

Evaluation: Student and Peer Review

Charles E. Branch, Ph.D.

*Department of Physiology and Pharmacology
College of Veterinary Medicine
Auburn University, AL 36849*

With each generation of technological development, educators have experimented with applications for the new technologies. Examples have included the use of slides, slide-tape presentations, linear videotapes, and computer assisted instruction (CAI). More recently newer technologies such as interactive videodisc (IVD) and digital video interactive (DVI) have received increasing attention. There is no question that well-written lessons using these technologies can improve education with significantly increased performance and efficiency.

These new instructional programs need critical review in order to ensure that:

- 1) The learning benefit meets or exceeds existing instructional methods.
- 2) The educational benefits are in some reasonable proportion to the cost of development, including the costs of human resources which are difficult to document.
- 3) Faculty receive adequate reward and recognition for the considerable commitment required to develop the programs.

We must discourage faculty from spending excessive time developing incomplete or ineffective CAI. Failure to direct attention to this problem may result in future generations accusing our generation of wasting precious human resources developing mere video games.

At the same time, we must encourage faculty to make a commitment to develop unique and improved educational tools which are useful throughout the profession. Incentives for faculty to expend the effort required to develop new instructional materials will necessitate a change in some traditional attitudes towards what constitutes scholarly endeavor.

This paper presents the need for evaluation and proffers possible mechanisms for student and peer evaluation.

Evaluation: Educators traditionally have not

conducted formal evaluation of changes they make in their instructional methodology or content. Usually, changes are simply made and put to use. That approach will not suffice with the expensive new technologies. Responsible administrators of limited resources will require evidence of quality and effectiveness. Likewise, other schools will not purchase expensive lessons unless they have some evidence that the lessons are at least as effective and efficient as those they will replace.

Evaluation may range from simple opinion surveys, internal evaluations of effectiveness and efficiency using randomly selected students, and peer review by external specialists. External peer review is the evaluation method most likely to affect faculty rewards for instructional developments.

A number of evaluation instruments have been developed for use in various types of media productions. Most developers of CAI have developed some type of internal evaluation form for use primarily to improve their own programs. For instance, Appendix 1 illustrates the evaluation form provided to users of the cardiovascular laboratory videodisc simulation. The form was designed to give feedback for use in changing the lessons and the authoring program used to write them. The first questions of the survey gather information on the user (affiliation, job function, and extent of use of the program). Then a five-item Likert type survey asks for responses to 44 statements grouped into categories of ease of use, quality, effectiveness, general suggestions, suggested applications, and general attitudes. A disadvantage of such a survey for evaluation purposes is its length.

Other forms have been developed for more generalized use. For example, Figure 2 illustrates a general media production evaluation form which was developed by Dr. Linda Johnson (Johnson, 1990) as a part of a doctoral research project on the evaluation of audiovisual media productions in veterinary medical education. The project also included an extensive review of evaluation methodology. Dr.

Johnson's survey includes 31 items grouped into categories of content, methodology, and technology. It is designed to be used in formal review, including peer review, of media productions. Dr. Johnson used a three round modified Delphi survey to develop the instrument with participation of 118 faculty from 20 schools and other institutions. Suggested criteria and questions in each of the three categories were solicited from the faculty members who then ranked the questions to be included in the final instrument. Such a survey instrument would be appropriate for many types of media evaluations. Selected questions could be added if information related to individual productions were needed.

Surveys can provide valuable feedback to the developers of lessons, and can give an indication of their quality. Their weakness is that they do not give objective comparative data. For comparative data it would be preferable to use a controlled experiment to measure the performance, efficiency, cost-effectiveness, and attitudes of those students who use a new method compared to those who attend a traditional lecture or laboratory. Unfortunately, many productions are not amenable to such a controlled study, and the use of a well-designed survey is the best instrument available.

Our interactive videodisc project is designed to develop alternatives to specific laboratories and lecture segments. An important reason for selecting such projects was the possibility of making controlled studies of effectiveness and attitudes using randomly selected students.

In three separate studies (Branch et al., 1987; Branch et al., 1988; and Fawver et al., 1990), a heart sound videodisc lesson, a preliminary cardiovascular laboratory lesson, and the final cardiovascular laboratory lesson were compared to existing methods using two groups of randomly selected veterinary medical students. For the final cardiovascular lesson a third group of students was originally planned; that group would have participated in neither the videodisc nor the traditional instruction--engaging instead in self-study. The third group was eliminated because of administrative concerns.

Initially there was some concern that students might feel disadvantaged by being placed into one or the other of the groups. Because of those concerns, the purpose of the evaluation was carefully explained to the students before the random groups were assigned. Furthermore, students were given the option of withdrawing from the study after

assignment. Finally, in the first study (Branch et al., 1987), questions were added to the survey to measure the students' attitudes to the evaluation itself.

The cooperation of the students in all three studies was remarkable. Only two students out of 270 expressed any dissatisfaction. Those students were offered the opportunity to withdraw from the study, but after the purpose of the study was explained again, they elected to participate. The responses to questions on the evaluation itself indicated that students would prefer that faculty simply adopt new programs, but the students strongly agreed that the evaluation was fair to all students.

In both the heart sound laboratory and the cardiovascular laboratory, students who used the videodisc lesson performed at least as well on a post-test, but the videodisc groups required significantly less time to cover the material. In addition, substantially less faculty and staff time was required to prepare and conduct the laboratory. The role of the participating faculty also was different. They spent most of their time answering questions on physiological principles instead of helping resolve technical procedural problems.

Peer Review: Peer review is needed to help faculty compare and select the most appropriate instructional software to purchase for use in their courses. Furthermore, peer review is a necessary ingredient in developing the faculty rewards that will encourage the best educators, especially younger faculty, to become involved.

Unfortunately, the problem of developing peer review mechanisms has seen little progress. Several organizations either have some type of peer review mechanism or have studies underway to establish a system. The AVMA for instance, did establish a peer review certification program, but that program did not extend to multimedia lessons using the new technologies. The Society of Applied Learning Technology formed a committee in 1988 to investigate peer review, but no report has been issued to date. EDUCOM also has made plans to study peer review.

The Health Sciences Consortium (Adcox, 1990) has some 20 years experience developing and marketing media productions, primarily videotapes. Using a review board of approximately 2000 members, that organization conducts peer review of videotapes before selecting them for inclusion in their catalog. That organization has only recently addressed IVD and CAI.

A peer review system for media productions must be understood as different from peer reviews for refereed research journals. Such journals have editorial boards which select the most appropriate reviewers for specific manuscripts. Investigators have incentive to serve as reviewers because of the opportunity to see advanced work in their field as well as the prestige associated with serving on review panels and editorial boards for major journals. Relatively little time is required to review a manuscript in one's field.

In contrast, peer review of media productions is more demanding in its time commitment and is less prestigious than manuscript review. Media production review requires expertise not only in subject content, but also in instructional methodology, and technical quality. While it may not require extensive time to review a 30 minute videotape production, review of a complex CAI or IVD production may require a day or more. Furthermore, there is little incentive for busy faculty to spend the time to provide such a peer review, since the activity may not weigh heavily in the faculty reward systems of most schools.

The very interactivity which is the essence of such productions requires the reviewer to examine many branching possibilities in the lesson. It is likely that a complex lesson will include a wide range of subject matter content, which may require that several reviewers participate in a single review to judge accuracy of content, technical quality, and methodology.

A critical difference between manuscript review and media production review is the problem of installing the production on an appropriate hardware system. This can be a problem even when the lesson is written for one of the commonly available hardware platforms; it can be exceedingly difficult for lessons requiring specific hardware.

With difficulties such as these, it is little wonder that effective peer review for CAI, IVD, and other complex media productions has yet to be perfected.

Proposed Peer Review System: It is proposed that a peer review system be developed for veterinary medical education to be administered by some existing organization such as the AVMA, CONVINCENCE, or some other organization which has the interest as well as the infrastructure necessary to develop and operate such a program. Grant support would be sought for initial development, but the final

system would be self-supporting either by fees from developers or by marketing of approved productions.

A permanent team of advisors in the areas of technical quality and educational methodology would be retained. Teams of subject matter specialists would be assembled as needed for specific productions. The reviews themselves would be conducted annually in conjunction with a national meeting such as the AVMA meeting. That meeting would have the advantages of an existing learning resources center with technical support for installation of the lessons, as well as potential reviewers already in attendance at the meeting. Reasonable honoraria could be provided to defray personal expense and compensate effort.

Such a review would ensure quality, cut costs, and provide an incentive of reward. It would be recognized by academia as critical acclaim. Standards of education would improve, and the economics of scale would justify the dedication of resources.

References

- Adcox, J: Personal communication, 1990.
- Branch C.E., Ledford B., Robertson B.T., and Robinson L. (1987). The validation of an interactive video disc as an alternative to traditional techniques: Auscultation of the heart. *Ed Tech* 27(3): 16-22.
- Branch C.E., Slimp M., Robertson B.T., Beckett S.D., Lewis M.M., and Fawver A.L. (1989). Interactive video simulated physiology laboratories. *Computers in Life Sci Educ* 6(11).
- Fawver A.L., Branch C.E., Trentham L., Robertson B.T., and Beckett S.D. (1990). A comparison of interactive videodisc instruction with live animal laboratories. *Am J Physiol; Advances in Physiol Educ*, in press.
- Johnson L. (1990). Personal communication.

APPENDIX 1
SURVEY ON VIDEODISC CARDIOVASCULAR LABORATORY

To improve future versions of this videodisc lesson and other lessons, we need your comments and suggestions. Please fill out this form, and use the attached page to add any specific comments or suggestions.

Please check the category in each column which most closely applies.

Your Affiliation:

- Medical school
- Veterinary medical school
- Other school _____
- Medically related Business
- Computer / video related business
- Other business or industry
- Government agency
- Foundation
- Research lab
- Private practice
- Other _____

Your Job Function

- Basic science instructor
- Clinical science instructor
- Research scientist
- Business or industrial training
- Professional school student
- Graduate student
- Intern, resident, clinical fellow
- Undergraduate student
- Media specialist
- Administration
- Other _____

Check one of the following 4 statements.

- I observed the program being demonstrated
- I used the program briefly (less than 15 minutes)
- I used the program more extensively (15 minutes to 1 hour)
- I used the program extensively (more than 1 hour)

D Please rate the following questions on a scale of 1 to 5

i 1 = Strongly Disagree

s 5 = Strongly Agree

a A Omit questions which do not apply to you.

g g
r r
e e
e e

Ease of Use:

- 1 2 3 4 5 1. Overall, the system was easy to use.
- 1 2 3 4 5 2. I was able to find the sections I wanted.
- 1 2 3 4 5 3. The section on instructions was useful.
- 1 2 3 4 5 4. I was able to use the program without using the instructions.
- 1 2 3 4 5 5. The progress indicators (asterisks) were useful.
- 1 2 3 4 5 6. The automatic (default) selections were helpful.
- 1 2 3 4 5 7. The TOGGLE (scratch list) function was useful.
- 1 2 3 4 5 8. I often got lost in the program.

Quality:

- 1 2 3 4 5 1. The content (subject matter) was acceptable and accurate.
- 1 2 3 4 5 2. The video quality was acceptable.
- 1 2 3 4 5 3. The quality of the animations was acceptable.
- 1 2 3 4 5 4. The narration was accurate and helpful.
- 1 2 3 4 5 5. The technical quality of the narration was acceptable.
- 1 2 3 4 5 6. The overall technical production quality was acceptable.
- 1 2 3 4 5 7. The material contained important inaccuracies. (Please identify important errors on attached sheet.)

Effectiveness:

- 1 2 3 4 5 1. The fact that the lesson covered more examples and variations than possible in a real lab was a desirable feature.
- 1 2 3 4 5 2. The lesson covered too many variations.
- 1 2 3 4 5 3. Being able to stop and review, in contrast to a real lab, is a desirable feature.
- 1 2 3 4 5 4. I have a better understanding of the subject matter from using this videodisc lesson.
- 1 2 3 4 5 5. The program held my interest.
- 1 2 3 4 5 6. The sequences were realistic.
- 1 2 3 4 5 7. There was too much material.
- 1 2 3 4 5 8. There was too little material.
- 1 2 3 4 5 9. The use of "we" helped make the material more realistic. (The program used words such as "we" and "let's" instead of "you will be shown..." to make it appear that the user was in the lab, not just passively watching a video.)
- 1 2 3 4 5 10. The use of "I" to refer to the computer itself was annoying.

General Suggestions:

- 1 2 3 4 5 1. Include more examples and variations of each experiment.
- 1 2 3 4 5 2. Include fewer examples and variations of each experiment.
- 1 2 3 4 5 3. Include additional experiments.
- 1 2 3 4 5 4. Make program less flexible (force student to go through prescribed sequences).
- 1 2 3 4 5 5. Make program keep records, so you could start where you left off.
- 1 2 3 4 5 6. Add practice self-testing.
- 1 2 3 4 5 7. Add actual testing for grading.

Suggested Applications:

- 1 2 3 4 5 1. Replace current labs on this subject matter for all students.
- 1 2 3 4 5 2. Use as optional replacement for those students who desire.
- 1 2 3 4 5 3. Make available in library for review,
- 1 2 3 4 5 4. Use as a visual aid in lectures.

General Attitudes:

- 1 2 3 4 5 1. It is desirable to eliminate the use of live animals in teaching when alternatives are available.
- 1 2 3 4 5 2. It is desirable to eliminate the use of live animals in teaching, whether or not alternatives are available.
- 1 2 3 4 5 3. No matter how good a video / computer lesson is, it could not fulfill all the important objectives for animal labs in teaching basic sciences.
- 1 2 3 4 5 4. Before I would replace a real lab with a computer or video alternative, I would need to see objective student performance data that the simulated lab is as effective as the real lab.
- 1 2 3 4 5 5. I could make a decision on replacing a real lab with an alternative by just viewing the material, much as I would select a new textbook.
- 1 2 3 4 5 6. This videodisc lesson covered material which is relevant to biomedical teaching.
- 1 2 3 4 5 7. This videodisc lesson is something I would consider using in my teaching.
- 1 2 3 4 5 8. This concept in general, not necessarily this specific subject matter, is something I would consider using in my teaching.

COMMENTS

List any subject matter errors or problems you had with the lesson. Please suggest any other simulated laboratories you would like to be available.

INSTRUCTIONAL MEDIA PRODUCTION EVALUATION CRITERIA

Classifications

1. Needs revision
2. Acceptable
3. Excellent, no changes needed

CONTENT

- | | |
|---|-------|
| 1. Is the information presented accurate, clear and thorough? | 1 2 3 |
| 2. Is the material current, up-to-date information | 1 2 3 |
| 3. Have the stated objectives been met? | 1 2 3 |
| 4. Are illustrations pertinent? | 1 2 3 |
| 5. Is the title an adequate description of the program? | 1 2 3 |
| 6. Does the content arouse and maintain interest? | 1 2 3 |
| 7. Are major ideas summarized? | 1 2 3 |
| 8. Were the objectives reasonable and the information appropriate for the intended audience | 1 2 3 |
| 9. Can the material be easily update? | 1 2 3 |
| 10. Does the material make an important contribution to veterinary medical education? | 1 2 3 |
| 11. Is the producer/author qualified as a subject specialist? | 1 2 3 |

METHODOLOGY

- | | |
|---|-------|
| 1. Are the objectives clearly stated? | 1 2 3 |
| 2. Is the information logically arranged and organized in an easy-to-follow format? | 1 2 3 |
| 3. Is there continuity between the visual and verbal/written? | 1 2 3 |
| 4. Are illustrations/examples given? | 1 2 3 |
| 5. Is a summary presented? | 1 2 3 |
| 6. Is the message clear and uncluttered (no irrelevant/distracting elements)? | 1 2 3 |
| 7. Does the production give a realistic picture of the ideas presented? | 1 2 3 |
| 8. Are directions given for the use of the program? | 1 2 3 |
| 9. Does the viewing time allow adequate development of the intended production concept without making it boring | 1 2 3 |
| 10. Is creativity used in presenting the material? | 1 2 3 |

TECHNOLOGY

- | | |
|---|-------|
| 1. Is there legibility of words, charts, graphs, captions? Is detailed information clearly visible and designed for easy comprehension? | 1 2 3 |
| 2. Is the narration fluent and easy to understand? | 1 2 3 |
| 3. Is the production in focus? Do the exposures used ensure correct visibility? | 1 2 3 |
| 4. Is there synchronization of sound and image? If multimedia is incorporated, are they in synchronization? | 1 2 3 |
| 5. Is the production free from background "noise" that is obtrusive and interferes with the narrative or visuals? | 1 2 3 |
| 6. Is the information presented at an appropriate pace? | 1 2 3 |
| 7. Is the medium appropriate to the message it seeks to communicate? | 1 2 3 |
| 8. Has the production been satisfactorily edited (cuts, dissolves, continuity)? | 1 2 3 |
| 9. Is there a comfortable sequencing of information? | 1 2 3 |
| 10. Does the production have aesthetic appeal? | 1 2 3 |