DOCUMENT RESUME

ED 373 750 IR 016 760

AUTHOR Rathbun, Gail A.; Goodrum, David A.
TITLE Evaluating the Impact of Instructional Multimedia: Workable Techniques.
PUB DATE 94
NOTE 12p.; In: 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 (16th, Nashville, TN, February 16-20, 1994); see IR 016 784.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Computer Software Development; *Computer Software Evaluation; *Educational Media; Educational Technology; *Evaluation Methods; *Formative Evaluation; Instructional Design; *Models; *Multimedia Instruction; Research Methodology; Social Science Research; User Needs (Information); User Satisfaction (Information)
IDENTIFIERS Examples; *Multimedia Materials; Stakeholders

ABSTRACT

A framework is proposed for the formative evaluation of multimedia. It describes techniques that have worked well in the evaluation of software development and gives examples of the use of evaluation results. The focus is primarily on the degree to which the instructional multimedia program supports the user's activities and tasks in the user's task environment. Key characteristics are (1) a focus on user satisfaction; (2) integration of evaluation into the design process; (3) use of a variety of techniques; and (4) inclusion of a range of stakeholders. The approach contrasts with the objectives-oriented social-science evaluation model, and it borrows heavily from naturalistic and participant-oriented evaluation approaches. Evaluating usability will provide 80% of what the developer needs. It must also be stressed that evaluation is a concurrent process through all stages of development. One table summarizes evaluation purposes and characteristics. An example of evaluation in practice is described in an attachment. (Contains 7 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

ED 373 750

Title:

Evaluating the Impact of Instructional Multimedia: Workable Techniques

Authors:

Gail A. Rathbun
David A. Goodrum

Instructional Support Services
Indiana University
Bryan Hall 106
Bloomington, IN

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S.Zenor

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

016 760

A variety of conditions and attitudes have made it difficult to develop and use evaluation schemes for "new media." Below we list four issues which pose obstacles to effective evaluation.

- **Continuous change**: A continuous blizzard of technological innovations often blinds designers of instructional multimedia to the fundamental questions to be answered by formative evaluation in the course of development. As a result, "evaluative activities" are put aside in order to expend effort on just getting the multimedia system to work smoothly in its technological environment, and on insuring that the system is compatible with the latest software.

- **Technology focus**: The delivery of multimedia programs can require sophisticated, non-standard, andoften costly hardware and software. Thus the discussion of Multimedia remains focused on resolving hardware/software issues: *who* can tell us *what* to buy, *how* to hook it up, and *how* to keep it working. The necessary concurrent discussion of learning or intellectual activity to be supported by multimedia tools pales by comparison.

- **Evaluation in isolation**: Evaluation is frequently labeled a "phase" unto itself, residing outside of the central development activities of analysis, design, and implementation. Frequently, the only people involved in the evaluation are the prototypical user and the development team, ignoring the interest and vested interests of a much wider group of stakeholders, further isolating the process by keeping it out of a real-world environment. Evlauation is conceived of as a measuring of outcomes and products, requiring statistically valid instruments and experimental groups, and takes place long after key decisions have already been made based on perceived value and usability. Thought of in this way, evaluation seems impractical.

- **"Hyped Media"**: A variety of claims are made as to the superiority of instructional multimedia to more traditional instructional media and methods. Among them are higher grades, improved critical thinking, accommodation of different learning and cognitive styles, and improvement of teaching. A coherent approach to confirming these claims has yet to be described and implemented.

In this presentation we propose a framework for the formative evaluation of multimedia, describe the techniques that have worked best in our software development efforts at Indiana University, and provide examples of the variety of ways in which we used evaluation results. Our emphasis is on what is workable and possible to do given the traditonally limited resources of the instructional designer in an academic setting. Our aim is get the instructional developer to use a variety of evaluation techniques more frequently and with greater confidence.

Our approach

The familiar social science experimental method has long been the accepted approach to evaluating the effects of mediated treatments on learning. Scriven (1987) lists these steps. In order to evaluate a program, you

1. Identify the goals of the program
2. Convert them into behavioral objectives
3. Identify tests (or construct them) that will measure these objectives
4. Run these tests on the target populations
5. Crunch the data
6. Report whether or not, or to what degree, the goals have been met

Characterizing this approach as the "naive social science model," Scriven continues

> There is a standard set of about fourteen questions that need to be investigated in most evaluations, of which only one is the traditional investigation of alleged or hypothesized effects so familiar in social science research. The only points that need to be made are that these questions, be they concerned with cost or alternatives or ethics or unexpected effects or historical background, cannot be ignored; and that a systematic approach to them is possible, with about the same chances of getting an answer as we can expect in the usual scientific or criminological hunt for explanations and theories. (p. 65)

In contrast, we seek to arrive at the answers to these questions from within the development process by focusing primarily on a single question: to what degree does this instructional multimedia program support the user's activities/tasks in the user's task environment? For if the product is not used then it doesn't matter whether it creates potential for or enables more or better learning. Furthermore, the desired learning is most likely to take place far from the machine, as the user reflects, applies, and synthesizes. It is therefore the process, or the quality of the mediation, not the product that should be the focus of instructional multimedia evaluation (Hutchings, 1992; Marchionini, 1990).

Our views on an appropriate evaluation model for multimedia do not, however, exist in a simple bi-polar opposition to what Scriven describes as the "naive social science model. " Rather, our approach can be seen as an eclectic one moving across a continuum from a management-oriented model (related to the social science objectives-oriented approach) to the naturalistic-participant oriented approach at the opposite end (Worthen and Sanders, 1987). (See Appendix A). According to Worthen and Sanders, the management-oriented approach focuses on user needs, de-bugging, and evaluation at all stages of development. The naturalistic approach is helpful for examining innovations-in-use, for portraying the complexities of and educational activity, and for responding to an audience's needs. From our experience of both success and failure in the development of multimedia programs we have seen the following key characteristics emerge in our approach to evaluation:

- a focus on user satisfaction (usability and valuing)
- integration of evaluation into the design process
- use of a variety of techniques
- inclusion of a range of stakeholders

A focus on user satisfaction

What do we mean by user satisfaction and usability? What attracts our faculty to using multimedia is the technological reality of access to a variety of media via the computer. The ability to draw together a huge volume of information in a variety of formats, makes it possible to create and place at the students' and teacher's disposal elaborate

learning/teaching environments not possible only 10 years ago. Thus, the issues of access and control are paramount; that is, the program must facilitate moving about, finding things, and control appropriate to the task and level of user. Romiszowski suggests that when the purpose of the program is to provide access or act as a tool, the appropriate evaluation approach is to measure user satisfaction (Romiszowski, 1990).

To the definition of user satisfaction we have added another dimension: valuing. This aspect is based on the simple idea that a tool viewed as relevant, critical, and of wide applicability is a tool that the user will come to rely on. The user's "valuing statements" comprise a face validity of the evaluated program. Therefore, in all of our evaluation efforts we look for valuing statements.

Integration of evaluation into the design process

Evaluation occurs simultaneously with the analysis, design, and implementation. Within the large scale development project there are many small scale design processes which go through a full problem solving process for purposes of exploring options and developing a fuller understanding of the requirements. The successive phases of development–each with concurrent problem analysis, development and evaluation–become increasingly concrete (Goodrum, Dorsey, & Schwen, 1994).

Evaluation could be said to lead the process. The creation of prototype versions of proposed solutions are for the purpose of evaluation and refinement of specifications. This allows users to have a realistic experience for basing assessments and revisions. Conceptual prototypes allow for early user reaction, feedback, and projection of consequences. Working prototypes allow for hands-on use in the context of the task. If the evaluation process begins early then parts of a design or even an entire design can be discarded before time investment and escalating commitment prevent such corrective action. The use of alternative prototypes in the process helps keep designers and users from locking into a design too early. Frequent user evaluation helps insure usability and provides new ideas for design, content, and evaluative categories.

Inclusion of a range of stakeholders

By including a range of stakeholders early in evaluative activities you collect ideas, gain buy-in and commitment, and avoid unforeseen technical and administrative problems. The key stakeholders most often missing from the iterative design cycle are those involved in implementation and delivery, for example, those responsible for the local area networks which must have the capacity and flexibility to get the multimedia program to the user. The developer may need to consciously expand his or her idea of what a stakeholder is. To determine stakeholders, consider who or what could block the user's access and ablity to work with the program. Another way to find all the stakeholders is to think of everyone having a vested interest, and let the non-stakeholders self-select out of the group.

Use of a variety of techniques

Examining a process requires gathering snap-shots at various stages along the way, calling for a concert of methods, each of which adds "color" to the description. The techniques we have used that have yielded the most usable data have their roots in qualitative inquiry. They are.

- Observation
- Self-report
- Interview

- Peer evaluation (showing and telling others)

The stage of development, the purpose of the program, the target user and other stakeholders, help determine the type of method used and degree of formality in conducting the evaluation and analyzing the results. Another way to approach the choice is to ask (Knussen, Tanner, and Kirby, 1991):

- For what reasons are we doing an evaluation?
- What will we do with the results?
- What resources do we have?

Our Experience

The nature of our multimedia development has been a) creation of presentation packages for faculty clients to use in developing multi-media lectures and b) creation of multimedia programs for students to use in networked computer clusters. The following techniques have proven beneficial to our evaluation process, especially when used together on the same multimedia project.

Client questions. Throughout the development process we ask the client to compare the "old" way with the "new way" of doing things. What was wrong with the "old way" of teaching a topic? Does the innovation help? How does it help? Could you use this in other courses? Do you think that you could take on more of the development yourself? Client questions like these keep the developer and client in a critical appraising mode focused on usability, worth, and value. (See Appendix B.)

Mock-ups. Paper mock-ups may be used at a variety of points in the development process. They are, in a sense, like a structured interview, if presented as a question rather than a *fait accompli.* At the beginning of the process they are an inexpensive and simple way to test ideas for fit with the client's or user's requirements. They can provide data on usability before committing to a design. We have also use paper mock-ups in mid-development when a project has foundered or been dormant. Removing the program from the "high tech" environment helps refocus everyone on task and usability, and the underlying reason for creating the product. It may also invite greater participation from those not at ease with computers and from those who may be new to the project if it is being picked up again after a period of time. The paper mock up signals an openness to critical evaluation. (See Appendix C.)

Observed initial use. This may be done as formally or informally as resources, data needs and stage of development require. After 4 months of development on one program we conducted a field test in two stages: a pre-field test review conducted by instructional developers, two students, and a faculty member; and a full field test with 20 students using the program in a network cluster. The full field test involved detailed observation of two students and a survey of the whole class. The tests identified current and potential problems, highlighted the program's positive points, provided design ideas, and involved three diferent groups of stakeholders. It di d require more coordination than some of the other methods, however, and preparation of the evaluators. (See Appendix D.)

Field trial. To be effective, the field trial must take place under actual conditions that test the limits of the products caabilities. Since a field trial is conducted during actual use, care should be taken to provide backup and support in case the system fails. Observations may be

highly structured or subjective evaluation of an expert (See Appendix E). The structure and nature of the observations again depend on what data is needed and how the data will be used. For example, are you still looking for design ideas, are you firming up a design to bring the process to closure, or do you want to know if the progam is usable in order to seek more resources?

Minute paper. Having the users/consumers respond in writing to an open-ended question can help you adjust your program in mid-course. The time limit forces the writer to divulge only those issues of primary concern. The short length makes it possible to quickly review responses and tabulate results. The openness allows you to pinpoint problems you may not have considered. (See Appendix F.)

Classroom artifacts. Student and faculty client questions, criticisms, compliments, value statements, problem statements are all evaluative in nature. They point directly and indirectly to places where the program either performed well or fell short. Electronic mail now helps keep a record of them. We encourage our faculty clients to keep a record of these artifacts as well as keep a record of mail received from faculty. Setting up a "help" phone service during specified hours can also help capture problems, and provide evaluative data.

Focus group. The focus group, conducted by a person perceived as impartial and open, is an efficient way to interview a number of users at the same time. We have had success with both a highly structured format in which a list of questions was prepared in advance, and an informal format in which small groups were asked one or two open-ended questions. The more informal focus groups helped in defining what issues would become important later in the process. The structured focus group was conducted at a later stage of development, at a point when the important issues had been more clearly identified through prior evaluation. (See Appendix G).

Survey. A survey conducted during later stages of development may complement the structured focus group. From our experience we have identified some critical survey categories:

- How does the student value the learning experience?
- Was the experience relevant to the course and at the same time widely applicable?
- Was the program useful in completing the assigned task?
- How does the student compare the activity with activities in other courses?

The purpose of the survey is to gather data for fine tuning performance and to project the "bottom line" results of using the program. (See Appendix H).

Peer Evaluation. Demonstrating or talking to peers about their multimedia programs helps clients remain in the critical, evaluative mode established at the beginning and maintained during the design process. Peer questions are opportunities for reflection, generate new ideas, open up new partnerships. Peer evaluation may take place among other stakeholders as well, with the same effect of informing current and future design, and identifying other criteria by which the program might be judged. (See Appendix I).

Summary

Our approach to the evaluation of multimedia contrasts sharply to the objectives-oriented

social science evaluation model. Rather than look to one method for answers to all of our evaluation questions we have borrowed heavily from naturalistic and participant-oriented evaluation approaches to obtain the answers. Our underlying philosophy toward evaluation in multimedia development process, based on our actual practice can be summed up this way:

- Find out if your multimedia helps the users do what they want to do.
- Evaluating *usability* gets you most (80%?) of what you need.
- Evaluation is a concurrent process through all stages of development.

References:

Goodrum, D. A. , Dorsey, L. T., Schwen, T. M. (1994). A socio-technical perspective of instructional development: A change in paradigms. Paper presented at the annual conference of the Association for Educational Communications and Technology, Nashville, TN, February 1994.

Hutchings, G.A. (1992). Authoring and evaluation of hypermedia for education. *Computers and Education, 18*(1-3), 171-177.

Knussen, C. , Tanner, G. R. and Kirby, M. R. (1991). An approach to the evaluation of hypermedia. *Computers and Education, 17*(1), 13-24.

Marchionini, G. (1990). Evaluating hypermedia-based learning. In D. H. Jonassen and H. Mandl (Eds.) , *NATO ASI Series, Vol. 67: Designing Hypermedia for Learning.* Heidelberg: Springer-Verlag.

Romiszowski , A. (1990). The hypertext/hypermedia solution–But what exactly is the problem? In D. H. Jonassen and H. Mandl (Eds.) , *NATO ASI Series, Vol. 67: Designing hypermedia for learning.* Heidelberg: Springer-Verlag.

Scriven, M. (1986). Evaluation as a paradigm for educational research. In E. R. House (Ed.), *New directions in educational evaluation.* Philadelphia, PA: Falmer Press.

Worthen, B. R. and Sanders, J. R. (1987). *Educational evaluation:Alternative approaches and practical guidelines*. New York: Longman.

	Management -oriented	**Consumer-oriented**	**Naturalistic and Participant-oriented**
Purpose	Providing useful information to aid in making decisions	Providing information about educational products to aid decisions about educational purchases or adoptions	Understanding and portraying the complexities of an educational activity, responding to an audience's requirements for information
Character-istics	Serving rational decision-making, evaluating at all stages of program development	Using criterion checklists to analyze products, product testing, informing consumers	Reflecting multiple realities, use of inductive reasoning and discovery, firsthand experience on site

Criteria	Utility. feasibility, propriety, and technical soundness	Freedom from bias, technical soundness, defensible criteria used to draw conclusions and make recommendations; evidence of need and effectiveness are required	Credibility, fit, auditability, confirmability
Benefits	Comprehensiveness, sensitivity to informaiton needs of those in a leadership position, systematic approach to evaluation, use of evaluation throughout the process of development, well-operationalized with detailed guidelines for implementation, use of a wide variety of information	Emphasis on consumer information needs, influence on product developers, concern with cost-effectiveness and utility, availability of checklists	Focus on description and judgment, concern with context, openness to evolve evaluation plan, pluralistic, use of a wide variety of information, emphasis on understanding
Limitations	Emphasis on organizational efficiency and production	Cost and lack of sponsorship, may suppress creativity or innovation, not open to debate or cross-examination	Nondirective, tendency to be attracted by the bizarre or atypical, potentiall high labor intensity and cost, hypothesis generating, potential for failure to reach closure

Adapted from Worthen, B. R. and Sanders, J. R. (1987). *Educational evaluation:Alternative approaches and practical guidelines*. New York: Longman

Client Questions

1. What's one of the things you'll be teaching today?
   (looking for a lecture part or even a specific point
   that we can make sure we capture and hilite)

2. How did you teach this in the past?
   (what were the frustration or limitation of teaching it that way.
   i.e., what was wrong with that way of teaching it)

3. How will you be using the technology to help you teach this in your class?

4. How difficult was this to create?
   (and what kind of help did you need?)
   (do you see yourself doing more and more of this on your own?)

5. Is this a better way of teaching? Why? (or why not)

6. What other courses or areas in your field would this be useful for?

Insert Graphic entitled "Mock-ups" here.

Insert Graphic entitled "Observed Initial Use" here.

Field Trial

This is an example of another style of observation, performed at a later point in development than the Observed First Use example in this packet.

M conducted the class like a two-person dance. He used the Komo program to display text mostly. It was an outline he traversed to give students an idea of where they were. It seemed to work reasonably well because he didn't go too deep in the hierarchy. He has someone in the back controlling two slide show projectors. Frequently the images from the slides and Komo overlap. Often, the person controlling the slide projectors moves the images to keep them from falling on the text M is projecting. Frequently, the tool palette from the program can be seen over the slides. M's irreverence is plain--he doesn't seem to care much if the slides overlap the tool palette or not. The person in the back swings the slides back and forth trying to find an empty space for them as M turns Komo on and off. There is no "grid"--there is no place where text always comes up and pictures always come. The presentation is like M--fluid, ever changing, refusing to be categorized and yet paradoxically, existing within the framework of a highly structured outline....

...He wants to be able to access things quickly in any way he wants. The interface is extremely important to him. I can imagine him wanting something with a fluid picture-showing capability. He presses "next slide" on his computer. Up pop two pictures. He points to one of them with his finger on a stylus and drags it off the screen. He pops up some text. He presses "clean up" and all of a sudden, everything fits." He presses another button and the two pictures blow up to fill the screen. Another button and the pictures disappear.

Minute Paper

This question was placed on an overhead projector. Students wrote brief answers to the question which the instructional consulting staff then reviewed to get a "reading" on how things were going at mid-semester. The purpose was to spot problems and fix them before the end of the semester.

What is the effect of the technology used in this course on your learning?

Insert graphic entitled "Focus Groups" here.
Insert graphic entitled "Survey" here.

Peer Evaluation

A faculty developer of a large multimedia project answers questions from her peers as she demonstrates her program. The following is a rough transcript from the video recording of

her demonstration.

Faculty developer: An advantage of this format ...it really does give student access to a much wider array data than ever before and more accessible we've started tools to help them make those connections in their own minds in different sorts of ways graphics, text, video give each student a little bit of something

...what we're trying to improve upon is the interactivity...

Faculty developer stops her demonstration.

Any questions?

Peer question: Do students type their answers out?
Delivery issues discussed

Peer question: How do you get them started without being overwhelmed?
Faculty developer: Simple questions and self-help tutorial. Different students got started in different ways.

Disadvantage: lots of data
Advantage: open ended

If you ask the right questions...The trick is the interface of the questions you ask.
I've challenged graduate students with this data set.

I wanted them to use this as a tool and a learning resource as much as I wanted them to see use it as a text so I didn't allow them to print it out.

Peer question: You're not involved with authoring? the programming?

Faculty developer: The programmers created a miniature set of authoring tools
that have allowed me to -- created a template -- I create the images and can very quickly...

I put it together mostly, I haven't done the scripting. (Faculty developer goes into stack and adds to it. She opens a palette to choose a card format, then she brings in a picture, types in the text)

That's how I put it together. It's been very easy to put it together. I know a lot more about hypercard than I ever did.
a cut and paste and a content problem for me

Peer comment: Very flexible architecture overall.

Faculty developer: Next: use this as a presentation tool
Astound -- allow you to select a few elements -- I'll be working
on that next semester

Peer question on copyright issues.
Faculty developer: Copyright issues -- got to mail out the letters! I've tried to replace

copyrighted material with my own material
any color stuff is mine.

I like to doodle, so I can always create something myself.

Prototype on a CD

At each stage we've gotten student evaluate information
I've been astonished how resilient the students are.