial validation procedures, including expert reviews and pilot testing. Information gleaned from these steps guided further refinements.

Analysis methods included both qualitative and quantitative aspects. Data were gathered from students, faculty, and instructional designers. Students were surveyed at three points during the semester, with each survey tapping a different aspect of the student’s experience with the online course. Interviews were conducted with Faculty and instructional designers at the beginning of the semester to gather data on the issues and processes they faced developing and delivering these courses. At the end of the semester, faculty members also completed a survey designed to gather information about their perceptions of course development and delivery. Open ended responses were thematically analyzed in order to identify promising practices for effective development and delivery of distance-learning courses.

In addition to interview and survey data, team members analyzed pertinent course documents and artifacts (e.g., course syllabi, instructional design plans and online course content). These data sources provided evidence of pedagogical and technological innovations, and served to verify the consistency between the planned and delivered courses.

Surveys and Interviews

Three student surveys were administered, each designed to tap a different aspect of the student’s experience with the online course. Faculty and instructional designers were also interviewed at the beginning of the semester to gather data on the issues and processes they faced developing and delivering these courses. Toward the end of the semester, faculty members completed a survey consisting of both selected and open response items related to
perceptions of course effectiveness. The collective responses were analyzed for themes in order to identify recommended practices for effective development and delivery of distance-learning courses.

The three student surveys and the faculty survey were developed for administration at various and distinct stages of course delivery to gain a holistic and in-depth evaluation of the online courses. This process resulted in a three-phase approach to gathering data from students enrolled in online courses. Initially, five domains were identified based on our review of the literature and then further collapsed into three categories after extensive review and mapping of items to domains by the research team (see Smith, Helmick, Kromrey, Carey, Lang, Ducher, et al., 2003 for further details). The three resulting domains were identified as: (1) Course Delivery, (2) System, and (3) Design. Additionally, a fourth domain was added for the purposes of overall evaluation, (4) Satisfaction. Once an initial version of each instrument was constructed, the entire team reviewed each for purpose, consistency of focus, wording and uniqueness. The instruments were then refined and finally sent for review by the instructors of six online courses for content, layout, comprehensiveness, applicability, etc. Their input guided further refinement and adjustments to the instruments. For example, one of the instructors suggested adding an item asking students about the source and processes needed to acquire materials necessary for the course (e.g., did they buy books in person at the school bookstore or online?). Table 1 contains a summary of the instruments that were developed as a result of this process.

The faculty survey, administered at the end of the semester, gathers data intended to inform about the preparation, delivery and effectiveness of their courses in an online environment. This instrument consisted of both Likert response items as well as open ended items to obtain information in both a structured and flexible manner. Initial validation efforts were taken through cyclic reviews by both the team and the faculty involved in the online courses.

Table 1
Description of Surveys for Online Courses.

<table>
<thead>
<tr>
<th>Survey</th>
<th>Purpose of the Survey</th>
<th>Expected Administration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Survey One</td>
<td>To gather data on the reasons students choose to take an online course</td>
<td>Within the first two weeks of the semester</td>
</tr>
<tr>
<td>Student Survey Two</td>
<td>To gather data on students’ perceptions of the usability and effectiveness of elements of the online course (e.g., navigation, online resources, etc.)</td>
<td>At approximately the mid-semester point</td>
</tr>
<tr>
<td>Student Survey Three</td>
<td>To gather data regarding student satisfaction with the course. This survey is somewhat analogous to the typical end-of-course survey administered to in-class students</td>
<td>Within the last two weeks of the semester</td>
</tr>
<tr>
<td>Instructor Survey</td>
<td>To gather data regarding instructor’s satisfaction with and perception of online course delivery, including support, resources, and pedagogical effectiveness.</td>
<td>At the end of the semester of initial course delivery</td>
</tr>
</tbody>
</table>

The protocols for interviewing both faculty and instructional designers underwent similar developmental steps as the surveys, including assembly of teams to develop the format of the interviews and questions to be asked. Due to the nature of the interviews as well as the differences in their respective jobs, they took on slightly different approaches.

Faculty members were interviewed individually by pairs of interviewers from the evaluation team using a protocol consisting of 36 questions falling within six general areas: Background/Training/Experience, Commitment, Support, Innovations, Satisfaction, and Concluding Remarks. Each team member took notes and independently transcribed their notes. Subsequently, these results were independently reviewed for themes.

Unlike the interviews that were conducted individually with each faculty member, a focus group approach was employed for the instructional designers using a semi-structured protocol. This was done primarily due to the team working environment in which the four instructional designers operated. They were more collaborative in their daily work since they all worked on multiple courses as compared to faculty members who were primarily responsible for one course a semester. A team conducted the interviews, asking the instructional designers questions from the protocol, audio-recording the conversation, and taking notes. All participants were encouraged to elaborate on any issue to ensure all designers were comfortable during the interviews and that all of the information obtained was correctly interpreted by the evaluation team.
Document analysis

In addition to information obtained from those directly involved with the courses, team members reviewed pertinent course documents (e.g., course syllabi, instructional design plans and online course content). This review was intended to provide evidence related to the breadth and depth of course documentation regarding pedagogical and technological innovations, as well as the degree to which different documents synthesized information for a given course and were reflective of actual course practices. These documents each have a unique purpose and serve to inform different aspects of course development and delivery.

Course Syllabus. The examination of a collection of current course syllabi was deemed essential to our inquiry. In order to conduct a content analysis of these documents, a team of measurement and technology specialists developed an instrument to analyze individual online course syllabi. This team identified six central categories that provided the framework for the content analysis. Individual team members reviewed each syllabus independently noting the inclusion of elements in each of the following categories: Contact Information, Online Resources, Course Delivery and Interactions, Processes and Products, Course Administration, and Software/Tools. Each category in the checklist contains at least five indicators representing the domain of interest. For example, with respect to Online Resources, syllabi were reviewed for explicit references regarding specific online resources for students such as course syllabi, class notes, readings, or specific websites to view. Within each syllabus we also searched for information regarding the Course Delivery mode (e.g., satellite, blackboard, WWW, teleconferencing, etc.). We anticipated that the mode of delivery might have implications for other aspects of our inquiry. In the Processes and Products category, evidence of students’ creation or interaction with technology was specifically sought. That is, do students have to use technology to create projects/presentations? Are they using audio/video? Are asynchronous or synchronous discussions part of their coursework? Is technology used to facilitate group or team assignments or interactions? In terms of Course Administration, we were interested in whether assignments were submitted online, if grades were available online, if course goals or objectives were related to the integration of technology and if there were any technology related prerequisites. Lastly, with respect to Software/Tools, we sought information related to software, hardware, or Internet requirements.

New online course syllabi were collected each semester and analyzed three at a time by each of the team members independently. After each set of three syllabi was reviewed, the members met as a group to discuss their findings. As courses were examined the content of the instrument was reviewed and revised. At this point, we feel that collectively these six categories and the individual items within the revised instrument adequately capture the breadth and depth of evidence and provide a broad overview of the integration of technology for teaching and learning.

Instructional Design Plan (IDP). Another key element of our document analysis was the examination of Instructional Design Plans. The primary purposes of the IDPs are to document the analyses and anticipations for the course and to provide a mechanism for communicating learning objectives, design strategies, and production schedules between the instructional designer and the course instructor. The instructional designers construct the IDPs based on preliminary syllabi and on-going discussions with the course instructors. Although individual IDPs differ slightly based on the content and the designer, they all cover the Analysis, Design, Technical Specifications, and Production Schedule of the course. The Analysis portion of the IDP consists of goals for the course, learning objectives, and audience analysis. The Design portion provides information relative to the instructional strategies, course structure, and specific tools (such as bulletin boards and chat) that will be implemented into the course. The Technical Specifications section covers aspects of both the development and delivery environments. The Production Schedule outlines the tasks, responsible parties, and due dates.

Once again a team of measurement and technology specialists developed an instrument to synthesize the information contained within the IDPs. The checklist was divided into three major areas: Innovative Features, Instructional Strategies, and Course Management System Options. A three-stage process was used to analyze the content of the IDPs. In stage one, each member of the group individually reviewed the IDPs and noted items evident in each of the three aforementioned areas. During the analysis of IDPs, each member of the group was encouraged to add to the checklist as necessary. During the second stage, group discussions of the individual reviews provided a means to consolidate the findings, resulting in a modified instrument. Using the revised instrument, each item was assigned a value between 0 and 4 indicating how many group members identified that particular item in the IDP. Items that were rated either three or four were noted as identified in the IDP. Items scored two were discussed further to reach consensus, while those receiving a score of one where excluded from the final results. Finally,
scoring on the instrument was reduced to a dichotomous scale with one indicating present (items with scores of three or four) and 0 indicating not present (items with ratings of one or zero). The resulting instrument contained 47 items representing Innovative Features, Instructional Strategies and features of a Course Management System. Frequencies were examined for each of these items to establish the extent of implementation.

**Course Content.** A review of the online course material was utilized to understand how the courses use technology and innovative strategies for teaching and learning. A multi-category checklist was created to assess technical and pedagogical innovations found within the online courses. Four major domains: Course Design, Interaction and Collaboration, Assessment and Learner Support, were evident in course documents.

These elements describe how the courses are implementing innovative strategies in ways to increase interactions and effectiveness. The review of the courses in this manner provides additional information that complements the other data sources, providing a vivid picture of the overall courses and the processes that cohere to create online programs.

When evaluating Course Design, document reviewers examined elements related to technology (e.g., audio, video, animation, graphics, etc.), content format (e.g., multiple file formats, availability of CD-ROM) and content (e.g., course organization, consistency, and navigation). Within the section entitled Interaction and Collaboration, items were grouped into three categories: technology, feedback, and level of interaction. Items concerning technology, outcomes, and student presentations were listed under the larger heading of Assessment. Lastly, Learner Support included the subcategories of specified requirements, explanations of how to use technology, evaluation, and technology help and resources.

Individual team members were granted course access to online materials by the instructors delivering each course. The timing of course entry was critical as not all course materials were available for the entire length of the semester. Once again individual group members reviewed materials independently before meeting as a group to reach consensus regarding the presence or absence of the aforementioned elements.

**Document Synthesis.** Once each set of documents was reviewed independently it was considered vitally important to examine the relationship between the information provided in the IDP and the course syllabus. Concurrently, an investigation of the consistency between the content of the IDP and the actual course content and delivery was conducted. In this vein, checklists were developed to guide the systematic analysis of these documents. This part of the development of the assessment system is still in progress. Once these independent assessments are complete these two groups will merge to examine trends across the entire set of course documents.

**Results**

**Student Survey Data**

The results from the surveys of approximately 200 students suggest that nearly 60% of the respondents live more than 30 miles from the campus where the courses were offered, and 47% live more than 60 miles away. Consistent with the geographic distance, more than half of the students reported that distance was an important factor in their decision to enroll in the courses. Additional important reasons for enrollment were work schedules (69%), class schedules (63%), and family obligations (49%). Survey questions related to previous experiences suggested that the students were relative novices in learning via online distance education. For example, 28% of the group is enrolled in their first web-based course, and an additional 10% have only one previous online course experience. Despite the large proportion of novices in these courses, more than 90% of the respondents indicated that it was easy to obtain an ID card and the network ID required to access the course delivery system, to access the Internet and the server, and to obtain a course syllabus.

Student survey data gathered at course midpoints suggested that the majority of the respondents did not experience any technical problems with connecting to their courses, uploading assignments, opening documents and displaying graphics (ranging from 70% - 76%). Minor problems were experienced with broken links (36%), and major problems occurred less than 4% of the time. In response to questions regarding the utility of various aspects of online instruction, the overwhelming majority of participants were most satisfied with the online submission of assignments and the online grade book (with more than 79% reporting these aspects of the course to be Very Useful). Sixty-six percent of the participants reported that navigation was Very Easy, and more than 90% believed that taking this particular online course was a good decision. Additionally, 61% of the students reported that their assignments took about the same time to complete as those in more traditional course formats, while 32% reported that the online assignments required more time.

Data gathered at the end of each course suggested that students found the course organization to be
agreeable (81% found it logical and easy to follow, 77% reported that activities and assignments facilitated their understanding, and 73% reported that assignments were aligned with course objectives). Two items raised concerns about course delivery: 53% of the students reported that they rarely or only sometimes could skip over course content that was already familiar, and 24% reported that their instructors rarely or only sometimes provided timely feedback on assignments.

**Faculty Survey Data**

This end-of-course survey included both selected-response and open-ended items designed to obtain information on faculty perceptions related to the course development and delivery. The instructors participating in the course offerings reported being Very Satisfied on 21 of the 30 selected response items, including items addressing student performance, availability of technical support, working with the instructional designer, and satisfaction with the course as a product. Further, the instructors reported being Satisfied with aspects such as interaction and communication with students, technology dependability and support received from the department and college. The instructors reported being Dissatisfied on only two items (support received from other instructors and ability to interact with other faculty using technology in teaching).

The open-ended responses corroborated the data obtained from the selected response items (e.g., the instructors reported being very pleased with the quality of the students’ work in the course and the experiences collaborating with the instructional designers). Further, the instructors indicated improvements to be made in future offerings of these courses in terms of clarifying instructions for students and facilitating communications. Although no dissemination activities have been undertaken, the instructors reported plans for presenting information on their experiences at professional conferences in the future.

**Faculty Interviews**

The interview protocol consisted of 36 questions falling within six general areas: Background/Training/Experience, Commitment, Support, Innovations, Satisfaction, and Concluding Remarks. The faculty members’ backgrounds and previous experiences with online course delivery varied greatly. One faculty member, for example, had delivered technology-driven distance learning courses in a variety of modes (e.g., web, email, satellite) for over five years. Conversely, other faculty members indicated no previous experience delivering courses from a distance and had received minimal, if any, training beyond their own self-study. All faculty members interviewed, regardless of their previous experiences with online instruction, were enthusiastic about delivering their courses via this method.

Thematic analyses of faculty interview transcripts suggested that virtually all of the instructors indicated that the time required to design, develop, and deliver their course online was intense and, in many cases, much greater than what had been anticipated. Assistance for faculty, in the form of teaching assistants or other graduate student help, also varied, with some faculty receiving no additional support through graduate or teaching assistants to others with more than one assistant. All faculty members had the benefit of working with an instructional designer to help develop and deliver their courses under the grant, and all indicated that working with their designer(s) was a very positive experience. The degree to which technological assistance (in the form of hardware and software support and troubleshooting) was considered adequate, however, varied across instructors. Instructors also indicated varying experiences with developing a positive and community-oriented atmosphere online. Although at least three instructors felt that they had been able to foster a very strong sense of community and partnership both between themselves and their students, as well as between the students themselves, other instructors considered this an issue of potential concern.

All of the instructors were satisfied with various aspects of innovations that occurred either as a part of their online course delivery or as a result. Three instructors indicated that putting together a class online required more thought and careful consideration of the organization and presentation of course material. A few noted that they had changed methods of assessment to both take advantage of the opportunities presented in an online environment (e.g., having students make web pages as a means of introducing themselves and communicating their learning) as well as circumventing potential problems with cheating (e.g., more project based assessment processes). Three instructors also indicated that their experiences developing and delivering online courses had helped them make changes in courses they delivered face to face. Two instructors provided feedback on the different characteristics required for students to be successful in online and in-class delivery modes. One instructor was particularly adamant that both modes of delivery are needed to meet the learning styles of different students.
Instructional Designer Interviews

Transcripts from the semi-structured interviews of instructional designers were analyzed for major themes. Thematic analyses of these interviews revealed seven domains (Faculty/Designer Interaction, Monitoring Maintenance and Evaluation, Course Management System, Resources, Designer Perspective, Instructional Design and Instructional Development). These themes elucidated issues the instructional designers found both challenging and rewarding throughout the process. One of the most prominent positive results was a strong collaborative relationship built between the instructors and the designers, while the most prominent negative theme stemmed from restrictions placed on design due to the available systems for online course development and short time frames for development.

Document Analysis

Course Syllabi. The analysis of course syllabi was guided by a checklist delineating evidence of the integration of technology with respect to six categories: Contact Information, Online Resources, Course Delivery and Interactions, Processes and Products, Course Administration, and Software/Tools. These data suggested variability with respect to the availability of online resources, with some providing online readings, while others provided class notes. With the exception of a single course, all of the courses required students to create or interact with technology to produce course projects or presentations, and required some form of online submission of student work. Further, as might be expected, all of the courses required some form of asynchronous discussion or posting. Surprisingly, only one course made reference to the use of an electronic portfolio. Lastly, all of the course syllabi contained references to the technology requirements of the course (e.g., software, hardware, plug-ins, internet connections).

Instructional Design Plans. Currently, the IDPs have been examined from the standpoint of the design strategies implemented and the “innovativeness” of the strategies. Within the IDPs, we noted a variety of approaches across the courses. From this review, two major themes emerged – an emphasis on collaboration and project-based learning. In every case, the design plans stipulated that students would engage in discussions (synchronous and/or asynchronous) and work in groups to produce projects related to the course objectives. Several innovative uses of technology were also noted – many of the design plans explicitly mentioned narrated presentations, and several plans called for the implementation of video, either as an introduction, presentation of content, or student project. All of the courses were divided into several units, with related content, links, and assignments for students. Only two of the courses included activities with immediate feedback (e.g., computer-assisted instruction tutorials).

Online Course Content. The review of the online course content provided an evaluation of the delivered course materials, with a special emphasis on uses of technology and innovative strategies for teaching and learning. The analysis of these data revealed many aspects of technology innovation being exhibited within the courses. Some of the innovations that were noted include: narrated, web-based presentations from the instructor; an online syllabus that hyperlinks content within the course to outside resources; web cast; video presentations that enhance content with authentic examples; and student use of technology such as making web pages and multimedia presentations.

Discussion and Conclusions

The results of this study provide strong evidence that delivery of distance-learning courses needs to be carefully and thoughtfully planned through collaborative efforts of individuals. The initial results of the student evaluations suggested satisfactory evidence of content validity with internal consistency estimates ranging from .70 to .90 for the individual subscales within the surveys. With respect to the essential component of interaction, students indicated that communication was a key element in course effectiveness as well as the availability of resources and assistance. Findings from the examination of the psychometric integrity of the instrument will guide further refinements and enhancements to the planned system of evaluation.

As with any research or evaluation effort, issues arose, both anticipated as well as non-anticipated that limited the strength with which we are able to make our inferences. One of the biggest problems arose from the use of technology in distributing the surveys and collecting the data. The evaluation team encountered reliability
problems with the server identified to be used and, at times, had to use alternate servers to deliver the surveys. This necessitated an unusual amount of backing up the data. These issues tended to arise all three semesters but with different problems encountered. Additional problems included lack of response by students. Whether this was because instructors did not send the surveys to students or because students were reticent to respond is unclear. Steps have been taken with each subsequent semester to enhance response rates and are, at this point, showing some improvement. These issues, while frustrating, serve as valuable lessons about the frailties of technology and the problems that can be encountered as a result of our growing dependence on technology tools and capabilities.

As the proliferation of distance-learning opportunities continues in this era of educational accountability, educators, administrators, and institutions will need to have tools and methods available to ensure that the courses and programs they offer continue to meet the requirements of accreditation, policy-making, and funding agencies. This study provides a foundation upon which to build and further develop a system to meet the needs of a host of individuals and groups involved in higher learning, including faculty, administrators, and regulating or granting agencies. Furthermore, although the focus of the current study is the evaluation of post-secondary online courses, the implications for other education environment, e.g., K-12, are numerous. The push for a P-16 or K-16 ‘seamless’ education system (Ercegovac, 2003; Reising, 1996), the increasing tendency for students to graduate earlier from High School and the increased emphasis on college preparation and college-credit classes in high school (Rajala, 2003; Olszewski-Kubilius, 2002;), and the prevalence of distance learning in the K-12 system (Donlevy, 2003; Collins, 2001) support the relevance of this work beyond post-secondary education.
References


Acknowledgements

This work was supported, in part, by the University of South Florida and the Fund for the Improvement of Postsecondary Education, under Grant No. P339Z000006. The opinions expressed are those of the authors and do not reflect the views of the United States Department of Education or the University of South Florida.