

GUIDELINES FOR EVALUATION OF INTERNET-BASED INSTRUCTION

Cheryl J. Hamel

**Institute for Simulation and Training, University of Central Florida
Orlando, FL**

David L. Ryan-Jones

Robert T. Hays

**Naval Air Warfare Center, Training Systems Division
Orlando, FL**

ABSTRACT

The Department of Defense (DoD) established the Joint Advanced Distributed Learning (ADL) Co-Laboratory to support the implementation of ADL within DoD. As part of this initiative, the Joint Co-Laboratory is to provide technical assistance to program managers responsible for the development and fielding of ADL systems. This paper describes one of the efforts under this initiative. The goal of this project is to establish documented guidelines for the design and evaluation of Internet-based training and performance aiding. These guidelines are being developed for the Joint ADL Co-Laboratory by the University of Central Florida's Institute for Simulation and Training. The guidelines are being identified by reviewing the literature on development of web-based instructional environments, web-site development and computer-based training, including issues relating to content, format, instructional support features, standards compliance, learning management, tutoring, usability, hardware, and instructor and trainee skill requirements. Procedures for evaluation of ADL materials are discussed, and a sample of the preliminary guidelines is provided.

AUTHORS

Cheryl J. Hamel, Ph.D. Adjunct Faculty member, Department of Psychology, University of Central Florida (UCF), and Research Scientist at UCF Institute for Simulation and Training. Dr. Hamel has participated in the design, development, and implementation of computer-based and Web-based training systems. She has a number of publications and presentations related to guidelines for evaluation of computer-based training, and features of Web-based instruction for Psychology. Dr. Hamel holds a doctoral degree in experimental psychology and holds additional degrees in psychology and mathematics. She may be contacted via electronic mail at *chamel@ist.ucf.edu*.

David Ryan-Jones, Ph.D. Research Psychologist in the Training Technology Development Branch at the Naval Air Warfare Center Training Systems Division, Orlando, FL. Dr. Ryan-Jones received a Ph.D. in Experimental Psychology from Florida State University. His interests are in the development and evaluation of web-based training and instructional support features. He may be contacted via electronic mail at *Ryan-JonesD@navair.navy.mil*.

Robert T. Hays, Ph.D. Senior Research Psychologist in the Training Technology Development Branch at the Naval Air Warfare Center Training Systems Division, Orlando, FL. He received a Ph.D. in general experimental psychology from Virginia Commonwealth University in 1979. Currently, his main areas of research are training applications of virtual reality technology, intelligent pedagogical agents, and advanced distributed learning. Dr. Hays is the author of over forty journal articles, technical reports, and other technical publications. Before working at the Training Systems Division, he was a Research Psychologist with the Army Research Institute. He may be contacted via electronic mail at *HaysRT@navair.navy.mil*.

GUIDELINES FOR EVALUATION OF INTERNET-BASED INSTRUCTION

Cheryl J. Hamel

**Institute for Simulation and Training, University of Central Florida
Orlando, FL**

David L. Ryan-Jones

Robert T. Hays

**Naval Air Warfare Center, Training Systems Division
Orlando, FL**

The Department of Defense (DoD) has established the Joint Advanced Distributed Learning (ADL) Co-Laboratory to support the implementation of ADL within DoD. The function of the Joint ADL Co-Laboratory in Orlando is to promote collaborative and rapid development of ADL prototypes and ADL system acquisitions, principally among the Services training systems development components. As part of the ADL initiative, the Joint Co-Laboratory is to provide technical assistance to program managers responsible for the development and fielding of ADL systems.

This report describes one of the efforts under this initiative. The goal of the three-year effort is to establish documented guidelines for the design and evaluation of Internet-based training and performance support. Scientifically validated principles for design are not yet available to support journeyman development of ADL material. It is hoped that the guidelines will reflect the needs of DoD and will be aligned with ADL goals and user requirements. The guidelines are being developed for the Joint ADL Co-Laboratory by the University of Central Florida's Institute for Simulation and Training.

The preliminary evaluation guidelines, which are the focus of this report, do not yet comprise a DoD directive, implying mandatory compliance with standards. Rather, they have been compiled to provide the best available guidance and direction to evaluation and improvement efforts that support the development of high quality ADL products. Adherence to the guidelines is recommended, but not required.

WHY IS EVALUATION IMPORTANT TO AN ADL PLAN?

Evaluation is about collecting data to judge the merit and worth of any aspect of a program. Recently the Department of Defense (DoD) has embarked on a DoD-wide endeavor to implement Advanced

Distributed Learning (ADL). According to the Department of Defense Strategic Plan for Advanced Distributed Learning (Office of the Under Secretary of Defense for Personnel and Readiness, April, 1999), the Secretary of Defense vision of training technology is to ensure that DoD personnel have access to the highest quality education and training that can be tailored to their needs and delivered cost-effectively, anytime and anywhere. (p.8)

DoD's ADL program is making an important investment in development of Internet-based resources that provide high-quality instruction and performance support strategies to include development of the network infrastructure needed to support them). Right now, while ADL efforts are in the planning stages, it is imperative to think in terms of the data needed to provide convincing evidence of the merit and worth of ADL products and the ADL program. In other words, we need objective evaluation guidelines that will help answer some important questions. The guidelines should address questions like, "Is the instruction and performance support of high quality?", "Do user-learners find ADL resources convenient, useful, and tailored to their individual needs?", and "Can ADL resources be delivered cost-effectively?"

WHY START EARLY?

Planning for ADL evaluation is important because it is easier and less expensive to make changes at the beginning of a program than at the end. Evaluation can't be tacked on as an afterthought. Evaluation planned from the start and kept high on the agenda can produce useful results with modest resources (Draper, 1999).

When evaluation is done early during the design and development stages of a product, it is by convention called *formative evaluation*. Formative evaluations contribute to an evolutionary approach to development. Early evaluative judgments of portions of an ADL product, as they are developed, can

efficiently affect improvements in the product, and evaluative feedback is a form of quality control. Judgments of effectiveness after implementation of the product are called *summative evaluations*. Summative evaluations address the whole ADL product (e.g., resource center, course) and promote evaluation of the product in the wider context in which it is used.

HOW SHOULD WE EVALUATE ADL?

Organizations with ADL programs will decide the scope of evaluation by making decisions about what outcomes will provide the best feedback for ADL improvement and best determine the effectiveness of their ADL programs. Figure 1 indicates a sampling of the variety of evaluation measures available to program planners. Programs can begin with an evaluation of the design and content of an ADL product. Instructional developers, content experts, and other interested stakeholders can continually assess the product as it is being designed and developed. Then, when all or part of the product is developed, user-learners can evaluate the usability of the product and the usefulness of the information it provides. Learning and performance outcomes can be obtained by assessing user-learners themselves, interacting with the product. Finally, after ADL implementation, its use and effectiveness within the organizational context can be evaluated. As the evaluation effort is expanded, the expense and complexity of measures increases. Many organizations are in the relatively early stages of the

ADL planning-design-development-implementation process, making this a good time for them to plan how ADL will be evaluated throughout all of the stages. In the best of all worlds, as ADL resource centers and other ADL systems are designed, formative evaluations will contribute to highly evolved, high quality products. Then, as ADL products are implemented, they will be subjected to summative evaluations according to an organization's ADL plan.

PROJECT DESCRIPTION

Formative evaluation of ADL products should be based on well-documented principles of design. The Institute for Simulation and Training, University of Central Florida has been tasked by the Naval Air Warfare Center Training Systems Division to develop design guidelines for the Joint ADL Co-Laboratory that will give direction to the evaluation of ADL throughout DoD. The project is developing guidelines for the design of ADL resource centers, and so the scope of the design and evaluation process includes guidelines for online courses, workplace performance support systems, discussion forums, information libraries, and the technical infrastructure to support ADL. These guidelines provide the foundation for formative evaluation procedures during the early stages of ADL system development. The project, which began in June 2000, expects to produce updated versions of the design and evaluation guidelines before project completion in 2003. In the first year, in

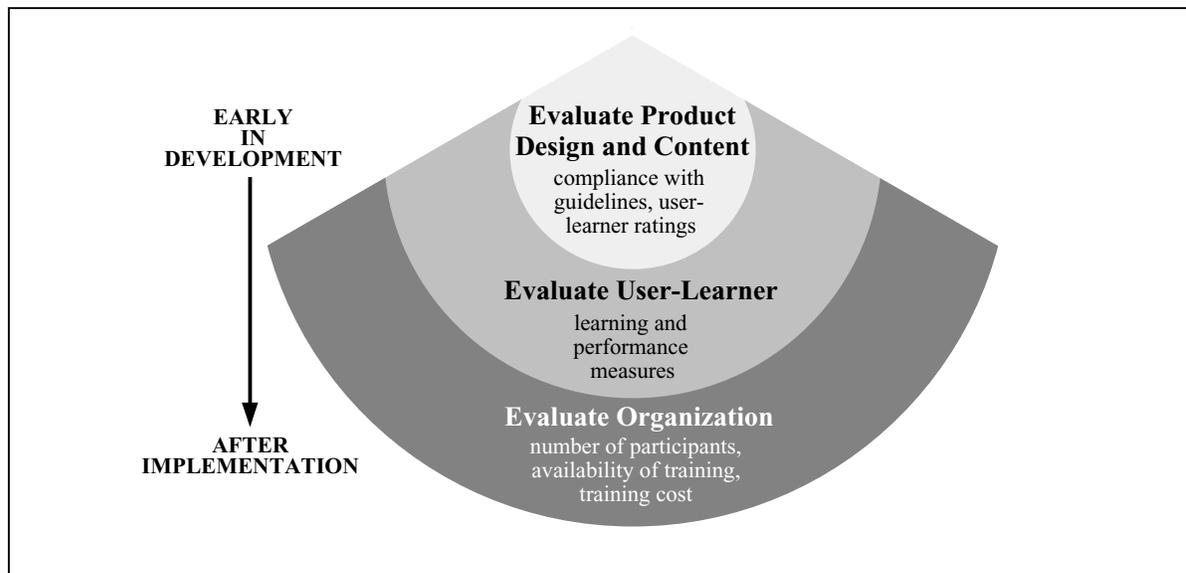


Figure 1. Scope of evaluation and associated outcome measures

keeping with the DoD vision of providing the highest quality education and training for DoD military and civilian personnel, the evaluation guidelines focus on user-centered evaluation of interface design and evaluation of high quality instructional content. The first set of ADL design guidelines are divided into two categories: interface design and design of content.

Some content guidelines pertain to those features particular to ADL resource centers and Web delivery of instruction, such as guidelines for text-based online conferencing (e.g., chats, discussion forums). Other content guidelines prescribe features of good quality instruction, many of which are independent of the technology used to deliver the instruction. Interface design guidelines are based on published research and theory about good human-computer interface design in the context of

Web-based learning and/or performance support systems.

The set of guidelines can be used to design ADL resource centers that support learning and performance, and compliance with the design guidelines can be used as criteria for evaluating the resource centers. Table 1 lists examples of content and design guidelines. The content and design guidelines are easily reformulated into a checklist that can serve as an evaluation tool for formative evaluations of ADL products. For instance, the first design guideline in Table 1 is Text line length should be around 60-70 characters, using standard font sizes. The corresponding evaluation checklist item reads, "Text line lengths are around 60-70 characters." With continued use, throughout the project, the evaluation checklist itself will be subject to evaluation and quality improvement.

Table 1. Examples of interface design and content guidelines and guideline categories.

<p><u>INTERFACE DESIGN GUIDELINES</u></p> <p><u>Screen Design</u></p> <ul style="list-style-type: none"> • Text line length should be around 60-70 characters, using standard font sizes (Beltr n, 1996). <p><u>Page Design</u></p> <ul style="list-style-type: none"> • Limit web page length to 1-3 printed pages of information (Lynch & Horton, 1997). <p><u>Site Design</u></p> <ul style="list-style-type: none"> • Include site identifiers on every page (Nielsen, Nov. 1997). • Limit the depth of menus to 4 or 5 levels (Schneiderman, 1998). <p><u>Aesthetics</u></p> <ul style="list-style-type: none"> • Limit a single screen to four colors (Remmers, 1998). <p style="text-align: center;"><u>CONTENT GUIDELINES</u></p> <p><u>Interactivity</u></p> <ul style="list-style-type: none"> • Provide feedback to students in a manner that is constructive and non-threatening (Phipps & Merisotis, 2000). <p><u>Content Relevance</u></p> <ul style="list-style-type: none"> • The content should be written clearly for target users (Shenot, 1999). • The content should support the goals and objectives of the chosen audience (Schleyer & Johnson, 1999). • Establish and maintain a work context (Gery, 1995). • Content must be supported by references (Schleyer & Johnson, 1999). <p><u>Instructional Strategy</u></p> <ul style="list-style-type: none"> • For teaching procedures, use this template: Provide an overview of the entire procedure, demonstrate each step and identify its critical elements, coach the trainees as they practice each step, require the trainee to demonstrate the mastery of each step, integrate all steps, provide systematic practice toward fluent application (Thiagarajan, 2000). <p><u>Pedagogical Approach</u></p> <ul style="list-style-type: none"> • Domains requiring mastery of lower-order skills should not use a constructivist approach (Campbell, 2000).

Several Web authors have distinguished between design principles, guidelines, and conventions (IBM Corporation, 1997; Ohnemus, 1997). The *guidelines* above are based on a set of established *principles* of cognitive/behavioral psychology and instructional design theory. *Conventions* prescribe specific design decisions compatible with an organization's needs and goals. For example, many government organizations have published internal guidelines for establishing and maintaining web information. It is expected that many instances will arise when there may be more than one way to design a feature. In those cases, organization *conventions* will dictate the final design.

In the next phase of the project, guidelines that support technical standards will be integrated into the existing set of content and design guidelines. Technical standards are likely to include required software, bandwidth limitations, file naming conventions, standards for site maintenance and internal security, procedures and standards for server compatibility, external security, and user access control. Learning management system features will be prescribed so that DoD components can share sections of courses (learning objects). As the DoD Sharable Course Object Reference Model (SCORM) (<http://www.adlnet.org>) is developed, guidelines will be expanded to include guidance for SCORM compliance.

Two ADL resource centers under development by IST provide examples of implementation of the guidelines and are being used to validate the guidelines. IST's development process involves rapid collaborative prototyping (Dorsey, Goodwin, & Schwen, 1997) coupled with ongoing formative evaluation, and usability and performance testing. IST developers are collaborating with representatives from all military services, the federal government, and academia to design and develop the two prototypes according to the guidelines. Evaluations, using the design checklist, are being used to measure the prototype ADL resource center's compliance with design guidelines. Design, development, and evaluation will be done in an iterative fashion until the final products are produced in years two and three of the project.

The Guidelines Development Process

The methodology for development of the design guidelines has begun with a search for existing design guidelines for Web-based training and Web-

based performance support systems, followed by a search of general guidelines for computer-based training and web design.

The following types of searches are being conducted: 1) Internet search, and 2) electronic library search (from 1985-present) using PsychInfo, First Search, ProQuest, ERIC (Educational Resources Information Center), and other library databases. On June 21, 2000 a search of the ERIC database using the keywords, "web based instructional guidelines" produced 18 hits. The keywords, "computer based instruction guidelines" produced 350 hits. An ERIC search using the keywords, "multimedia guidelines" produced 251 hits. Several thousand hits are expected by the time a search of all topics is completed.

Guidelines are being derived from the following sources: 1) articles including guidelines *per se*, 2) literature reviews of guideline topics, 3) empirical research that establishes the authenticity of a guideline, and 4) articles about successes with guideline implementation. Priority for inclusion in the set of ADL guidelines will be given to guidelines in articles that come from reputable sources, e.g., well-respected research institutions, successful commercial enterprises, experts in the field, and authors of juried (peer-reviewed) articles.

Evaluation Rating Scale Development

The project plan is to include other evaluation measures besides those obtained from the guidelines checklist. The checklist is a valuable evaluation tool during the early stages of ADL design. But once intact portions of the product are developed, users will be able to evaluate the product as they interact with it. The next step is to produce evaluative rating scales to judge site design and content from the user point of view. This type of evaluation will move the product closer to the important goal of being user-centered.

Rating scales will be derived from the guidelines. Figure 2 provides examples of how several design guidelines correspond to important criteria (principles) of a well-designed ADL resource center. A set of questions can then be produced from these evaluation criteria. The evaluation rating scales will be used to rate the evolving prototype ADL resource centers and make improvements throughout the development process. The Joint ADL Co-Laboratory will make the rating scales available to other ADL resource centers and ADL course developers.

Table 2. Example evaluation criteria and associated design guidelines.

EVALUATION CRITERIA	ASSOCIATED GUIDELINES
<p><u>Content Relevance</u> Question: <i>Was the information provided useful?</i></p>	<ol style="list-style-type: none"> 1. The skills and competencies presented in the instruction should correspond to workforce competencies and foundational skills (Dougherty & Ellibee, 1995). 2. Use real-world predictable and unpredictable situations in the instruction (Dembowski, (1999).
<p><u>Navigability</u> Question: <i>Were you able to navigate the site without getting lost?</i></p>	<ol style="list-style-type: none"> 1. Draw a site map or use other orienting device to illustrate the main areas of the site (Nielsen, Nov. 1997). 2. Websites should always be organized as offshoots of a single home page (Lynch & Horton, 1997).
<p><u>Accessibility</u> Question: <i>Were you able to access all the information?</i></p>	<ol style="list-style-type: none"> 1. Include ALT tags on all image files to ensure that sight deficient users can access the information (Lynch & Horton, 1997). 2. Keep underlines intact for monochrome viewers (Nielsen, Nov. 1997).

CONCLUSIONS

Evaluations can be conducted in many different ways and the evaluation data can serve many purposes. The two methods of evaluation described in this article -- the content and design checklists and the evaluation rating scales -- are needed in the early stages of ADL program planning in order to give direction to developers of ADL resources. The scope of future ADL evaluations will depend on the answers to a number of questions: 1) who is the evaluation for? 2) what is the purpose of the evaluation? and 3) what can be measured, given the time and resources allocated?

ACKNOWLEDGEMENTS

The authors extend appreciation to Stephen Dixon, Rollins College, Orlando, FL, who assisted in the gathering of guideline materials and preparation of the References section of this paper, and to Robert Bishop, University of Central Florida, Orlando, FL, who produced the graphical materials.

REFERENCES

Beltran, B. (1996). *An EPSS interface that people can use*. Retrieved June 20, 2000 from the World Wide Web: http://ourworld.compuserve.com/homepages/bea_beltran/

Campbell, K. (2000). *The web: Design for active learning*. Retrieved May 2, 2000 from the World

Wide Web: <http://www.atl.ualberta.ca/articles/idesign/active1.cfm>

Dembowski, F. (1999). Rigor and relevance in the training of educational administrators. *The AASA Professor*, 23(1). Retrieved June 21, 2000 from the World Wide Web: <http://www.aasa.org/TAP/fall99dembowski.htm>

Dorsey, L., Goodrum, D., & Schwen, T. (1997). Rapid collaborative prototyping as an instructional development paradigm. In Dills, C., & Romiszowski, A. (Eds.). *Instructional development paradigms*, (pp. 445-465). Englewood Cliffs, N.J.: Educational Technology Publications.

Dougherty, B., & Ellibee, M. (1995). *Curriculum quality standards for school-to-work: A guidebook*. Retrieved June 21, 2000 from the World Wide Web: <http://ncrve.berkeley.edu/AllInOne/MDS-955.html#Heading13>

Draper, S. (1999). *Observing, measuring, or evaluating courseware: A conceptual introduction*. Retrieved June 20, 2000 from the World Wide Web: <http://www.icbl.hw.ac.uk/itdi/implementing-it/measure.htm>

Gery, G. (1995). Attributes and behaviors of performance-centered systems. *Performance Improvement Quarterly*, 8(1), 47-93.

IBM Corporation (1997). *Web design guidelines*. Retrieved June 20, 2000 from the World Wide Web:

http://www.ibm.com/IBM/hci/guidelines/web/web_design.html

Lynch, P. & Horton, S. (1997). *Web style manual (2nd ed.)*. Yale Center for Advanced Instructional Media. Retrieved June 20, 2000 from the World Wide Web: <http://info.med.yale.edu/caim/manual/contents.html>

Nielsen, J. (1997). *Interface design for Sun's WWW site*. Retrieved June 20, 2000 from the World Wide Web: <http://www.sun.com/sun-on-net/uideesign/>

Nielsen, J. (Nov. 1997). *The tyranny of the page: Continued lack of decent navigation support in version 4 browsers*. Retrieved June 20, 2000 from the World Wide Web: <http://www.useit.com/alertbox/9711a.html>

Office of the Under Secretary of Defense for Personnel and Readiness. (1999). *Department of defense strategic plan for advanced distributed learning*. Washington D.C.

Ohnemus, K. (1997). Web style guides: Who, what, where. *Proceedings of the 15th Annual Conference on Systems Documentation*, 189-197. Snowbird, Utah. Retrieved July 3, 2000 from the World Wide Web: <http://www.acm.org/pubs/contents/proceedings/doc/263367/>

Phipps, R., & Merisotis, J. (2000). *Quality on the line: Benchmarks for success in internet-based distance education*. Washington D.C.: National Education Association.

Remmers, E. (1998, January). *Guidelines for WWW-based support environments for education professionals*. Retrieved June 20, 2000 from the World Wide Web: <http://users.edte.utwente.nl/remmerse/litstudy/>

Schleyer, T., & Johnson, L. (1999, March). *Guidelines for the design of educational software*. Retrieved June 26, 2000 from the World Wide Web: <http://www.temple.edu/dentistry/di/edswstd/>

Shenot, A. (1999, September). *Designing distance learning courses*. Retrieved June 21, 2000 from the World Wide Web: <http://www.education.wisc.edu/outreach/deweb/design.html>

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human-computer interaction* (3rd ed.) . Reading, MA: Addison-Wesley Publishing.

Thiagarajan, S. (2000, January). *Rapid Instructional Design*. Retrieved June 21, 2000 from the World Wide Web: <http://www.thiagi.com/article-rid.html>