

Session # 1331

A Model for Teaching Materials Evaluation: Development and Testing of Interactive Computer Simulations Modules for Undergraduate Education

**Anne E. Donnelly¹, Emilia Hodge¹, Melis Budak¹, Heath Wintz²,
Randy Switt², Chang-Yu Wu², Prakash Kumar³, Pratim Biswas³
Priscilla Chapman⁴, Anne L. Allen⁴**

¹University of Florida, Engineering Research Center for Particle Science & Technology, Gainesville, FL 32611/²University of Florida, Department of Environmental Engineering Sciences, Gainesville, FL 32611/³Washington University in St. Louis, Environmental Engineering Science Program, St. Louis, MO 63130/⁴University of Florida, Office of Academic Technology, Gainesville, FL 32611

Abstract

A comprehensive evaluation program was developed as part of an NSF Course Curriculum, and Laboratory Improvement grant to develop three undergraduate computer simulation modules. Aerosol science and technology is generally taught at the graduate level and the goal of this program was to develop materials that would bring this subject to the undergraduate level. To achieve this goal, an evaluation plan was developed that included formative and summative, and cognitive and affective measures. This was a collaboration between content and evaluation experts who were involved from the proposal stage. The program included formative evaluation of pilot versions of the materials. Both undergraduate students and faculty were involved and provided significant feedback on how the materials could be modified to be more effective. The modified versions were formally tested in classroom settings to determine if students could master the material and if they enjoyed using the modules. Students made statistically significant gains in knowledge as a result of the modules and appreciated the ability to go through the simulations at their own pace. The evaluation program used here was instrumental in ensuring high quality products and can serve as a model easily exportable to other educational product development projects. The full model and the lessons learned will be described.

Introduction

Evaluation of engineering education programs and products is critical to ensure quality and enhance the dissemination of these materials. The National Science Foundation has taken a leadership role in ensuring adequate program evaluation by making it an integral element of

proposals and by publishing several evaluation guides.^{1,2} Education program and product evaluation is a mature discipline with a variety of best practices and procedures already developed. However, it traditionally has been studied and taught in Colleges of Education. The increased emphasis on education evaluation in engineering programs has placed evaluation responsibilities on engineering professors and researchers who may not have had any formal training in this area. As a result, evaluation programs conducted in support of NSF proposals have been of mixed quality.

There are several common mistakes that hamper evaluation efforts. Evaluation programs that are added on at the end of the proposal writing process; that are under-funded; that do not include a variety of potential stakeholders; and that are not conducted with the same degree of rigor that would be required of any research effort, will not produce the best results.

Methods for Developing an Evaluation Plan

The steps below were used in the development of the Aerosols project evaluation plan and can serve as a guide to developing an evaluation program for any type of educational materials. How each step was used in the development of the Aerosols modules is included.

Step One: Identify and include the evaluator at the proposal writing stage. Collaborating with someone with experience in evaluation is a good first step. It is important to involve the evaluator at the earliest stages of the proposal process.¹ This will help clarify the goals of the project, the expected outcomes, and result in an evaluation plan that will meet the program developers' needs. Education evaluators can be located in Colleges of Education or Psychology or in other education resources on campus.

For this project an education program evaluation expert who was located in the UF NSF Particle Engineering Research Center was recruited at an early stage in the proposal process. Many campuses have large NSF Centers such as ERC's, Science and Technology Centers (STC's) and Materials Research and Science Engineering Centers (MRSECs) where there are Education Administrators who have expertise in evaluation.

Step Two: Identify stakeholders that should be included in the evaluation process. In addition to the obvious group involved in a study (for example, if you are developing materials to teach freshman, freshman students would be tested) there are often other groups that should be included.

Since the goal of this project was to develop educational materials for undergraduates, it was clear that undergraduate education students would be part of the evaluation process. In addition however, potential faculty users of the finished products were an important group to include, as the acceptance of the materials by them would impact how well the materials were disseminated.

Step Three: Identify what question(s) the evaluation will seek to answer. There are multiple purposes for the evaluation (to ensure the products are effective, to enhance dissemination of the materials, to report to the funding agency, etc.) and each of these purposes should be considered when determining what questions the evaluation program should address.

The purpose of the Aerosols evaluation program was to determine if the modules were successful in teaching the selected content and to determine how well accepted the modules would be to undergraduate students and faculty. In other words, the evaluators wanted to answer the questions: (a) Were students able to learn aerosols content from the modules? (b) Could students navigate the online modules with ease? (c) Was the format of the modules acceptable to students and easy to understand? and (d) How did faculty rate the content and usefulness of the modules?

Step Four: Identifying the type of evaluation that will be used. There are two types of evaluation that a comprehensive evaluation program should include, which produce two types of information for the program developers.^{2,3} Formative evaluation is conducted as the materials are being developed and can avoid costly errors by providing feedback to the developers as the materials are being formulated. After the materials have been improved based on information collected during development, summative evaluation can determine the degree to which the finished product meets the goals established in the beginning of the project. Both formative and summative evaluation was built into the Aerosols evaluation.

Step Five: Determining what type of data is required. In educational materials development, it is necessary to collect cognitive data to determine if the materials were effective at teaching the intended content. This can be achieved through a test designed by the content experts. Additionally, it may be useful to collect affective information that investigates how the participants enjoyed using the materials, and how they perceived the usefulness of them.

It was determined that five types of information would be collected for the Aerosols project. Since these were computer simulations, it was important to determine if students found the format understandable and the program easy to navigate. We were also interested in how they enjoyed using the modules. Of critical importance was if the modules effectively taught the specified content. Finally, data from potential faculty users with respect to both content and format were also desired.

Step Six: Select the experimental design for the evaluation: Quasi-experimental and experimental designs can be used and often the ability to randomize the sample will dictate which you use. If randomization of subjects can be achieved, an experimental design can be used. In many cases, existing classes of students are used to test materials, and this will dictate the use of a quasi-experimental design. Additionally, it must be determined if a posttest only, pre and posttest, posttest only, or some other design will be utilized. A pre-post test design is best when looking for learning gains as it allows you to control for prior knowledge of the students.

A pre-post test, quasi-experimental design protocol was developed for the Aerosols project. The use of a pretest allowed evaluators to establish the level of knowledge of the content that the student possessed prior to using the module.⁴ While it was expected that it was unlikely that these students would have prior instruction in this area, it was important to test this assumption. The posttest following use of the module then allows the determination if the students mastered the content as a result of using the material. The selected design was quasi-experimental as there was no randomization of the participants, rather a convenient sample was employed – those

students who were in existing undergraduate Environmental Engineering classes at UF and WUSL.

Step Seven: Design of the evaluation program instruments. Written tests and surveys are often used to collect both content and affective data, but there is a wide variety of choices when selecting instruments. A discussion of some of the types of instruments that can be used is found in Aiken (1997).⁵

Two instruments to be completed by student participants were developed to evaluate the usefulness of the Aerosols modules. The first questionnaire (Table 1) collected the affective data for all three modules by asking for participants' evaluation of the navigability of website, format of the modules, and this mode of learning. This questionnaire was composed of 12 questions. Eight of the questions asked students to provide written comments, while the other four asked them to indicate on a scale of 1 (weak agreement) to 5 (strong agreement), their level of agreement with the different statements regarding the website and this mode of learning.

A written test was developed for each module to collect cognitive data and was used as a pre and posttest. These tests related to the content of the module, therefore evaluating the usefulness of the module to teach the actual aerosol content. They were developed by the content experts and consisted of open-ended questions. The data were analyzed using Statistical Package for the Social Sciences (SPSS). Descriptive statistics were computed, a one-sample *t*-test, as well as simple regression analysis were conducted at the $\alpha = .05$ level of significance.

A third questionnaire for faculty was developed to collect data from potential faculty users after their review of the three modules. The questionnaire consisted of 11 open ended questions because it was important to find out the perceptions of the usefulness of the modules from this group. Questions focused on finding out what this group liked most and least about the modules; whether they would use this type of curriculum material in their courses; and how they would improve the modules. The responses were organized under categories by question and displayed so that developers were able to incorporate the feedback into their revision of the modules.

Step Eight: Implementation

Formative Evaluation:

Initial formative evaluation of the modules involved (a) the primary evaluator who conducted the first test of the material; identifying problem areas; and providing feedback to the programmers and suggesting modifications to format, and (b) a small number of engineering undergraduate students who tested the materials and provided additional feedback for the software designers.

Following suggested modifications, more formal formative evaluation of the modules was carried out by students in the classroom environment using a pre-post test design; as well as by potential faculty users. Students were administered the pretest. They reviewed the module and

Table 1. *Questions for Review of Web Site and Format of the Aerosol Modules*

1. Did you run into any part of the website that didn't work? If yes, explain which part and the problem you encountered.
2. Did the site load in a reasonable length of time?
3. Was there anything on the website that caused you any confusion?
4. Would you have preferred a paper and pencil version of this website or did you feel this website was an effective means of learning? Explain your answer.
5. What did you like most about this web site?
6. What did you like least about this web site?
7. How long did it take you to complete the module?

Please indicate your level of agreement with the following statements by rating the items from 1 to 5 where 1= weak agreement and 5=strong agreement.

8. I liked this form of learning

Weak agreement					Strong agreement
1	2	3	4		5

9. I thought I learned more from this website than from an actual lecture.

Weak agreement					Strong agreement
1	2	3	4		5

10. I liked the idea that I could complete the module on my own time.

Weak agreement					Strong agreement
1	2	3	4		5

11. I liked the idea that I could complete the module at my own pace.

Weak agreement					Strong agreement
1	2	3	4		5

12. If you answered 1 or 2 for questions 10-11, please explain your answer below.

were given as much time as they wanted to complete the material. When they finished the module they then completed the posttest as well as the format survey. The protocol was as follows: the principal investigator administered the pretest to the group of students in the classroom and organized a schedule for when participants would report to the PERC (Particle Engineering Research Center) to use the computers to review the module and take the posttest. For this study students were allowed to take notes as they completed the module, but were not given access to the notes or the module while completing the posttest. This was done in an effort to control the conditions under which all the participants reviewed the module and took the posttest. Participants worked at individual computer stations and were allowed as much time as needed to work through the modules. After completing the online module review, they completed the posttest in another room. Two graduate students from the university's College of Education who were trained to move participants through the evaluation process facilitated the procedures.

To reach a broad cross-section of potential faculty users, the modules were available for testing and surveys were provided to participants at the *2002 International Aerosols Conference in Taipei* and the *2002 Annual American Association for Aerosol Research*. Forty-eight individuals examined the on-line interactive modules and responded to questions about the usefulness of the materials. They were also asked to make suggestions on how to improve the modules. Input from this group added significant content expert suggestions as well as additional formatting ideas.

Summative evaluation of this module involved undergraduate students enrolled in Environmental Engineering classes at UF and UWSL. The same procedures described for the formative evaluation were followed. A complete report of the evaluation results of all three modules developed as part of this CCLI is currently in preparation. Preliminary results for the *Cyclone* module are below.

Cyclone Results

Initial formative evaluation of the modules focused on navigability of the module, clarity, and ease of understanding, and provided feedback from the primary evaluator and four undergraduate engineering students. At this level of evaluation, significant improvements can be made from feedback from small numbers of participants. Significant revisions were suggested to the designers to help improve the materials.

Some examples of the type of feedback on the format that was provided to the content experts and designers are shown below. Participants' responses to the first question (Was there anything on the website that caused you any confusion?) included such answers as:

- the equations page seemed to overwhelming, cluttered with lots of equations, better if made a separate page for each equation
- unexplained chart (no info on trends)
- no sample calculations
- some of the formulas needed to be explained better
- yes, animations should have been clearer

Following these small-scale tests the designers made changes to improve the format as suggested by these participants.

Formative evaluation of the module next involved a class of 13 undergraduate students. The first questionnaire collected the affective data by asking for participants' evaluation of the navigability of website and this mode of learning.

The data from the questionnaire, which tested student's understanding of the content of the module, were analyzed using Statistical Package for the Social Sciences (SPSS). For every participant, the means of the posttest scores were higher than the means of the pretest scores. The standard deviations of the posttest scores were lower than those of the pretest scores in all instances except in only one instance when the figures were the same. The results of the one-sample *t*-test showed a statistically significant difference between the means of the pre and posttest scores, $t(12) = 4.731, p < .05$. However, this sample consisted of only thirteen participants and a sample of over 30 would provide a more powerful test.⁶ Nevertheless, these results showed that, at this formative stage, there are differences between the pre and posttest scores and that students did learn some content from the module. This initial evaluation effort indicated that the students could learn Aerosols content from the *Cyclone* Module. Again, feedback was provided to the module developers for revision of the module.

Formative results from 48 University Professors showed that ninety-two percent of the respondents indicated that the *Cyclone* module would be a useful tool in teaching particle transport related concepts, or topics related to inertial behavior of particles. The responses showed that it was clear the modules were very well received by this important group.

Of even more help to project managers were the comments on how the modules could be improved to make them more effective. Some of these comments were:

- It seems to include a lot of the “what” and “how” but almost one of the “why” – for example, in the cyclone, the equations are there, tables of parameters are there, but no bringing it all together, or explaining trends, reasons, etc.
- Not enough detail on inputs, at least at first glance
- The only complaint is that in the instructional part, the animations require scrolling on a typical screen. Maybe the animations can be reduced in size.
- Error conditions should be flashing or more noticeable immediately.
- It would be nice to have links for each item selected to give an explanation of the choice to be made.
- Tell when outside of range

There were also many detailed suggestions with respect to specific module features provided. This information was provided to the web designers and faculty to allow them to make improvements in the modules. After these modifications and improvements were made, formal testing of the cyclone modules with students was conducted.

Summative evaluation of the modules involved engineering students at UF. With regard to student's ability to learn the actual aerosols content from the module, the descriptive statistics revealed that for each participant, the means of the posttest scores were higher than the means of the pretest scores. The results of the one-sample t -test showed a statistically significant difference between the means of the pre and posttest scores, $t(17) = 5.182, p = 0.00$. This confirmed the findings of the formative evaluation that the module successfully taught the desired concepts.

Other Considerations

Institutional Review Board Requirements

For this study and any others that are conducted with the goal of publishing evaluation results, it is necessary to consult the campus Institutional Review Board (IRB) to ascertain what the local requirements are. To conduct the evaluation program described above it was necessary to obtain IRB approval that consisted of supplying the IRB office with a completed protocol, including all testing instruments and participants' letter of informed consent. Per UF IRB requirements, the data collected for the evaluation was collected by the evaluator (not the classroom professor), was confidential, and did not affect their grades. As each institution has regulations specific to each campus, it is necessary to consult with the local IRB office.

Human and Budgetary Requirements

The total evaluation program was budgeted at 10 % of the grant. In addition to one-month salary of the Evaluator, two College of Education Graduate students were employed on an hourly basis to assist in the collection of data and to collate and complete statistical analysis of the data. It is estimated that evaluation of each module required approximately 100 hours of effort.

Discussion

Evaluation of educational products ensures that the materials designers achieve their goals. The steps outlined in this paper that were used to evaluate the Aerosols modules resulted in engineering education materials that were well received by both student and faculty users and that effectively taught the desired content. The development of these modules was greatly enhanced by the evaluation component built into the project.

This collaboration between engineering faculty and education evaluators required significant discussions through which the team came to an understanding of the different languages and area of expertise they each represented. There were important lessons learned in several areas.

The first lesson is that there is a need for initial conversations to come to a consensus on what questions are appropriate for the evaluation to address. For example, the Aerosols researchers initially believed that the first step in evaluation would be to compare the use of the module in one class to a class in which the module was not used. Through additional discussions, the team clarified that the more appropriate first question to ask was if the modules were effective at conveyed the desired content. Clarifying the questions then led to the development of the evaluation design.

Secondly, Engineers and Education Evaluators have different understandings of education evaluation. Not only should a protocol be developed, but also the reasons for the protocol selected and the constraints imposed by the protocol must be made clear. There are a variety of threats to the validity of an education study that may not be transparent to all, and if these are not clearly articulated ahead of time, researchers may inadvertently break protocol, resulting in a loss of data. Once again, thorough discussions between the team members will resolve this issue.

A third lesson involves the use of measurable objectives. Not everyone has a clear understanding of the need for measurable objectives in an evaluation study or how to write this type of learning objective. It is important to help clarify how a given study will be more useful if objectives for the materials are phrased in measurable terms that can then be incorporated into the instruments. The evaluators worked with the Aerosols content experts to develop this type of objectives for each module.

It is without a doubt that the final products produced in the Aerosols project were superior due to the incorporation of multiple suggestions and comments collected from both students and faculty as part of the evaluation process. The results of this evaluation program suggest that the following recommendations are key features to be included when developing engineering education teaching materials:

1. An evaluation expert should be included in the project team from the beginning
2. An adequate budget (about 10% of the total budget) to support the evaluation effort is recommended
3. The project team must work together to develop the goals of the materials that ultimately become part of the pre and post test instruments
4. Formative evaluation should be conducted with a variety of stakeholders
5. Feedback derived from the formative evaluation instruments must be provided to the materials developers to allow for appropriate modifications
6. Summative evaluation of the materials should be conducted with appropriate stakeholder groups as identified in the initial evaluation steps.

A final lesson learned from this exercise was that even though one might use the term “summative” to describe an evaluation method, it is important to keep in mind that there is always room for improvement when developing education materials. Future plans are to provide the Aerosols modules to faculty across the country to collect additional data on how well they export to other institutions.

References

¹ Frechtling, J. (2002). *The 2002 User-Friendly Handbook for Project Evaluation*. NSF 99-12175. Arlington, VA: NSF.

*“Proceedings of the 2004 American Society for Engineering Education Annual Conference & Exposition
Copyright © 2004, American Society for Engineering Education”*

² Frechtling, J., & Sharp, L. (1999). *The User-Friendly Handbook for Mixed-Method Evaluations*. NSF 97-153. Arlington, VA: NSF.

³ Morse, J. M. (2003). Principles of mixed methods and mulitmethod research design. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research* (pp. 189-208). Thousand Oaks, CA: Sage Publications.

⁴ Smith, M. L., & Glass, G. V. (1987). *Research and Evaluation in Education and the Social Sciences*. Englewood Cliffs, NJ: Prentice Hall Incorporated.

⁵ Aiken, L. R. (1997). *Questionnaires and Inventories*. New York: John Wiley and Sons Incorporated.

⁶ Norusis, M. J. (1999). *SPSS 9.0: Guide to data analysis*. Upper Saddle River, NJ: Prentice Hall.

Author Biographies

Dr. ANNE E. DONNELLY is Associate Director for Education and Outreach at the NSF Particle Engineering Research Center (PERC) at the University of Florida. She manages the ERC educational programs including course development and evaluation activities, graduate student activities, and the Undergraduate Research Scholarship Program. She also consults with the ERC Industrial Partners Program on particle science and technology issues of mutual concern.

Dr. EMILIA E. HODGE is Education Research Consultant at the Particle Engineering Research Center, University of Florida. She earned her doctorate in Educational Leadership from the University of Florida and has conducted project and program evaluation of NSF funded grants involving graduate and undergraduate students and faculty in Science and Engineering.

Ms. MELIS BUDAK is a graduate assistant at the Particle Engineering Research Center, University of Florida. She is currently a graduate student in Special Education scheduled to complete her Master's degree in May 2004.

Mr. HEATH WINTZ is an undergraduate assistant in the Department of Environmental Engineering Sciences at University of Florida.

Mr. RANDY SWITT is the IT coordinator in the Department of Environmental Engineering Sciences at University of Florida.

Dr. CHANG YU-WU is an assistant professor in the Department of Environmental Engineering Sciences at University of Florida. His research and teaching interest focus on air pollution control and aerosol science. He has authored/co-authored 21 refereed journal articles and given 85 presentations on these subjects.

Mr. PRAKASH KUMAR is currently pursuing D.Sc. working on transport, capture, and synthesis of ferromagnetic aerosols in the Department of Chemical Engineering, Washington University. He received M.S. from Washington University in 2003 and has been nominated as a web page designer for Association of Graduate Engineering Students at this University.

Dr. PRATIM BISWAS is the Stifel and Quinette Jens Professor of Environmental Engineering Science and Director of the Environmental Engineering Science Program at the Aerosol and Air Quality Research Laboratory at Washington University in St. Louis. His major research interest focuses on the examination of particle formation and growth dynamics in high temperature environments. He has published extensively in his field and has made several presentations at National and International Conferences, and has been invited to lecture worldwide.

Ms. PRISCILLA CHAPMAN is Program Assistant in the Office of Academic Technology at the University of Florida.

Ms. ANNE L. ALLEN is Manager/Instructional Designer in the Office of Academic Technology at the University of Florida.

Acknowledgements

This work is made possible by an NSF Course, Curriculum, and Laboratory Improvement (CCLI) grant, Award # DUE 0127429. Disclaimer: Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

The authors acknowledge the financial support of the Engineering Research Center (ERC) for Particle Science and Technology at the University of Florida, The National Science Foundation (NSF) (Grant EEC-94-02989), and the Industrial Partners of the ERC for support of this research.