The Online Academy Formative Evaluation Approach To Evaluating Online Instruction

Edward L. Meyen, Ronald J. Aust, Yvonne N. Bui, Eugene Ramp, and Sean J. Smith

The University of Kansas

e-Learning Design Lab

3061 Dole Human Development Center

The Center for Research on Learning

1122 West Campus Road

Lawrence, KS 66045

785-864-0675

meyen@ku.edu

Abstract

The Online Academy (HO29K73002) was funded by the Office of Special Education Programs (OSEP) to develop online instructional modules in the content areas of reading, positive behavior support and technology across the curriculum. Targeted to preservice teacher education programs in institutions of higher education (IHE), to date, the modules have been adopted for implementation by 162 institutions. A requirement of the funding agency was that the content of the modules be research-based. A total of 75 lessons in 22 online modules and an authoring software tool were developed. Each module is approximately equivalent to a one semester credit course. This paper describes the formative evaluation processes that were employed in creating

the instructional design, design and production processes, content development, usability and navigational features of the modules, and the national implementation process.

Key words:
Online modules
Teacher education
Interactive
Streaming media
Extranet model
Higher education evaluation
Beta testing research-based
Jurors
Content experts

Background

The Online Academy was funded in 1996 as a three-year project to develop online instructional modules for preservice teacher education programs. This represented a major development effort as few commercially available development tools were available for web-based instruction at that time, and existing ones were geared primarily for institutional use, where the fee structure was mainly passed on to the institution on the basis of student use. Additionally, streaming media was still in its infancy, with minimal use of streaming media in online instruction. Besides, the status of online instruction in higher education institutions was in an early stage, with few institutions engaged in creating or offering online instruction. Consequently, there was little experience or literature to draw upon in the development of online instruction of the magnitude included in the scope of work for the Online Academy.

The development tasks facing the Academy included creating an instructional design applicable to the proposed content and target audience, developing systematic processes and authoring software to automate the module production process and to facilitate writing content, and implementing the modules as instructional resources in preservice teacher education programs

nationally. Each development task introduced the need for a major emphasis on formative evaluation to enhance the decisions to be made relative to each of the tasks. The formative evaluation process was further complicated by (a) the commitment to scalability nationally not knowing the changes in technology that would occur over the life of the project; (b) the integration of streaming media as a central mode for presentation of lessons; and (c) the requirement that the content be research-based. The decision was made early in the project to disseminate the modules to institutions of higher education via an extranet model that would require each institution to manage the modules on their own servers. This eliminated the dependency on the Academy during implementation and ensured continued access to the modules following the end of the grant.

The development effort involved creating and field-testing 22 online modules that contained approximately 14,000 graphics, 4,600 streaming media files and over 50,000 unique files. Each module was designed to be taught in either an interactive mode with an instructor or as a self-contained unit. The modules were also designed to be Bobby compliant for access by persons with disabilities.

Formative Evaluation Procedures

The scope of the development effort, combined with the limited time period, dictated that a formative evaluation model be adopted. The challenge faced by the Online Academy in designing a formative evaluation model was not the lack of traditional evaluative strategies. Rather, it was an almost total absence of literature on evaluation of online instruction when the project was funded in 1996. This was especially true of development projects of the magnitude of the Academy. The evaluation efforts were concerned with the effectiveness of the technology

as well as the instructional design and content. The model drew heavily from the evaluation of traditional forms of distance education or curriculum development (Beyer, 1995). For example, a body of literature focused on course development and the utilization of instructional television (Price & Repman, 1995). There was also an emerging literature base on the evaluation of teaching (Nelson & Smith, 1994). Dodge (1994) offered direction on the formative evaluation of a computer-based tool in the development of lessons. In 1998 Bodily and Mitchell published a source book on the evaluation of Challenge Grants funded by the U.S. Department of Education. Subsequently, resources began to appear in the literature on the formative evaluation of web-based courses (Maslowski, Visscher, Collies, & Bloemen, 2000; Youngman, Gotcher, Vafa, Dinsmore, & Goucher, 2000). Thus, in addition to drawing heavily from the traditional literature on formative evaluation in the design of evaluation procedures, the Academy sought consultation from individuals engaged in the development of online instruction at the time for input on evaluation of technology related features of online instruction.

The formative evaluation process needed to minimize the time required to translate evaluation results into indicators for revising or refining the modules and the major elements of the project. While formative procedures were integral to all aspects of the project, five processes served as the main foci. These included:

- *Instructional design of the online modules:* As mentioned, literature on the status of online instructional design was very limited in 1996. However, online courses were being developed at the University of Kansas that were fully online, employed multimedia including streaming media and incorporated a wide array of instructional features characteristic of face-to-

face instruction (Meyen, Lian, & Tangen, 1997a). These instructional design features became an early focus of evaluation as they served as the basis for the foundation for module content development.

- *The production system*: Two early decisions set the framework for evaluation strategies related to production systems. The first was that a production system would be created to allow content to be written according to prescribed specifications to meet the requirements of the production tool. Once the content met the requirements of the specifications, it would be processed in the production system and the operational online modules created. This meant that the instructional design had to be validated prior to creation of the tool, i.e., the production system, and the tool had to be completed prior to the development of specifications for content writers.

- *Content development:* Online design delivery required that the content be in final form and in compliance with all design features prior to being delivered. (see Figure 1) This added a level of precision in the creation of content that is not essential in face-to-face instruction. In addition, the funding agency requirement that the content be research-based added a dimension of content validation more rigorous than is typical of course development. Two primary strategies were employed in the development of content. The first involved the use of content experts as jurors to identify applicable research and validate the content to be incorporated into the modules. The second was the use of writing teams knowledgeable of the content areas that worked with the jurors in the validation process and in writing the content to meet the specifications of the instructional design and production system.

- *Usability and navigation:* In contrast to face-to-face instruction where the instructor manages the students' progress through the content and has real-time options to modify the presentation, if necessary, during online instruction all instructional features have to be developed in advance and thus are not subject to "on the spot" revision. Further, online instruction requires that the instructional features be integrated into an electronic form and managed by the students as they progress through the program. Thus, it is important that the features and the process (navigation) for moving through the instructional program be consistent. They must also be somewhat intuitive so the students can easily learn how to navigate through the instruction and thereby focus their attention on the instruction and not the technology features. The user interface design adopted standards of information organization that are widely applied and understood, including a page structure, table of contents, nested menus, chapter headers and way-finding markers that indicate the current location and scope of information (e.g., page 14 of 76). The flexible navigation design accommodates the preferred styles of diverse learners, including a linear navigation from first to last page, spoke-and-wheel navigation using the table of contents and chapter menu navigation.
- *Implementation:* The goal of the project was to make the online modules available to any preservice teacher education program nationally. This requirement meant that it was essential to design the technology so that the modules would be maximally accessible regardless of the institution from which students might take the modules. It was also essential to develop the modules so that they could be installed on local servers of participating institutions within the context of the extranet model. The latter was considered essential for ensuring access to the modules following completion of the project.

A mixture of formative methodologies was employed depending on the purpose of the function or process being evaluated. Both internal and external strategies were used. The external strategies involved a third-party evaluation firm that was contracted to carry out specific evaluation tasks. Their involvement was primarily related to evaluation of the usability/navigation and implementation processes. Additionally, because the Academy project was carried out in an academic environment, student research studies were encouraged and, when appropriate, the results were utilized in the formative evaluation process.

The five evaluation targets listed above were addressed sequentially; however, each was also part of a continuous process. For example, the instructional design was the first to be addressed as the design had to be in place prior to beginning work on the development tool and production system. Likewise, the tool had to be completed before content generation could begin. The focus of each evaluation effort was on the targeted functions, but if data supported changes in other target functions, they were made. The exception was in the instructional design and the tool during the third year of the project where module production was heavy. At that time, a policy of "no feature creep" was adopted (i.e. no new features). Any change in the design required a change in the tool and in the development of all modules. Thus, even if a feature change was found to be beneficial, it was not cost effective to make it at that stage in the project. Only one design change was made in the third year of the project and that was the indexing feature made available through the Real Player. This feature simplified the task of the learner moving from one place to another in the multimedia presentation.

In the following discussion each of the five evaluation areas will be addressed in terms of methodology and results.

Instructional Design

Methodology. The instructional design for the Academy modules derived from the online instructional design developed by Meyen for graduate-level courses at the University of Kansas (Meyen et al., 1997b). Meyen had developed two online instructional models. One was designed for graduate level courses (this model received Honorable Mention in the international Paul Allen Virtual Education competition for online courses), the other approximated a graduate seminar. Each had been taught at least once under normal conditions and subjected to Alpha testing in the design process. Focus sessions had been conducted with students unfamiliar with online instruction engaged in working through lessons and evaluative data on each lesson were collected on student satisfaction and performance. A study was also carried out using the standard course evaluation system employed by the University of Kansas and the results were compared with other web-based courses and with similar traditional face-to-face courses. Also, the Meyen model was deconstructed to analyze the individual elements as the initial process for creating the instructional design to be employed by the Academy. This design, in turn, served as the framework against which other online instructional models nationally were compared.

The process for comparison with other models was carried out in two steps. First, a series of standards were developed to guide in the selection of research studies for the content of the modules as the basic OSEP requirement for research-based content. The second set of standards was in the form of pedagogical statements that served as design standards. The latter were

principles that were viewed as important to the final design and content to characterize the modules. The evaluation of design features was carried out by initially identifying online courses that were operational in 1996. In addition, emerging technologies were also reviewed. For purposes of review, courses from the fields of education, medicine, nursing and engineering were screened for the features that appeared to be unique or of sound pedagogy. The requirements of the funding agency (OSEP) were also reviewed for instructional design implications. This process was completed via group sessions during the first three months of the project by the core staff, along with the technology and writing staffs that had been hired at that time.

Additionally, it was necessary to finalize the design so that development could proceed with the confidence that the format would not change. However, the aesthetics design (i.e., the look and feel of the interface) continued to remain fluid until near the end of the second year of the project. The look and feel, as well as the navigation, became the focus on the Beta testing once modules were completed in draft form. (See section on usability and navigation.)

Once the features were agreed upon, an instructional design in the form of a Table of Contents (TOC) for a module was configured. (See Figure 1) Each feature in the TOC was structured to reflect the instructional design and defined to minimize misinterpretation. This instructional design was then refined as the basis for the prototype tool that would be created to automate the production process of producing the modules once the content had been created and validated. It was essential to achieve consensus on the instructional design features so the development of the tool could proceed and so specifications could be written on each element as a guide for the development of content. Ultimately, the content had to be crafted in a format that matched the requirements of the tool. As mentioned, given the amount of content and the number of modules

to be developed, a decision was made that once the instructional design was agreed upon, no changes would be made in the format or instructional design.

Insert Figure 1 about here

Results. Three levels of results in the evaluation process were directed toward creating the instructional design. These included (a) the identification of features to be included in the instructional design, (b) the nomenclature used to label the features, and (c) the nature of the interactivity and the configuration of the features. Of the 21 features selected, only four were not part of the original Meyen model. These included the addition of critical questions, research summaries, a separate practice level and a text version of the multimedia presentation (see Figure 1). In the Meyen model the text version served as the script for the multimedia version of the presentation. The most significant change was the configuration of the features into four levels (i.e., orientation, support, lessons and practice). The results of the instructional design evaluation contributed significantly to the usability and navigation design. The aesthetic design also gave the modules their own branding. That is, it was determined that the modules should all have the same look and feel, as well as features. This enabled students to quickly acclimate to the instructional design and technology, allowed them to focus on the instruction and not the features or the technology.

Production System

Methodology. The module production system is a complex series of procedures, protocols, and software that transforms content, initially developed as text, into interactive media for online delivery. The production software includes:

1. *The content delivery forms*, which are used by content developers to enter text-based content such as the outline, glossary, critical questions, presentations and graphic descriptions.

2. *The source management structure*, which is used to control the state and security of the original text and media data. When prepared for distribution on the Internet, modules typically include over 500 graphic and 2000 data files.

3. *The user interface template,* which describes the look and feel of the modules and are designed for rapid global modification of the user interface.

4. *The content rendering engine*, which marries the harvested content with the user interface templates to publish web-ready data.

5. *The distribution system,* which prepares the web-ready data produced by the content transformation engines for downloading to remote servers.

The evaluation of the production software system was ongoing throughout the project. The initial evaluation focused on developing high-quality, reliable and consistent modules that adhered to the Academy design. The user interface templates progressed through several iterations prior to the release of the first module. Formative evaluation of the initial template designs was primarily conducted internally using faculty and graduate students. Reviews employed email, online threaded discussions and "all-hands" face-to-face sessions. Each of these "Alpha" sessions was

followed by design revisions and follow-up review sessions that progressively involved broader audiences, including diverse settings, undergraduate students, preservice teacher educators, the Academy jurors and the board of governors.

The alpha evaluations of the user interface templates were followed by three Beta evaluations during which the designs were fixed and the focus shifted to uncovering bugs or incompatibilities. The initial requirements called for modules that were accessible to a wide variety of learners, technology and network connectivity options. These requirements demanded extensive Beta evaluation using various versions of computer platforms, operating systems, browsers, streaming media players and connectivity options.

Results. Once the initial modules were released the user interface templates were stamped as golden and the evaluation of the production system turned from improving the module template design to improving efficiency. Formative evaluation of the content delivery component involved email communications and several face-to-face meetings with the writers. Revisions to both the procedures and the software were made to streamline the content entry process. The source management and security system was evaluated regularly during the weekly meetings, through email and the internal production web site. These evaluations resulted in greatly improved version control and a system for managing and backing up the source files.

The evaluation of the content rendering engine also took place during the weekly meetings of the Technical Applications Group (TAG), email and the internal production web site. The content

rendering engine evolved through three major revisions. Each version has become more reliable, easier to use and more flexible.

Content Development

Methodology. The parameters of the content were set by OSEP. That is, the content areas, which were to be research-based, included instruction for preservice teacher education in reading, positive behavior supports and technology across the curriculum. The evaluation methodology for content centered on identifying the research to underlie the content of the modules, developing content according to the specifications of the tool and production system, validating the content to ensure accuracy and currency, and integrating the content into a pedagogy that worked effectively with the online design.

The approach taken by the Academy in the evaluation of content was closely aligned with OSEP's research-to-practice requirement, combined with the validation of content for each component in the instructional design (see Figure 1). A writing team was appointed for each content area, and a board of jurors comprised of national experts in the respective content areas was appointed. The jurors' roles were to provide leadership in the identification of research to be incorporated into the modules, participate in building content maps for each module, and to review the content as developed by the writing teams. Paralleling the work of the jurors in the identification of research, writing teams also generated a review of the literature.

Detailed directions were developed as a writer's guide for each element in the instructional design to facilitate the writing of content to meet the specification requirements of the template.

Modules were written by each team, so several team members were involved in writing each module. They were responsible for generating the content to meet the specifications of the development tool and the editing process carried out by a third party. Once content was drafted, it was subjected to internal review for accuracy, completeness and compliance with the template specifications prior to being edited by a copy editor. Evaluation continued into the production process where the production staff applied their criteria to the content to ensure compliance with the development tool specifications. Once completed, the modules were Beta tested with a focus on usability and navigation.

The most formal approach to the evaluation of content from a validation perspective occurred at the Board of Jurors level. While the writing teams were experienced and knowledgeable in their respective content areas the jurors were selected nationally for their expertise. The jurors served as the external source of evaluation and validation of content. This occurred in several forms. Group meetings were held involving the jurors and writing teams in each content area. Individual jurors were consulted via telephone and some were engaged to participate in the writing. The content maps resulted from collaborative efforts and were collective decisions. Each content map was equivalent to a brief outline of the content.

The content for each module was developed to meet the specifications of the template. Four features within each module were specific to the content of the lessons. For example, lists of 10-15 key points were generated to reflect the intent of the modules, based on the content maps. Subsequently, outlines were developed from which the presentation of content for each lesson was written. This provided four levels of content detail that could be evaluated, each serving as

an independent element in the instructional design. The content for each of these features was subjected to internal review.

The four content elements were combined with the remaining features of the module design (see Figure 1). When this was done, the content was expanded to include elements such as critical questions, directed questions, activities, assessments, glossaries, and readings. The content-based features were derived from the process of generating the content for the presentation element of lessons. It was the presentation element that comprised the majority of the content from each lesson. A presentation was equivalent to a comprehensive multimedia lecture with audio support and represented the focus of each lesson.

Results. The formative procedures applied to the content development process to ensure appropriate content within the research-to-practice model are reflected in the richness of content and quality of the final modules. The evaluation steps included the involvement or national experts as jurors in the selection of research, the generation of content maps, reviews of the different features of content, and the modules in final form. The role of the jurors varied, in that some became more extensively involved as individuals. The role of the jurors was complemented by the functions of the writing teams within each content area.

The major results from the evaluation of the content generation process included: (a) the need to refine the specifications used as guidelines for writers to a more formal manual to minimize the probability of any variance from the specifications; (b) greater reliance on contracting with nationally recognized experts as writers; (c) infusion of a content manager in the content

development process to serve as the quality control liaison between the jurors, the contracted writer(s), and between the writer and the production team; and (d) refinement of procedures for moving content from the writing stage to the production stage.

Usability and Navigation

Methodology. Throughout the development of the instructional design, user interface templates and content, attention was given to those features that would enhance usability and navigation. Two major evaluation targets related to scalability and meeting the goals of the Academy were: (a) producing online modules that had a high level of usability for the students and (b) designing modules that were easy to navigate and that provided flexibility in navigating through the instructional experience. Both targets were considered critical due to the newness of the online mode of instruction at the time the modules were developed. Even though most universities were offering some courses online and some were offering degrees online, most students at that time had no prior experience with online instruction. This also applied to instructors. Evaluation of usability and navigation could not occur until the instructional design was finalized, the user interface template was operational and content had been produced to meet the specifications of the template.

A series of rigorous formative evaluation procedures were employed in the development of the e-learning instructional modules in an effort to release modules to the public with maximum assurance that they would meet state of the art expectations.

The evaluation of usability and navigation was carried out in four stages: (a) Alpha testing was carried out within the Online Academy; (b) Beta 1 testing, conducted with the assistance of 32 colleges and universities nationwide, involved subjecting the online modules to actual use conditions in teacher education programs; (c) Beta 2 testing took place at the University of Kansas as part of a summer course; 40 teachers from two states were enrolled in a module taught by the staff under conditions approximating typical enrollment conditions; and (d) Beta 3 testing, which involved seven institutions and was designed to specifically test the extranet model that involved downloading modules to local servers at participating institutions. While the first stage was conducted internally, the last three were carried out with the assistance of University of Kansas and other institutions of higher education, and a third party evaluator. The four stages combined to provide a very high degree of assurance that the modules would meet the highest standards of usability and navigation feasible.

Alpha testing was the first level of evaluation. Much of this testing involved observing individuals using modules and procedures involving individuals self-reporting on their personal experiences with the modules at different stages of development. Nearly everyone who used the modules while in development, including, but not limited to, the Technical Applications Group (TAG), communicated with TAG about any usability or navigation problems observed during the development process. The TAG team focused predominantly on the technical design and navigation tools of the modules, while program managers closely at both during the first year of the project. All data were given to the TAG team, who in turn oversaw changes to the modules—often in conjunction with the writing teams. The result was the development of a set of modules ready for Beta testing.

Beta 1 testing began after revisions dictated by the results of Alpha testing were made. Beta 1 involved a formal set of procedures for peers outside of the Academy to use in testing the modules. This phase of testing was limited to a small, but representative, sample of 32 institutions of higher education. Beta 1 testing was conducted in 1998. Few institutions or students had prior experience with online instruction. Extensive research was done to determine what would constitute a representative sample of institutions. The first step was to determine the 25 largest teacher-producing preservice programs in the United States and then develop a list of potential institutions representing different demographic profiles. It was determined that a sample should be between 20 and 30 test sites, representing the various types and sizes of teacher education programs. Once this was done, invitations to participate in Beta testing were sent to the 750 education deans, directors and department heads, who constituted the entire membership of the American Association of Colleges of Teacher Education (AACTE).

The invitation consisted of a letter describing the mission of the Online Academy and detailed information about module content and design. Anybody interested in participating as a Beta test site for evaluation of the modules was asked to register online. The registration page was detailed regarding the names of key individuals at the institution that would be using the modules, deans' and directors' names, and many other items. The form also requested information on the size of the institution, the department that would be involved, the number of students enrolled each year and the number graduated, as well as other demographic details of the applicant. Within one month more than 200 institutions had registered at the Academy web site, indicating a desire to participate. After applying the selection criteria and negotiating with potential applicants, 32

preservice programs were selected to become Beta 1 test sites. Of these, 10 were rural, 12 urban, and 10 were suburban/mixed. Twenty-one states were represented. Four institutions were private, while 28 were public institutions.

A total of 230 students were selected as module users. Of those, 182 were women and 48 were men. 165 used PCs and 64 used Macs. Modules were accessed on desktop computers by 210 students; 20 worked on laptops. They connected to the Internet via modems (151) and LAN connections (49). Students were primarily first-year graduate students or seniors in college. More than 100 were between 21 and 30 years of age with approximately 100 being 30 years or older. Nearly 150 students were special education majors, with the rest distributed among early childhood, elementary, secondary, and other majors. Most students reported they worked on modules from home (119), with the rest working in labs, classrooms or other locations.

Beta 1 testing lasted two semesters, and was carried out through a series of formal and informal evaluation methods. The formal measures were recorded online as part of the module completion process. Each module and lesson within the module included an evaluation form to be completed by students online as they finished each part of the module. Students completed a demographic information form before they could begin a module. Finally, there was a form for instructors to complete online. The informal part of the evaluation took the form of phone calls and email messages largely from faculty using the modules in their classes.

Beta 1 testing focused primarily on module structure, navigation system, feedback on special features, the technology capacity of the participants, departmental characteristics, the attributes

of the faculty and students involved in the test, and the instructional environment in which the modules were deployed. Modifications in the module design, usability features, and navigation options were made as a result of feedback from this stage.

Academy personnel conducted beta 2 testing with 40 teachers enrolled for credit under conditions anticipated during implementation. While the teachers were not enrolled in preservice programs, they were considered typical of teachers who enroll in masters degree programs where the modules might be used. The Online Academy worked with directors of special education in Kansas and Missouri in recruiting teachers to participate. A total of 118 applications from K-12 special education teachers, administrators, and students preparing to teach were received. The final sample consisted of 40 subjects pulled from this applicant pool in a stratified, intentional (nonrandom) procedure based on prioritized features (gender, age, and technology experience). If students dropped, replacements were selected from those remaining in the subject pool. Over the course of Beta test 2, 58 individuals were drawn from the residual pool to maintain the continuous enrollment at a minimum of 40 subjects.

The subjects' age ranged evenly from 22 to 55 years of age, with a mean of 38.5 years. Although an equal number of each would have been preferred, the ratio of men to women was unbalanced with 3 males and 37 females even though all males who applied to participate were included. When technology experience was defined as “comfort using e-mail and World Wide Web on a daily basis,” the subjects balanced, with 48% claiming to be comfortable using telecommunications and the remaining 52% claiming to be novices. By contrast, when technology experience was defined as “experience with online courses,” only 6 of the 40 subjects declared experience, meaning 85% had not taken courses online. Because Academy staff

conducted the Beta 2 testing a variety of options were employed to pursue in-depth interviews on any concerns identified. Phone interviews with subjects were conducted to ascertain problems experienced by subjects and the implications of those problems.

Beta 3, the fourth and final phase of the module formative evaluation program, began with the identification of colleagues that were engaged in online instruction and willing to help in this process. Beta 1 testing had involved institutions with little online experience. The review then moved to institutions that were known to have the technological capabilities and demonstrated interest in working with the Academy in this final Beta test phase. This search also included Beta 1 test sites that had been especially knowledgeable and helpful. Approximately 20 sites were considered, and after initial contacts and discussions, the list was narrowed to 10 possible institutions. Following further discussions about schedule conflicts, department requirements, and several other matters, seven institutions were finally selected to participate in Beta 3 testing. Most were familiar with Academy modules, and one was a former Beta 1 test site. This was important because this stage of testing was carried out using modules that were accessed exclusively from servers located at the Online Academy, rather than from local servers at individual institutions. This was done so the Academy could monitor access and technology problems.

The Academy did not intervene in how the institutions used the modules. The intent was to observe their experience in the installation process under normal conditions to ascertain any modification required in the final version. Detailed instructions were developed for Beta 3 test sites. Special Help files, FAQs, and a Getting Started Tutorial were developed, and made available online. Special help lines for both phone and email communications were set up, and

the entire process lasted approximately two semesters. The goal was to take one last, hard look at the modules and their use by institutions under typical implementation conditions.

Results. The following discussion of results is broken down by testing stage as above.

Alpha testing. During this phase, changes were made in the module design. Specifically, navigation and design elements were modified to minimize user error and to maximize the users effective and appropriate interaction with the content of the module(s). Because of the complexity and size of modules, additional help features and navigation aides were included. The Orientation section was reworked to better explain how the modules were supposed to be used and what results should be expected. Also, module assessments had to be changed in order to present the look and feel of professional software and quality content.

Beta 1 testing. Experience with various levels of technology usage varied among all students. On a scale from 1 to 5, with 5 being the highest level of experience, students ranked themselves 4.1 on use of word processing, 3.6 on use of e-mail, 3.2 on use of the Internet, and only 2.2 on knowledge of online instruction. While students had experience in using technology, students in the sample had very little prior knowledge of online instruction. This was to be expected, since online instruction was very new at the time Beta test 1 occurred in 1998.

Students were asked to rate the modules on several dimensions, using a rating scale of 1 to 5, with 1 being most negative and 5 being most positive. They report having the most difficulty in controlling the RealNetworks audio player—with an average ranking of 1.9 (see Table 1). This

was anticipated, as streaming media was new and few had experienced using RealNetworks audio player. In subsequent testing and in implementation that problem disappeared.

Insert Table 1 about here

They were also asked to comment on the lessons within each module (see Table 2). Most students (120) did not use the audio player, probably because of the difficulty they had in controlling it. However, all 141 of the students indicated that the instructional goals, as presented in the modules, were clear. About half felt the technology helped in learning the content. Most (125) said instructions were clear, and 103 thought that the layout of the module design helped them learn. Most (135) believed that the lessons were well integrated, and 103 felt that they were engaging and interactive. This was a reasonable response considering that this was a Beta test version.

Insert Table 2 about here

The students' rating for use of the audio player was 1.9 on a 5-point scale and, as might be expected, only 21 of the 141 students used it. At the time of this test, RealNetworks audio player was unstable on the Mac platform and almost none of the students had it installed on their PCs, even though it was available. It was cutting-edge technology at the time, and too new for most to know about—let alone understand how to download and use. It was known that by the time implementation would occur, the technology would be stable across platforms. This is now considered old technology and students are using audio players such as Napster to download and play audio in increasingly large numbers. Being a little ahead of the cutting edge can be

problematic when Beta testing. However, emerging technologies must be considered in the development of online instruction.

Students and faculty were provided open-ended opportunities to comment on the lessons and modules. The following contains a representative sample of comments by students:

Student Feedback Comments

1. It was definitely a new approach to learning for me. I liked it.

2. The lesson was ok if listened to on audio. However, when reading, it was too long.

3. The lesson is hard to navigate when you wish to go forward or back several pages.

4. I learned a lot of valuable information from these modules that I will apply to my teaching situation.

5. Lessons need to be shorter and more help needs to be provided.

6. I enjoyed the activities and discussion with an immediate response.

7. The glossary was great.

Although the focus of this Beta test was the student as user, unsolicited comments and questions were received from faculty over the two-semester testing period. A few representative comments include the following:

Faculty Feedback Comments

1. In the spring I will be a member of the teaching team, which includes this topic. I think it offers a really good 'validation' of the techniques and procedures that Arizona educators are following. So please continue to count me on the Beta test site team!

2. For me, I learned that I would make the use of the resource as part of an out of class assignment after I am sure students know how to use it. This would be explained on the syllabus - I particularly like the organization, literature from which you have drawn for functional assessment information.

3. A number of the students in my classes do not have access to e-mail. They have been attempting to complete the modules and send their responses through our open access labs here on campus. These open access labs are not configured for individuals and therefore, their responses have not been coming through … the students who have their own computers and e-mail accounts at home are having no difficulty. In fact yesterday one student announced she had even received a response back from [an Academy staffer]. She was impressed. I'm enjoying working with the modules, and plan to continue to use them.

4. That [module] is all I need and I am a happy camper. I appreciate all you are doing with the Online Academy.

5. I think that you folks have provided us with some of the best instructional materials that exist online right now.

Academy staff, in order to determine if an error or problem could be replicated, reconstructed very technical problem identified by students or faculty. This was essential to the revision process.

Beta 2 testing. Researchers collected and analyzed data with a variety of qualitative and quantitative methods. For example, one dissertation (Chiung-Sui Chang, 2000) compared the effect of the level of technological proficiency on variables such as performance, self-efficacy, and attitude toward instructional features. No significant differences in performance, self-efficacy, and attitude were found across the levels of technological proficiency. A second project took the form of an in-depth case study of module features. This work helped module developers understand "server-side" versus "user-side" features necessary for addressing co-mingled technical problems (e.g., when subjects used the module navigation and multimedia RealPlayer technology with different browsers and platforms during implementation). The primary finding was that the amount of experience with technology had no relationship to the level of success in the class.

Beta 3 testing. This stage of testing was important because it focused on the extranet model. It was found that system administrators at extranet sites had difficulty readdressing the links to audio content of modules to be delivered from their servers. The number of links that were necessary to change and the complex coding of Online Academy modules complicated the process of readdressing pointers to audio resources. Some system administrators also experienced difficulty properly configuring their Real Server. As a result, additional help resources were created, including several procedural guides about Real Server and installing Online Academy modules. Additionally, the Online Academy modules were reengineered to change the way audio resource readdressing was accomplished. This significantly decreased the time and effort system administrators had to expend to make the audio resources work from servers under their control, from as long as three hours to 30 minutes to permanently install on local servers. Improvements to the download system were also implemented, which allowed for improved personalization by institutions.

Implementation/Distribution

Methodology. The requirements for the Academy module distribution system called for an-easy-to use extranet system that would permit authorized system administrators at participating institutions to access modules compressed in two formats (.tar.gz and .zip). This was to be done through a controlled environment that tracked the downloads of each module. The intent was to give institutions maximum control of the maintenance, use and administration of the modules. The Academy also wanted faculty members at participating institutions to be able to reconfigure the modules to use them in part or in total. Several solutions were discussed in weekly and special meetings of TAG.

A number of designs and systems were created and tested before development of the final system, which uses a database to store user information and download status. The download system was tested extensively in-house on multiple platforms before an external beta test was conducted. With assistance from system administrators at the University of Kansas campus computing center and the Edwards Campus of the University, several tests were conducted to determine the reliability and usability of the download system. Based on these tests, several changes were made to increase usability. The most substantial change was reworking the system to address media links in the module to reduce the workload of the system administrator related to making the module function in unique server environments. Certain refinements of the download interface and instruction set were also accomplished. Once full public release of Academy modules was underway, trouble reports were monitored to ensure usability and reliability of the system. Because of the small number of trouble reports, it was possible to contact most of the individuals reporting trouble with the download process to provide them assistance and to learn more about the trouble that they experienced. These contacts resulted in further refinements to the instruction set and Help Files provided to system administrators who downloaded and installed modules on servers under their control at participating institutions.

Results. The original system required the system administrator to make multiple changes on multiple pages. This was a laborious process, taking considerable time to locate each page, locate the links on the page and the change them. It is difficult to estimate the typical or average time for installation per module because of individual differences. Several system administrators were unable to perform an installation because their skills were inadequate. During the evaluation

stage of implementation using the extranet model it was found that three hours was not a bad time in installing a single module. This was viewed as unsatisfactory both from the perspective of the Academy and the user. With the new system the links have been made and only one modification needs to be made to an easy-to- find file. This cut installation time to less than 30 minutes per module. Given that a typical module contains over 2000 data and five hundred graphic files, and employed streaming media, this seems very reasonable.

Summary of Lessons Learned

The primary purpose of the formative evaluation procedures was to improve the product, processes employed in development, effectiveness of the implementation and usability by students. Staff reviewed the results from the perspective of determining what changes needed to be made in the design of the online modules or the processes used to create the modules. Following is a summary of the lessons learned as a result of the evaluation processes.

Instructional Design

1. As instructional design features are created, it helps to define them to ensure a common understanding among the staff as to their function and the features impact on content and/or navigation.

2. The development of online instructional designs needs to be coordinated with the development of production systems.

3. In online development projects that produce significant amounts of online instruction design and production decisions must be made prior to initiating the content generation process.

4. Online instruction intended for national dissemination must be sensitive to the technology capacity of the target audience.

5. Participation of content writers and production process developers in decisions on instructional design is important in projects that create significant amounts of online instruction.

6. Setting standards to guide the development of instructional designs facilitates collaborative work in creating the actual design.

7. As production begins, new technologies will emerge that merit consideration. However, the value of adding new features must be weighed against the time and costs of retrofitting the production tools that have already been developed.

8. When creating the instructional design of online instruction, although aesthetics and navigation are important issues, those decisions can remain fluid while content development begins.

9. In developing instructional designs for online instruction, the focus should be on sound instructional principles and the nature of the content in contrast to focusing primarily on the capabilities of the technology.

10. Students value the ability to access any instructional resource in an online instructional program.

11. Many students prefer text resources as an option to multimedia presentations. For example, sometimes they are unable to access the multimedia resources, but can access the text resources or even print them.

12. Providing several levels of content to students for review is advantageous as it allows students to determine what they need to study and increases their efficiency.

Content Development

1. The use of national experts as jurors responsible for setting the content parameters and identifying supporting research was effective.

2. When focusing on research to practice, it is important to be sensitive to the variability in the amount and quality of research carried out relative to content areas.

3. Compliance with content specifications is closely related to the intensity and quality of training writers receive.

4. When a number of people are writing content for different features of a single module, coordination is essential to ensure that all parts come together on schedule.

5. Writing content against specific specifications with structure and length limitations is a new experience for many. It calls for close monitoring of work in progress so that corrections can be made in process if necessary.

6. Content writers must have an understanding of the ultimate product so that they appreciate the importance of adhering to specifications.

7. While there are some advantages in having writers on staff, since their work can be more easily monitored and communications are facilitated, there are also advantages in contracting for writing. For example, the latter allows for a larger pool of content experts.

8. Carefully defining each feature of the instructional design prior to beginning to create the production system enhanced the structuring of the content development process.

9. Placing a moratorium on new features once content development began proved to be a wise decision. Had this not occurred and new features were added, the task of retrofitting all completed modules would have been very costly in time and resources.

Production System

1. It is essential that agreement be reached on the instructional design and that the content parameters be well defined before the technical aspects of the production system are developed.

2. Responsibility for quality control needs to be vested and those responsible should be fully knowledgeable about the template requirements and capable of communicating the specifications to the content writers.

3. Once the design is final and the production system is in place, training must be provided to all writers to ensure compliance with the specifications.

4. Alpha testing involving naive participants is central to building a quality system. It needs to be carried out prior to finalizing the system.

5. Because the content validation model employed involved national content experts, the Beta testing focused on navigation and usability. This worked well as it allowed the staff to concentrate on the features that enhanced student progress through the online modules.

6. Creating our own production system and development tool proved to be a good decision, as it allowed us to control the features of the system without being dependent on an external source to maintain support for the system. Because we were committed to making the modules available to institutions of higher education without cost, we were not tied to a fee structure.

Usability and Navigation

1. Testing usability and navigation using staff members is of value even though they are familiar with the content and knowledgeable of the navigation options.

2. The involvement of individuals in Alpha testing who are naïve about the content and inexperienced in e-learning is essential to determining the effectiveness of navigation systems and the preferences of learners.

3. The level of interest on the part of institutions of higher education in e-learning was sufficient to generate a significant response among universities wanting to volunteer to participate in beta testing.

4. A common look and feel for the instructional design is important. Student's value not having to learn a new design with each new course.

5. Multiple navigation systems are helpful. Students prefer having several options and quickly form a preference for a navigation pattern.

6. The level of computer skill possessed by students does not affect their performance in e-learning instruction.

7. Formative evaluation procedures can be designed to yield specific results that can be employed in the revision of e-learning content and design features.

8. Students can adequately describe problems they encounter or features that are missing.

9. If a problem is major, students will be consistent in identifying the problem and in offering proposals for changes.

Implementation/Distribution

1. The decision to use an extranet model for installing the online instruction on the participating institution's server needs to be made early in the development process.

2. The time required to install online instruction via the extranet model can be reduced to a few minutes per course.

3. The extranet model is valued by participating institutions as it allows them full control of the online instruction.

4. Having signed agreements with the participating institutions designating contact people and committing server space adds a level of accountability that facilitates installation.

5. Instructions for installing online instruction via an extranet model can be placed on online and effectively used in making installations.

6. Implementation is most effective when on-site systems administrators and faculty work as a team in the installation process.

7. Maintenance of a web site is an effective vehicle for communicating with participating institutions during the implementation process.

ACKNOWLEDGMENTS

Preparation of this article was supported in part by the Office of Special Education Programs and Grant #HO29K73002.

We acknowledge the contributions of the Online Academy staff, the individuals who served on the various advisory boards, jurors and the many institutions, and individuals who contributed to the evaluation of the online modules and the production tool. Appreciation is expressed to the Center for Research on Learning and the Information & Telecommunications Technology Center whose collaboration created the e-Learning Design Lab in which this work is being conducted.

References

Beyer, B.K. (1995). *How to conduct a formative evaluation.* Alexandria, VA: Association for Supervision and Curriculum Development. (ERIC Document Reproduction Services NO. ED 391830)

Bodily, S., & Mitchell, K.J. (1998). An educator's guide to evaluating the use of technology in schools and classrooms (on-line). Available: http://www.ed.gov/pubs/EdTechGuide/appa.html

Chang, C.-S. (2000). *Efficacy in college student performance in online instruction*. (Unpublished doctoral dissertation, University of Kansas).

Dodge, B.J. (1994, February). *Design and formative evaluation of PLANalyst: A lesson design tool.* Proceedings of selected research and development presentations at the National

Convention of the Association for Educational Communications and Technology Nashville, TN.

Maslowski, R., Visscher, A.J., Collis, B., & Bloeman, P.P.M. (2000). The formative evaluation of a web-based course-management system within a university setting. *Educational Technology, 40*(30), 5-19.

Meyen, E.L., Lian, C., & Tangen, P. (1997a). Teaching on-line courses, *Focus on Autism and Other Developmental Disabilities, 12*,(3), 166-174.

Meyen, E.L., Lian, C., & Tangen, P. (1997b). Developing online instruction: One model. *Focus on Autism and Other Developmental Disabilities, 12*(3), 159-155.

Nelson, W. A., & Smith, R.E. (1994). *Learning how to teach: a computer-based system to enhance teacher reflection*. Proceedings of selected research and development presentations at the 1994 National Convention of the Association for educational communications and Technology Sponsored by the Research and Theory Division.

RAND (1997). *Evaluating challenge grants for technology in education: A source book.*. Santa Monica, CA.

Price, R.V., Repman, J. (1995). Instructional Design for College-Level Courses Using Interactive Television. *Journal of Educational Technology Systems, 23*(3), 251-63.

Youngman, T., Gotcher, L., Vafa,S.,Dinsmore, S., & Goucher, O.B. (2000, February). *A university design team approach: Developing courses for on-line distance education*. Proceedings of the Society for Information Technology and Teacher Education International Conference, San Diego, CA.

Table 1

Student Evaluation of Modules

Controlling audio player	1.9
Easy access to features	3.4
Effective glossary	3.8
Module review features	3.6
Quality of screen design	3.9
Usefulness of screen design	3.7
Response time of features	3.5

(1=negative and 5= positive N=32)

Table 2

Student Evaluation of Lessons

	Yes	No
Use of audio player?	21	120
Instructional goals clear?	141	0
Does technology help learning?	71	70
Clear instructions?	125	16
Did layout help?	103	38
Appropriate lesson integration?	135	6
Interactive and engaging?	103	38

(N=141)