Information and Environmental Education Session

Evaluating Environmental Education Programs

James B. Armstrong, *Department of Zoology and Wildlife Science, Auburn University, AL 36849*

Peter T. Bromley, *Department of Zoology, North Carolina State University, Raleigh, NC 27695-7646*

James C. Impara, *Buros Institute of Mental Measurements, University of Nebraska-Lincoln, NE 68588-0348*

Abstract: We use a case study to illustrate how challenging it can be to apply rigorous evaluation procedures of environmental education programs in a real world setting. We report here our experiences in conducting 2 major evaluations. In Case Number 1, the identity of the client is not disclosed for reasons that will become obvious. This client was a national, non-profit conservation organization that requested an independent evaluation of an educational program designed for volunteer use. Case Number 2 reviews the evaluation of the Virginia Hunter Education program conducted in the early 1980s.

Proc. Annu. Conf. Southeast. Assoc. Fish and Wildl. Agencies 49:647–655

Nearly all state wildlife agencies conduct environmental education programs. These programs have audiences that range from sportsmen to youth and vary in the form of presentation from outdoor camps to interdisciplinary classroom supplements (e.g., Project WILD, Project Learning Tree, Nature-Scope). In an era of increasing accountability, these programs should be evaluated as to format, content, and effectiveness. As observed by Stout and Peyton (1988), neither the agency nor the producer should assume program value even though each program may have been designed with the best of intentions and implemented to address a real need. Only through objective evaluation can the developer and the user have some measure of confidence that programs are worthy of continued use or modification.

Even though developers and/or promoters of environmental education programs may have a sincere interest in evaluation of their materials, they may

not be familiar with evaluation processes. Thus, professional evaluators are often employed. However, a lack of understanding of evaluation processes may result in communication breakdowns between the evaluator and the agency personnel. If the evaluation effort is to be successful, it is critically important to maintain open lines of communication. Typically, the evaluator educates the client as to what to expect during the evaluation process.

Objectivity and independence are as essential in program evaluation as they are in a scientific experiment. This is particularly true where agency administrators have called for the evaluation of programs produced and distributed by special interest groups. The evaluation must be protected from undue political pressure that would prevent arriving at clear judgments and recommendations about a program.

In this paper, we use a case study to illustrate how challenging it can be to apply rigorous evaluation procedures in a real world setting. Due to the sensitive nature of many of the observations, the identity of one client has not been disclosed.

Evaluation Theory and Management

Worthen and Sanders (1987) define evaluation as a process for determining worth or merit. One product of curriculum evaluation is the provision of information useful for revision of materials (Gronlund 1985). Frequently, selected components of programs are evaluated; e.g., specific outcomes, implementation, or planning (Patton 1980). These kinds of evaluations are formative in nature, as opposed to summative evaluations which result in "yes" or "no" decisions on the fate of entire programs.

The uses of evaluation results are dictated by client needs and influence evaluation strategies (Isaac and Michael 1985). Several authors (House 1980, Patton 1980, Isaac and Michael 1985, Worthen and Sanders 1987) have provided overviews of evaluation strategies. These strategies fall along a spectrum from the highly-structured utilitarian model, such as the objectives-oriented approach, to the less-structured naturalistic, and participant-oriented models. The methods by which data are gathered and analyzed should not be confused with models for managing evaluations. Evaluation management is much broader and addresses the following concerns: work scope, client responsibilities, contract provisions (legality of evaluation, confidentiality, publication rights, negotiation, arbitration), audience definition, instrument development, data analysis, and final reporting. Guidelines for managing evaluation have been proposed by Stufflebeam (1973), Joint Committee on Standards for Educational Evaluation (1981), Isaac and Michael (1985), Worthen and Sanders (1987), and Worthen and White (1987). Their recommendations differ very little from one another and can be adapted to most evaluations.

Internal evaluations are conducted by program personnel who may or may not have been involved in designing the program. External evaluations are con-

ducted by individuals not affiliated with the program to be evaluated. Both internal and external evaluators may have expertise in evaluation methodology. However, because internal evaluators may have a vested interest in the outcome of the evaluation, human bias may result in the selection of evaluation variables that indicate program success. While not foolproof, the use of external evaluators increases objectivity and may produce a much clearer picture of program strengths and weaknesses.

Social, economic, and political factors involved in program development and implementation dictate that evaluation management will be uniquely tailored to fit each program. The first priority of the evaluator is to establish and maintain excellent communication with the client, especially to identify and agree upon the end products of the process. It may take several sessions of thoughtful discussion and reiteration to reach this agreement. Evaluation management is dependent on the relationship between the evaluator and the different audiences associated with the program: (a) the client who initiates the evaluation and funds the study (e.g., the state fish and wildlife agency), (b) the school administrators who control access into the classrooms where interdisciplinary environmental education supplements are used, and (c) the classroom teachers who incorporate the materials into their lesson plans and, ultimately, provide the data. Interactions between each of these groups and the evaluator are highly variable and strongly influence the evaluation's success, cost, and validity.

We thank N. Holler, L. Stribling, D. Vogler and several anonymous reviewers for constructive reviews of this manuscript. This project was supported by Alabama Agricultural Experiment Station Project and published as Alabama Agricultural Experiment Station Journal Series 15-955018.

Methods

We report here our experiences in conducting 2 major evaluations. In Case Number 1, the identity of the client is not disclosed for reasons that will become obvious. This client was a national, non-profit conservation organization that desired an independent evaluation of an educational program designed for volunteer use. Case Number 2 reviews the evaluation of the Virginia Hunter Education program conducted in the early 1980s. The original evaluation management model for our case studies incorporated: (a) focusing the evaluation, (b) collecting information, (c) organizing information, (d) analyzing information, (e) reporting information, and (f) administering the evaluation (Stufflebeam 1973). Categories (a) through (e) relate to management of the evaluation, while (f) addresses evaluation outcomes. Our studies operated from Stufflebeam's (1973) evaluation that model, incorporating revisions and modifications where appropriate. Details of actual evaluation results are in Bromley and Hampton (1981), Hampton and Bromley (1982), Bromley et al. (1988, 1989), and Armstrong and Impara (1990, 1991).

Results and Discussion

Concerns were manifested in several issues which required resolving before the evaluation could proceed. We highlight many of these issues and the resulting resolutions.

Case Number 1

Several groups influenced the direction of the evaluation: the program developers (referred to as clients), school administrators, and teachers. It was the responsibility of the evaluation team to address the concerns of each of these groups in order to provide a useful evaluation of the program.

Client Issues.—Approximately 5 face-to-face meetings and 10 telephone consultations with the program developers (i.e., evaluation client) provided the evaluator with an internal assessment of the client and the client's political nature—factors which affected the use of evaluation results. As the evaluation progressed, hidden, politically-driven agendas, such as a desire to produce positive results at the expense of randomization, surfaced that required modification of the evaluation plan to meet the needs of the client.

Issue 1—Weighting of Objectives

Initially, the client agreed to evaluating several sets of program materials by testing student attitude and knowledge change. The number of times a learning objective was mentioned in the text was used to weight that objective on the evaluation instrument. This relative weighting was presented to the client for review. However, there was not a match between the amount of weight an objective received from the written materials and the priority of objectives as stated by the client. These discrepancies made it difficult to reach agreement on the evaluation instrument and threatened to invalidate the entire evaluation.

Issue 2—Design or Selection of Instruments

Selection and ultimate development of the attitude measure was influenced by the political nature of the client. The Environmental Awareness Inventory (EAI), developed by Passineau (1976), was originally selected as the measure of the program's effects on environmental attitudes. This instrument was selected due to its general environmental focus and acceptable reliability estimates. However, the client rejected the instrument on the basis that it did not measure attitudes in a way it felt was appropriate. The client consulted outside evaluators and solicited their support in rejecting the instrument. This was done despite approval of the instrument by evaluation experts on the proposal review team. To maintain rapport with the client and continue the evaluation, the EAI was rejected and a new environmental attitude scale developed.

Issue 3—Desired Evaluation Outcomes

Formal and informal meetings about the evaluation made it clear that the client was very concerned about the possibility of a negative evaluation. The client had clear opinions of the relative merits of each section of the materials and expressed concern that sections be selected which might produce positive

cognitive and affective results. This resulted in loss of a truly random selection process and necessitated modifications in the goals and methods of the evaluation. It should be noted with respect to Issues 2 and 3 that the instrument developed by the evaluation team still possessed the psychometric rigor to produce unbiased evaluation results (see Armstrong and Impara 1991).

Issue 4—Final Evaluation Plan

A final evaluation plan was developed and submitted to the client for approval. The plan indicated the general methods to be used in collecting data, expert review team to be employed, and a proposed timeline and budget. The timeline was flexible and allowed for the numerous modifications that were necessary due to evaluator and client delays.

Issue 5—Communication with the Client

Throughout the evaluation, the need for frequent and direct communication with the client was reinforced. Lack of communication resulted in a misunderstanding of the client's desired emphasis for the knowledge and attitude tests. This resulted in delays which threatened the timely completion of the evaluation.

Issue 6—Contract

The client/evaluator contract specified what would be provided by each party. The evaluation team agreed to manage the evaluation, collect and analyze data, and provide a final report. The client agreed to provide financial support for instrument development and testing, travel, and printing. In addition, the client agreed to provide free educational materials to teachers as an incentive to participate. The contract did not specify the liberties that the client could take with instrument modification and approval. This resulted in client/evaluator conflicts concerning the knowledge and attitude measures.

School Administrative Issues.—It was necessary to follow the "chain of command" in soliciting participation by teachers in the evaluation. This was done by contacting and working with school administrators (i.e., superintendents, principals) throughout the study area. This produced some difficulties, as well as some enlightenment regarding the nuances of working within the political structure of school administrations. Approximately 30 school administrators were contacted regarding participation in the study resulting in only 8 agreements. A final sample of 88 teachers from 2 states began initial participation in the evaluation.

Issue 7—Teacher Time Schedules

Many administrators refused to participate in the evaluation simply because it is easier to decline than try to cooperate. Others rejected the evaluation as a way of protecting their teachers from an increased workload, even though teachers were to be asked to volunteer for participation in the study. The evaluation was also rejected on the basis that the selected topics did not fit with the existing curriculum.

Issue 8—Cooperation of Administrators

By far, the greatest success was experienced with those administrators who were personal and professional acquaintances of the evaluation team members.

This raises the issue of the political aspect of producing successful evaluations within the school systems and reinforces the need for evaluators to work closely and cooperatively with educational professionals to develop a network of contacts.

Teacher Issues.—Once permission was gained through school administrators, the evaluation team began the process of contacting teachers to volunteer for the study. Once a list of volunteers was developed, workshops were initiated to distribute the materials.

Issue 9—Distribution of Materials

The evaluators worked closely with volunteer teachers at the beginning of the study to distribute materials and clarify the teachers' responsibilities. Most of the teachers volunteered to participate in an orientation meeting where materials were distributed. Some could not attend for legitimate reasons and requested that another teacher pick up their packets.

Issue 10—Subject Knowledge

Some teachers who attended the orientation meeting expressed concerns about their lack of knowledge about environmental topics to be evaluated. However, others seemed quite confident and had taught the subjects before. The evaluators served as a source of information on the topics and evaluation methodology prior to and during the study.

Issue 11—Incentives

The free materials offered by the client to participating teachers seemed to promote participation. All teachers who participated in the study completed and returned the form indicating which of the free materials they would like to receive. In addition, several teachers contacted the evaluator to inquire about the materials and when they might expect to receive them.

Case Number 2

In this evaluation, the Virginia Hunter Education Program was evaluated because the current program coordinator was about to retire and because standards for program length and content were being increased by the U.S. Fish and Wildlife Service through their Office of Federal Aid (Bromley and Hampton 1981).

Issue 1—Agency Expectations

Initially, agency administrators felt that a pre- and post-test of students exposed to the program combined with an undercover study of deer hunter behavior would suffice. It was agreed that the knowledge gain would be measured by students in the usual manner, but dove hunters were substituted for deer hunters. It was felt that the larger sample size afforded by hunters at public dove fields and their increased observability over deer hunters made dove hunters a better subset of the hunting public to observe (Bromley et al. 1989).

Issue 2—Expenses

The first real problem between the evaluation team and the agency concerned cost of the evaluation. The agency agreed to triple the budget after the

realistic costs of conducting a comprehensive evaluation were explained. The second problem concerned the test that was in use to measure knowledge gain.

Issue 3—Instrument Development

When the content of the student course and the relative importance of each section of the course were determined, it became abundantly clear that the existing test was inadequate. At least 50% of the questions were invalid and the balance of the test did not reflect agency priorities. A new test was developed and utilized (Hampton and Bromley 1982).

Throughout the evaluation, the program evaluation team maintained close contact with the agency coordinator and other staff involved in the hunter education program.

Issue 4—Unexpected Outcomes

The third and most serious problem encountered by the evaluation team came when the report was presented to the agency administrators. The team found serious problems with the course and made 30 pointed recommendations for improvement. However, the major obstacle to rational consideration of the recommendations came from the soon-to-be retired program director who was personally embarrassed. The team and its report were thoroughly scrutinized for weaknesses, but because a defensible evaluation management model had been followed and because the evaluation techniques used were based on strong social science methods, the recommendations stood the test. Within 6 months of the receiving the report and taking over the program leadership, the new program manager implemented 88% of the recommendations of the evaluation (Bromley et al. 1988).

Conclusions and Recommendations

Our experiences with evaluating educational programs relating to environmental education and sportsman education indicate that the actual process of evaluation will rarely if ever conform to the textbook, linear format. Rather, the process is likely to be complicated by numerous feedback loops. Frequent and meaningful communication between evaluators and agency administrators and staff is needed to produce useful recommendations for program improvement.

Educational program evaluation is not to be undertaken lightly. The value of an evaluation is directly proportionate to the support provided by the agency administrators and the competence of the evaluation team. However, in an era of increasing attention to program accountability, rigorous evaluation of all agency educational programs is recommended. Although it might be desirable for agencies to have staff capable of conducting routine measurement of program effectiveness, the use of periodic external evaluation should be encouraged. This is especially true whenever principal program personnel change, external evaluation is called for. When an external evaluation is contracted for, agency administrators need to respect the professional standards of evaluation. Meddling with the actual measurement instruments and other techniques vio-

lates the process just as certainly as biasing data collection in assessing wildlife populations voids it's use for making management recommendations. To make the evaluation a success, the evaluator needs to establish a high level of trust with agency personnel from the beginning by explaining the evaluation process and by actively listening to the needs of the client. Attention needs to be given to the differences in perceptions by agency administrators as well as by agency program delivery staff. That benefits of evaluation will outweigh the costs needs to be accepted from the beginning. When the evaluation is successful in assessing the merit or worth of a program, those successes need to be presented to interested people. Evaluation recommendations only have the potential to bring on change and improvement if those recommendations are acted upon by agency personnel. Taking a positive view of evaluation from beginning to end lends credibility to all involved.

Literature Cited

Armstrong, J. B. and J. C. Impara. 1990. The effects of order of testing on attitude scores. J. Environ. Ed. 21(3):37–39.

——— and ———. 1991. The impact of an environmental education program on knowledge and attitudes. J. Environ. Ed. 22(4)36–40.

Bromley, P. T. and E. L. Hampton. 1981. Considerations in evaluating hunter education programs. Proc. Annu. Conf. Southeastern Assoc. Fish and Wildl. Agencies 35:585–589.

———, ———, and J. D. Wellman. 1989. A study of ethical and safe behavior by dove hunters in Virginia. Wildl. Soc. Bull. 17:450–454.

———, H. Foster, E. L. Hampton, J. D. Wellman. 1988. Response to evaluation of the hunter education program in Virginia. Proc. Annu. Conf. Southeastern Assoc. Fish and Wildl. Agencies 42:585–589.

Gronlund, N. E. 1985. Measurement and Evaluation in Teaching. MacMillan Publishing Co., Inc., New York. 540pp.

Hampton, E. L. and P. T. Bromley. 1982. Development of new student examination for the Virginia Hunter Education Program. Proc. Annu. Conf. Southeastern Assoc. Fish and Wildl. Agencies 36:800–804.

House, E. R. 1980. Evaluating with validity. Sage Publ. Co. Beverly Hills, Calif. 295pp.

Issac, S. and W. B. Michael. 1985. Handbook in evaluation and research for education and the behavioral sciences. (2nd ed.) EdITS. San Diego, Calif. 234pp.

Joint Committee on Standards for Educational Evaluation. 1981. Standards for evaluation of educational programs, projects, and materials. McGraw-Hill Publ. Co. New York. 161pp.

Passineau, J. F. 1976. The development of the Environmental Awareness Inventory (EAI)—an assessment instrument for the evaluation of environmental education. Diss. Ab. 36(9):5654-A. (UMI 76-6241; 387pp.)

Patton, M. Q. 1980. Qualitative evaluation methods. Sage Publ. Newbury Park, Calif. 379pp.

Stout, R. J. and R. B. Peyton. 1988. The need for wildlife education program evaluation: a case study. Trans. 53rd North Am. Wildl. and Nat. Resour. Conf. Pp. 552–559.

Stufflebeam, D. L. 1973. Evaluation as enlightenment for decision making. Pages 143–150 *in* B. R. Worthen and J. R. Sanders, Educational evaluation: Theory and practice. Wadsworth Publ. Belmont, Calif.

Worthen, B. R. and J. R. Sanders. 1987. Educational evaluation: Alternative approaches and practical guidelines. Longman Publ. Co., New York. 450pp.

——— and K. R. White. 1987. Evaluating educational and social programs: Guidelines for proposal review, onsite evaluation, evaluation contracts, and technical assistance. Kluwer-Nijhoff Publ., Boston, Mass. 347pp.