The Journal of Environmental Education, 2001, Vol. 32 No. 2 18–27

Stake's Countenance Model: Evaluating an Environmental Education Professional Development Course

BARBARA BONSALL WOOD

ABSTRACT: The purpose of this study was to evaluate an environmental education professional development course using Stake's Countenance Model as the organizational framework. A complex analysis of 10 course features using 8 evaluation instruments focused on congruence between what was intended to occur and what was actually observed to occur before, during, and after instruction. The Stake model demonstrated its effectiveness by facilitating a thorough examination of both quantitative and qualitative data during all phases of course execution.

Professional development courses are a recognized means to assist teachers in increasing their knowledge bases throughout their careers (Tammi, 1991). As a course is developed and implemented, it is important to obtain periodic information about the quality of the outcomes to determine if the course goals are being met (Gottman & Clasen, 1972). Unfortunately, most course evaluations are cursory, consisting of short self-report surveys. For the impact of a course to be fully understood, however, it must be thoroughly described and judged (Stake, 1977). Assessing any course is a complex process, and professional development courses in the environmental science arena are particularly so because they have multiple learning goals, including the acquisition and synthesis of (1) new science content, (2) social context issues, and (3) effective pedagogical techniques.

Barbara Bonsall Wood *teaches at Thomas Jefferson High School for Science and Technology, part of the Fairfax County Public School System in Virginia. She is also a consultant for the National Wildlife Federation.*

Collecting data on such a range of outcomes as well as operational and curricular features can be a daunting task; consequently, a systematic process is needed to make decisions on the worth of the course. The purpose of this study was to evaluate an environmental professional development course using Stake's Countenance Model as the organizational framework. If the Stake model manages a complex evaluation effectively, then it may prove a useful tool to structure evaluations of other environmental professional development courses.

Background for the Evaluation

An evaluation may be defined as the assessment of goal achievement through the collection and analysis of data useful in making decisions on the worth of a program (Hamm, 1985; Scriven, 1967; Stake, 1967). Evaluation models are useful in guiding the management of data collection and analysis. Because the design of this study was complex and used diverse quantitative and qualitative methods, Stake's Countenance Model (1977) served as a framework for this evaluation. This model and the subject of the evaluation, the Chesapeake Watershed Ecology course, are described as follows.

Stake's Countenance Model

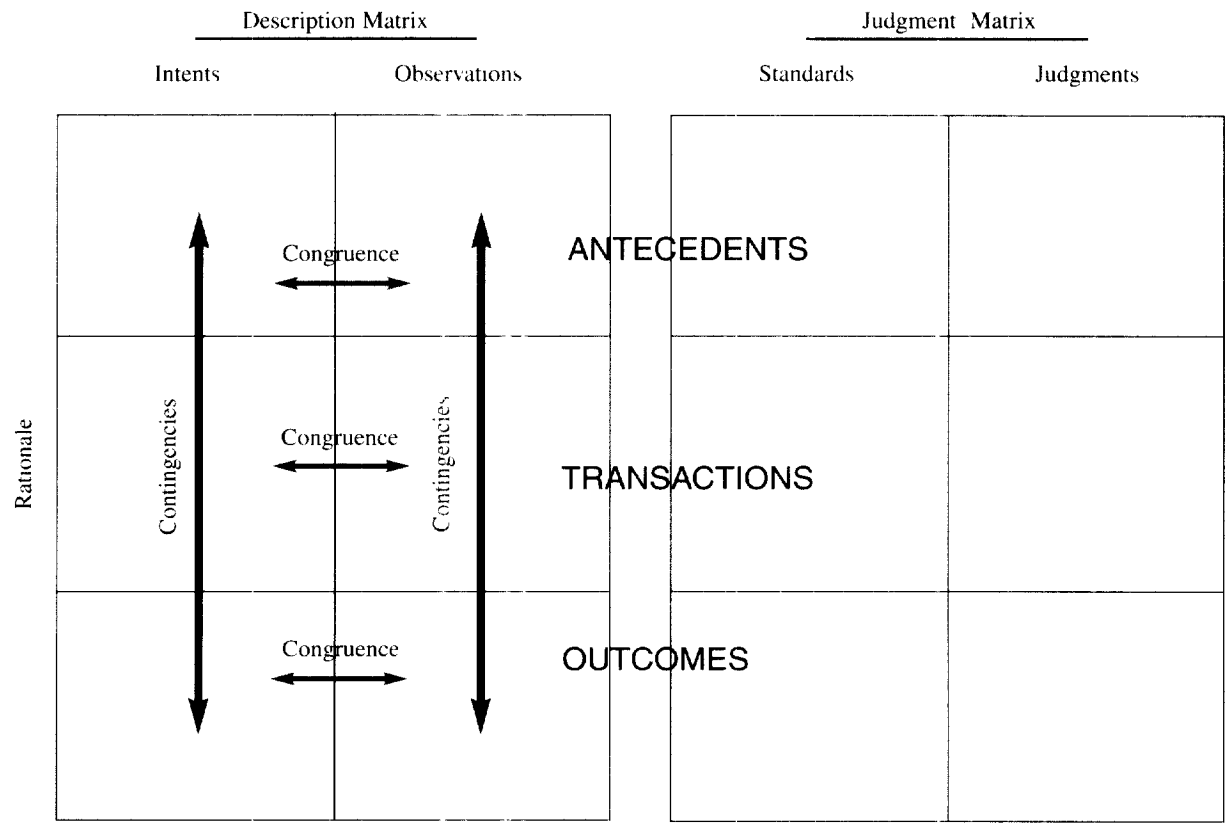
According to Stake, the two major operations, or countenances, of any evaluation are complete description and judgment of the program. Stake created an evaluation framework to aid an evaluator in collecting, organizing, and interpreting quantitative and qualitative data for these purposes (Worthen & Sanders, 1987). His scheme draws attention to the differences between descriptive and judgmental acts according to their phase in an educational program: antecedent, transaction, and outcome (Popham, 1993). An antecedent is a condition existing prior to instruction that may relate to outcomes (Popham, 1993; Stake, 1977). Transactions are successive engagements or dynamic encounters constituting the process of instruction. Outcomes are the effects of the instructional experience (Popham, (1993; Stake, 1977).

Stake further divided descriptive acts according to whether they referred to what was intended or what was actually observed to occur. Judgmental acts were divided according to whether they refer to the standards used in making judgments or to the actual judgments (Popham, 1993). The evaluator records what the developer intended, what observers perceived, what patrons expect, and what value is placed on each aspect (Stake, 1977). Figure 1 depicts the layout of statements and data to be collected by the evaluator.

Stake's model is useful because it provides extensive evidence to support the success of a course (Hamm, 1985) by documenting plausible links between all components of the course program. An evaluator begins by looking at intents. *Intents* are planned-for conditions and behaviors, including desired, anticipated, and feared effects. The evaluator will next analyze information in the descriptive matrix by looking at congruence between intents and observations. Observations are descriptions of environmental events and consequences (Stake, 1977). Moving across the chart in a similar fashion, the evaluator will then apply standards to the descriptive data to form judgments.

The evaluator, however, also studies the relationships, or contingencies, between antecedents, transactions, and outcome variables. Program antecedents are conditions in existence before the program is run; program transactions are dynamic encounters and processes occurring during the operation of the program; outcomes are the effects, both

Figure 1. Stake's layout of statements and data to be collected by the evaluator of an educational program. (Sources: Stake, 1967; Worthen & Sanders, 1987)



intended and unintended, of the program. Outcomes are dependent on transactions and on antecedent conditions (Stake, 1977; Worthen & Sanders, 1987). An understanding of these vertical relationships on the chart is important for course improvement. Formative evaluations are more interested in contingencies; summative are more interested in congruence (Stake, 1977).

Providing a conceptual framework for completely thinking through the procedures of an evaluation, the Stake data matrices engage the evaluator continuously in description and judgment at the beginning, during, and at the end of an instructional program (Popham, 1993). The matrices divide a complex interaction into manageable units and identify points at which instruments should be used to collect data (Hamm, 1985).

Chesapeake Watershed Ecology

Format

The course met for 6 hours daily over a period of 2 weeks. The school system requested this highly concentrated schedule to accommodate the limited time available for teacher training.

Academic Content

The Chesapeake Watershed Ecology course was designed to educate teachers about research and instructional strategies used to investigate community environmental issues. (In this case, the course focused on the reconstruction of an important local bridge that would necessitate the destruction of nearby wetlands and aquatic habitats.) Instruction was divided into a series of instructional units, or components:

Experimental design. This component showed teachers how to create, implement, and critique scientific investigations.

Map orientation. Before going into the field, teachers studied area maps and proposals for construction of the bridge.

Deepwater monitoring. Using sophisticated monitoring equipment located on a boat, teachers collected data about a water column at an offshore location to determine profiles of oxygen, temperature, conductivity, and light. They also calculated the concentration of zooplankton, estimated sediment particle size, and calculated descriptive statistics for collected fish.

Chemical analysis. Nine water quality tests were performed and analyzed by the teachers: oxygen, biological oxygen demand, phosphate, nitrate, temperature, total solids, turbidity, pH, and fecal coliform.

Biological indicators. Teachers learned to collect, identify, and categorize macroinvertebrates collected from a stream.

Wetland delineation. Teachers evaluated plant communities by identifying and determining the wetland indicator status of dominant plants. They also determined the presence of hydric soils, gathered hydrological evidence for saturated soils, and integrated findings from field data with results of a map study to determine the boundaries of a wetland.

Internet. The data summary forms from each component of this course were displayed on an Internet location. Teachers were shown how to access the Internet location and submit all forms.

Instructional Sequence

The program began with introduction to general principles of experimental design, map orientation, field methods, and Internet use. As a final project, teachers created their own investigations using their choice of field techniques learned in class.

The study of the field methods was accomplished in three phases. In the first phase, the teachers prepared for the field work in the classroom by practicing laboratory procedures and learning to recognize organisms. The second phase consisted of three data collection trips; the third phase was in-class data analysis.

Evaluation Methodology

Feature Description

Course features based on evaluation questions, which reflected the broad goals and enabled the objectives of the course, were arranged according to Stake's program phases. For each feature, criterion levels were established to judge discrepancies between what was intended and what was observed to occur. These features, related evaluation questions, and data collection methods are summarized as follows.

Program Antecedents

Teacher background. Were the teachers in the course representative of the county teacher population? It was intended that the treatment group be representative of the county teacher population and that the teachers be initially unfamiliar with course components. Consequently, demographic characteristics were collected from background information cards on which were recorded gender, age, academic preparation, school distribution, and teaching experience. A pretest was given to alert the researcher about specific components familiar to course participants.

Appropriate curriculum. Were the content and design of the course academically and scientifically appropriate? Experts on science education and on the techniques employed in the course judged the course materials using an expert review questionnaire.

Resource availability. Were resources available? The teachers reported concern about equipment in an opinion survey given at the end of the course. This survey asked the teachers to rate, on a scale from 1 to 5, seven characteristics for each component. Besides "resource availability," these

characteristics included "scientific accuracy," "academic appropriateness," "organization," "usefulness," "interest," and "future use."

Program Transactions

Component participation. In which components did the teachers participate? The frequency of absences was determined by referencing attendance records and triangulating them with observational records and portfolio documents.

Behavioral interactions. How did the teachers interact during the course? What concerns did they have when exposed to the new material? Teachers were asked to keep a daily journal for recording their reactions to the program. Six questions were given to the teachers daily to focus their comments. Following procedures established by Bogdan and Biklen (1992), journal responses were classified and interpreted, then triangulated and validated with other instruments and observations by the instructor

Course choreography. Did the program run smoothly? The journal maintained daily by the instructor documented classroom management, and input from teacher journals augmented these observations. Also, the opinion survey of the teachers solicited responses related to course organization.

Program Outcomes

Improved performance. The researcher developed a short examination of cognitive enabling skills to establish the teachers' previous knowledge of component procedures. Pretest and posttest versions of this instrument were then checked for content validity and reliability (0.871). The pretest was then compared with the matching posttest with *t* tests to statistically measure cognitive treatment outcomes.

Evaluation of teacher performance was accomplished using a control-group time series format (Issac & Michael, 1972).

For this design, teachers enrolled in the class were given an objective, multiple-choice pretest (O_1) before instruction (X), and they were given a similar posttest (O_6) afterward. A similar posttest was given 2 weeks after the course was completed. The control group was a sample of 30 county high school science teachers who took the pretest and one posttest under similar conditions but who did not have instruction in the intervening period between the tests.

Except for O_2, which refers to background information, the other three instruments (O_3, O_4, O_5) continually monitored the academic progress of the teachers throughout the course. O_3 referred to instructor and teacher journal observations about the transactional abilities of the teachers; O_4 referred to observations of academic performance obtained from portfolios; and O_5 was the attendance record that substantiated teacher participation in a component.

Teacher attitudes. What were teacher attitudes toward the course? Data were collected using the teacher opinion survey and daily journals maintained by the instructor and by the teachers.

Intent to use. How did teachers intend to use course knowledge and skills in their classrooms? Evidence of the teachers' intentions to use both subject matter and pedagogical techniques in their high school classes later was collected quantitatively with the opinion survey and qualitatively from comments in the teacher journals.

Unexpected outcomes. Were there unanticipated effects? Entries in the teacher's journals documented unplanned events and reactions and helped to explain unclear results and developmental issues.

Analysis

Eight data collection instruments were used to collect data on the 10 course features described above: a pretest, the posttests, a teacher opinion survey, an expert review questionnaire, attendance records, background information cards, the teacher journals, and an instructor journal. A summary of these results for each program feature is listed below. These results are also displayed in the complete Stake chart (see Tables 1–4). The intents, observations, and criteria judgments for each of the 10 program features are grouped in the tables according to whether they are antecedents, transactions, or outcomes.

Evaluation Results

Congruence

The completed countenance matrix for this evaluation is shown in Tables 1–4. The outstanding characteristics of each course feature are summarized as follows:

Teacher background described the "representativeness" of the teachers enrolled in the course relative to the county population of teachers. Course participants were found to be similar to the county teachers in age, gender, years of teaching experience, academic degree status, school distribution, and initial familiarity with course components.

An *appropriate curriculum* is one that is scientifically accurate and academically appropriate. All components were found appropriate except *experimental design*, which emphasized basic knowledge already familiar to the teachers rather than developing needed advanced skills.

Resource availability was important for this technical course. Most equipment and materials were in place and functioned as intended; but there were problems with the format of some materials (maps, directions for complex calculations, organism keys) and access to some equipment (computers and boat meters).

Component participation ensured that each teacher was exposed to all course material. Teacher attendance was high. However, the teachers did not experience the entire experimental design component because much time was spent developing introductory concepts rather than advanced skills. As a result, many portfolios were incom-

TABLE 1. Countenance Matrix, Completed: Program Antecedents

Description matrix		Judgment matrix	
Intents	Observations	Standards	Judgments
Program antecedents			
Teacher background			
Treatment and control groups should be similar to the county population on age, gender, years of teaching experience, academic degree status, school distribution, and initial familiarity with course components.	Teachers who participated were similar to the county teacher population both in their demographic characteristics and in their previous knowledge of course components. Although they taught in schools all over the county, most of the teachers in the course had about five years' teaching experience, were usually in their 30s, had a master's degree, and were familiar with experimental design basics but not other course components.	Schools should be evenly distributed for all groups, gender distribution should be similar; the median for age, experience, and degree status should be the same for all three groups; pretest results should be similar for experimental groups.	Teachers enrolled in the class are roughly representative of the county teacher population.
Appropriate curriculum			
Science content should be accurate and easy to understand; material should be challenging to teachers but adaptable for high school student use.	Minor corrections were made in response to expert advice. "Scientific accuracy" and "academic appropriateness" averaged over 4.5 on the opinion survey. The ratio of positive to negative comments in the journals was 2.36. Pretests indicated that the science teachers were generally acquainted with basic knowledge of experimental design, but portfolio scores for this component were under 90%.	Corrections made from expert review; responses on the teacher survey should average higher than 3.0 on academic appropriateness and scientific accuracy; more positive than negative comments in expert review and in journal questions; teachers not previously familiar with components	Course curriculum was appropriate except for experimental design. Many of the activities in this component related to developing introductory concepts were not necessary. Emphasis on advanced objectives would have been more appropriate to consolidate required skills.
Resource availability			
All equipment and materials should be available and functioning.	Resource availability was observed to be very high (4.73) Teachers reported problems with map legibility, light calculation directions, labeling and preparation of chemicals, plant and fish key organization, textbook funding, boat equipment availability, and continuous computer access. Also models demonstrating estuarine circulation were needed for conceptual understanding. Material strong points included the specimen sets, the resource manual, and the audiovisual aids	Responses on teacher survey should average higher than 3.0 for resource availability; more positive than negative comments in expert review and in the journal; researcher agreement.	The large number of complex, technical materials necessary were usually available, functioning, and served to enhance the course. Specimen sets, audiovisual aids, and the resource manual were well done. Corrections should be made to improve map legibility, light calculation directions, and key organization. Accuracy and arrangement of chemicals should be checked. Conceptual models are needed to demonstrate estuarine circulation. Computer, boat equipment, and textbook availability must be more secure

TABLE 2. Countenance Matrix, Completed: Program Transactions

Description matrix		Judgment matrix	
Intents	Observations	Standards	Judgments
Program transactions			
Component participation			
All teachers complete all components.	Teacher attendance averaged 96.4%. No component fell below 91% attendance, and only two teachers were present less than 85% of the time. Any daily absence cost a teacher six hours of contact time. Portfolio completion averaged 93.8%. Three teachers completed less than 90% of the portfolio. The experimental design component averaged only 86.2% completion, with half of the teachers' scores falling under 90%. Experimental design and Internet use were not completely covered as intended.	Teacher attendance rates exceed 90%; teachers complete 90% of their portfolios.	Overall teacher participation exceeded the standards for all components. Nevertheless, the experimental design and Internet components were incomplete and did not allow the teachers to fully experience their application. Insufficient numbers of functioning computers impeded integration of the Internet component. The emphasis of the experimental design component did not allow the teachers to move beyond basic knowledge.
Course choreography			
Activities run smoothly, course is well organized.	Teachers gave good ratings for "organization" (4.63) on the Teacher Opinion Survey, and they made many positive comments in their journals concerning the operation of most activities; however, they did not like the fast pace of the course. The effects of the rushed schedule were negative and pervasive.	Average for responses on teacher survey should be higher than 3.0 for organization; more positive than negative comments on teacher journal related to logistics. Instructor supporting evidence.	Most activities of the course were well coordinated and organized. However, teachers need additional time after one topic to study and reflect on their learning before moving on to a new subject; and the instructor needs more time to remove lab equipment from one exercise and properly set up the next.
Behavioral interactions			
Teachers are actively engaged in classroom activities.	There were twice as many positive as negative comments regarding teacher interest, attention, and group relationships. In spite of some negative reactions to one teacher, personal interactions with the instructor and within the rest of the class complemented the activities. Teachers appeared interested.	More positive than negative comments in teacher journal questions. Instructor reports evidence of active participation	Cooperative activities with teachers and the instructor were a positive aspect of the course. Teachers were observed to be actively engaged in course activities.

plete for this component. Also, teachers did not fully experience the application of the Internet component due to restricted access of computers at the host school.

Course choreography described the organization and coordination of the course. Most of the activities were found to run smoothly, but the overall pace of the course was too fast.

Behavioral interactions described the engagement level of the teachers in the classroom. Teachers seemed to be interested and actively engaged during all components.

Improved performance collected information on cognitive

TABLE 3. Countenance Matrix, Completed: Program Outcomes, Part 1

Description matrix		Judgment matrix	
Intents	Observations	Standards	Judgments
Program outcomes			
Improved performance			
Teachers are able to perform all skills specified in the instructional objectives.	The pretest and posttests indicated that overall cognitive gains were significant for the treatment group. Gains were also significant for the separate deepwater monitoring, chemical analysis, biological indicators, and wetland delineation components. Accomplishment of instructional objectives was also observed in the portfolios and reported by the teachers in the journals for these components. All responses in the journals to the first question on Day 9 were more detailed than on Day 1. The teachers, however, had high pretest scores for the experimental design component, which did not improve significantly during the course	Posttest results are significantly higher than the pretest. Teachers make positive comments in the journals on their learning and respond more fully to the first journal question on Day 9 than to the first question on Day 1. Ninety percent of teachers finish portfolios.	The course accomplished the instructional objectives for deepwater monitoring, chemical analysis, biological indicators, and wetland delineation. The experimental design component was not effective in improving previous skills.
Teacher attitudes			
Teachers enjoy the course and find the material useful and interesting.	There were three times as many positive as negative comments made by the teachers in the journal related to teacher attitudes. "Interest" averaged a rating of 4.31 on the Teacher Opinion Survey, and "usefulness" averaged 4.52. Experimental design and Internet use scored relatively low for "interest" and "usefulness." The researcher observed that one of the strong points of the curriculum was the variety of topics covered in the many components, which allowed the teachers to expand knowledge in their current field as well as explore new disciplines	More positive than negative comments in the teacher journals. Responses on teacher survey should average higher than 3.0 for "interest" and "usefulness". Instructor evidence supports.	The teachers generally enjoyed the course and found the material useful, interesting, and realistic. Teachers were less enthusiastic about experimental design and Internet use than other components. In spite of the diversity of their academic backgrounds and experience, the teachers found many areas of individual interest

gains made by the teachers related to component knowledge. A battery of pretests and posttests established treatment effectiveness for all components except experimental design. The teachers demonstrated basic knowledge of these investigation concepts, but scores did not improve significantly after treatment. Final portfolio scores were low for experimental design.

Teacher attitudes are an important gage of educational experiences. The teachers found Chesapeake Watershed Ecology to be useful and enjoyable. The variety of course subjects complemented the diversity of teacher backgrounds so that each teacher found many areas of interest. Although all components scored above the criteria standard

TABLE 4. Countenance Matrix, Completed: Program Outcomes, Part 2

Description matrix		Judgment matrix	
Intents	Observations	Standards	Judgments
Intent to use			
Teachers will plan to use course components and instructional stategies in their classrooms later.	The ratio of positive to negative comments regarding their intention to use course activities in the future was 4.48. "Future use" rated 4.38 on the Teacher Opinion Survey Teacher comments indicated enthusiasm for using most components. They also expressed concerns about administrative resistance to field trips. Activities sometimes required adaptations for younger or special needs students, and some calculations would not be included due to the perceived difficulty level.	Responses on teacher survey should average higher than 3.0 for intent to use; more positive than negative comments pertaining to potential usage in the journals.	The teachers indicated that they intended to use all major components in their classes although they anticipated administrative resistance to the field trips. The course was successful in encouraging teachers to consider including course activities in their curricula.
Unexpected outcomes			
Unforeseen consequences enhance the course goals and objectives.	*Positive:* • Teachers developed enhanced professional confidence from knowledge gains. *Negative* • Long class days precluded adequate study time and absorption of new learning. • Teachers' perception of administrative barriers hampered intentions to use course components in their classrooms.	There were more positive unexpected outcomes than negative	There were more negative unexpected consequences. Problems related to the concentrated schedule hindered course objectives, and must be addressed. Administrative barrier concerns impact longer term goals related to intent to use. The positive outcome of increased professional confidence, in spite of the two negative findings, is directly supportive of course goals and objectives.

for interest, the deepwater monitoring component had the highest average, and Internet had the lowest

Intent to use describes teacher interest in applying a course component to the secondary science classroom. In this study, the intention to use course material was above the criterion level for all components, but the Internet and deepwater components were significantly lower than other components.

Unexpected outcomes are major unanticipated effects of the program. Three were discovered: (1) The teachers displayed a marked enhancement of professional confidence related to knowledge gains; (2) the teachers reported that the concentrated schedule of the course did not allow them enough time between lessons to study and reflect; and (3) the teachers perceived strong administrative barriers to science field trips that they believed would hinder incorporation of these activities into their own curricula

An analysis of congruence indicated that 9 of the 10 course features exceeded criteria standards for that feature. Only unexpected outcomes fell below the criterion level for the course as a whole. Based upon the results of this summative assessment, it was concluded that the Chesapeake Watershed Ecology course was effective and worthy of continued implementation.

Contingencies

A formative evaluation, however, uses more than specified discrepancies to assess course merit during the developing stages. Many qualifying aspects were also found to be important in considering course improvement. Four relationships, or contingencies, between the above features were found to affect the course:

1. Resource availability was a critical feature because this technical course depended heavily on equipment and sup-

porting materials. Although most resources were appropriate and functioning, occasional difficulties altered the intended exposure of the teachers to the curriculum (course choreography and component participation), which affected their opinion of the activities (teacher attitudes).

2. It was important to adapt the curriculum (appropriate curriculum) to accommodate the various abilities and interests of the teachers (teacher background). The value teachers placed upon course activities (teacher attitudes) was related to the match they perceived with their professional needs.

3. Significant gains in learning (improved performance) influenced the teachers' intentions to use course activities in their own classrooms (intent to use). New science skills and pedagogical knowledge brought the teachers professional confidence (teacher attitudes) to expand their curriculum.

4. The professional–political context in which the teachers work was another important factor affecting their curricular choices (intent to use). Course experiences were measured against administrative policies that discouraged field application. An appropriate curriculum must also be feasible (appropriate curriculum).

Course Recommendations

Based on feature congruence and contingency analysis, the following modifications were recommended by the evaluator:

1. Reduce the length of the experimental design component. Place emphasis on advanced design skills. Reduction of this component would also free time for the Internet component, which was not completed.

2. Decrease the length of the class day and increase the number of days during which the course is given. Shorter days would allow more time between lessons for the teachers to study, and a longer course would ensure exposure to all components as intended.

3. Correct resource problems. More computers must be available throughout the course in order to properly integrate the Internet component. Activities with identified format flaws should be redesigned.

4. Work with county school officials to obtain more field equipment and alleviate administrative barriers for field trips.

Discussion

Course evaluators frequently collect information on participant demographics, cognitive gains, and attitudinal changes. A standard assessment of Chesapeake Watershed Ecology would have determined that the teachers were roughly representative of the county population of high school science teachers, that they made significant cognitive gains, and that they enjoyed the course, although they thought it was too fast paced. The Stake evaluation, however, facilitated an in-depth understanding of all aspects of the course program. It not only allowed the evaluator to determine both the anticipated and unanticipated outcomes of the course but also uncovered reasons and consequences for the effects.

Three relationships discovered in the Stake analyses would have been difficult to detect without a structure to organize the complex findings of the numerous instruments:

1. Pretest scores from the experimental design component were high and stayed high on the posttests, but the portfolio scores on this component were lower than expected. Triangulation of these results with teacher background information and journal responses confirmed that the teachers were familiar with the basic concepts of experimental design but not the advanced techniques needed to complete the portfolios.

2. Teacher interest in the Internet component was relatively low, and interest in the deepwater component was extremely high, although the results of the teachers' survey indicated the teachers did not expect to use curricular material from either the Internet or deepwater components in their classrooms. These apparently contradictory outcomes were found to be related to conditions in the antecedent and transition phases. Component participation analysis confirmed that teachers were not exposed to the entire Internet component as a result of computer inaccessibility during the course, resulting in low teacher interest. The deepwater component, on the other hand, was executed completely; consequently, teacher interest was high. Although the teachers enjoyed the deepwater unit, concerns about administrative barriers to obtaining equipment dampened their enthusiasm. The Stake analysis of contingencies established the importance of these ties between perceived resource availability (an antecedent condition), class participation (a transactional interaction), and curricular choices (an outcome variable).

3. A thorough quantitative and qualitative analysis of teacher journals in this study documented a link between knowledge gains and the improved professional confidence expressed by the teachers. Additionally, the analysis of contingencies also tied these outcomes to an antecedent condition, that is, an appropriate curriculum designed to meet the diverse professional needs of the participants.

These feature interactions did not become evident until the horizontal and vertical analysis of matrix variables was done By facilitating the discovery and documentation of these complex relationships, the Stake model provided a solid basis to support detailed recommendations and a confident judgment on the worthiness of the course.

Summary

The purpose of this study was to evaluate an environmental science professional development course using Stake's Countenance Model, an evaluation framework that sequentially organizes research questions according to their impact on intended course features. Ten course features were analyzed using eight quantitative and qualitative data collection instruments. The model managed the analysis of congruence between what was intended to occur and what

was actually observed to occur as well as contingency interactions between evaluation variables. By allowing the simultaneous analysis of multiple facets, the evaluation discovered significant interactions important for course improvement. This study has demonstrated effectiveness of the Stake model for the formative evaluation of an environmental professional development course because it facilitated a thorough examination of both qualitative and quantitative data during all phases of program execution.

REFERENCES

Bogdan, R. C., & Biklen, S. K. (1992). *Qualitative research for education: An introduction to theory and methods.* Boston: Allyn and Bacon.

Gottman, J. M., & Clasen, R. E. (1972). *Evaluation in education: A practitioner's guide.* Itasca, IL: F. E. Peacock.

Hamm, R. W. (1985). *A systematic evaluation of an environmental investigation course.* (Doctoral dissertation, Georgia State University). (ERIC Document Reproduction Service No. ED-256-622)

Issac, S., & Michael, W. B. (1972). *Handbook in research and evaluation.* San Diego, CA: Robert R. Knapp.

Popham, W. J. (1993). *Educational evaluation.* Boston: Allyn and Bacon.

Scriven, M. (1967). *The methodology of evaluation.* In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Chicago: Rand McNally.

Stake, R. E. (1967). *Toward technology for the evaluation of educational programs.* In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 1–12). Chicago: Rand McNally.

Stake, R. E. (1977). *The countenance of educational evaluation.* In A. A. Bellack & H. M. Kliebard (Eds.), *Curriculum and evaluation* (pp. 372–390). Berkeley, CA: McCutchan.

Tammi, J. K. (1991). The perceived content and pedagogical needs of selected Illinois teachers of science: Implications for staff development. (Doctoral dissertation, Northern Illinois University, 1991). *Dissertation Abstracts International, 52-08A,* 2812.

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