Curriculum Evaluation

EDCI 547 Foundations of Curriculum Studies

By

Debra Martin
Curriculum evaluation is essential to curriculum development, implementation, and maintenance. However, ideally, evaluation determines the value of some action or program, the degree to which it helps students meet standards, and its importance. Implicitly and explicitly, evaluation reflects value judgments about previous curricula and instructional designs. Evaluation critiques previous documents, plans, and actions. We believe that assessment (evaluation) involves value judgments as to merit and worth. These judgments affect which data we gather and how we view that data.
Further complicating curriculum evaluation are the explosions of knowledge regarding how the brain functions, how people learn, how the political realm affects schooling, how new pedagogies can address the needs of diverse student populations, how curricula can be created using various modern and postmodern approaches, and how assessment devices can be created and modified to get at the essences of learning. However, we fail to get an accurate picture of the depth and breadth of students, knowledge and cognition. They do not adequately address student creativity, compassion, commitment to action, and enthusiasm, and not a "view of how, say, students' understandings and skills evolve. It does not necessarily indicate the amount of learnings. Adding to the difficulty of evaluating the curriculum, is the increasingly voiced demand that assessment fair and appropriate for diverse students.
The Nature and Purpose of Evaluation

- Evaluators gather and interpret data to determine whether to accept, change, or eliminate aspects of the curriculum, such as particular textbooks. Curriculum evaluation is necessary not only at the end of a program or school year but also at various points throughout the program’s development and implementation.

- At the beginning of curriculum development the very concept of the program must be evaluated. Does the program have worth and merit? Evaluation identifies the curriculum’s strengths and weaknesses before and after implementation, in terms of effectiveness.

- Pellegrino, Chudowsky, and Glaser view assessment as a process of reasoning from evidence. The process of reasoning from evidence in curriculum evaluation can be conceptualized as an hourglass.
Process of Reasoning from Evidence in Curriculum Evaluation

Placing curriculum at one corner of the top of the hourglass with cognition at the other corner.

The neck of the hourglass represents the observation stage of reasoning.

The base of the hourglass represents interpretation.

The curriculum organizes subject matter in terms of scope and sequence.
Educators must make evaluative judgments regarding the worth of the subject matter being considered and organized as well as the political and social climates within which the curriculum will exist. And is consistent with curriculum theory?

- **Cognitive theories** inform us in our data gathering. Assist teachers in shaping their instructional approaches and evaluating students' learning.

- **Observation** includes all the means by which data are gathered, it may involve written tests, reviews of students' work (e.g. their portfolios), observation includes questionnaires, checklists, inventories, interview schedules, and video performances. It also includes data on teachers-for example, from observations of teachers.

- In the interpretation stage of curriculum evaluation, draw on their assumptions about curriculum and cognition. They process data into evidence regarding the curriculum's success. At the classroom level, interpretation tends to be informal and qualitative, including interpretation regarding teachers, instructional approaches. Interpretation implicitly draws on theories of testing, statistical models of data analysis, and theories of decision making.

- **Evaluation** must remain connected to the totality of curricular activities. Evaluators investigate the appropriateness of a particular assessment procedure or form of assessment. Evaluation focuses on the effectiveness of a school environment. However, most evaluation focus on curriculum and/or instruction.
Harriet Talmage posed five types of questions that educators can consider when evaluating curricula; questions of intrinsic, instrumental, comparative, idealization, and decision value.

1. **Intrinsic value** addresses the curriculum's goodness and appropriateness. It deals with both the planned curriculum and the finished (delivered) curriculum. They perceive the curriculum in light of the purpose of education that they see as paramount. (Should we stress critical thinking, citizenship, or preparation for employment?) They also see curriculum in terms of their preferred learning theory.

2. **The question of instrumental value** asks, "What is the curriculum good for, and who is its intended audience?" attempting to link the planned curriculum with the program's stated goals and objectives. It also addresses which students will accomplish what is planned in the curriculum and to what extent. Evaluation efforts should identify the types of students who are likely to benefit the most from the planned curriculum.

3. People faced with possible new programs often ask the question of **comparative value**. Is the proposed new program better than the one it is supposed to replace? Usually, new programs are created because people feel that the existing program is inadequate.

4. **Idealization value addresses** ways to improve a curriculum. They consider information on how the program is working and ask themselves if there are alternative ways to make the program even better. They continually reconsider how they might fine-tune the program's content, materials, methods, and so on, so that students will optimally benefit.

5. **Decision value deals** with the vital role that the previous four questions play in the evaluation process. The evaluator and the curriculum decision maker should now have evidence documented in such a manner that they can decide whether to retain, modify, or discard the new program.
Definitions of Evaluation

To Name a Few

- **Blaine Worthen and James Sanders** define evaluation as "the formal determination of the quality, effectiveness, or value of a program, produce, project, process, objective, or curriculum." Evaluation includes inquiry and judgment methods.

- **Abbie Brown and Timothy Green** define evaluation as the process of judging, based on gathered data, the success level of an individual learning or product effectiveness.

- **Daniel Stufflebeam** has defined evaluation as "the process of delineating, obtaining, and providing useful information for judging decision alternatives.

- **Collin Marsh and George Willis** indicate that evaluation permeates all human activity. It deals with questions such as these; Is something worth doing? How well is it being done? Do I like doing it? Should I spend my time doing something else?

Most evaluators maintain that while the presence and importance of values cannot be ignored, they can only be considered within a particular context. We judge whether a program reflects its values and if those in charge of a curriculum have made their values explicit.
Fred Kerlinger defined Measurement as assigning numerals to objects or events according to rules. Measurement describes a situation or behavior in numerical terms. We make observations and then assign numbers to aspects of the observed phenomena. Measurement enables educators to record students' degrees of competency.

Evaluation assigns value and meaning to measurement. For example, an evaluator might decide that a score of 70 percent correct answers means "passing" or successful performance.

However, educators must do something with the gathered data. They must decide whether a student who spells 18 of 20 words correctly should get an A, an A, or some other grade. Measurement always precedes evaluation.
Approaches to Evaluation

- How people, process data is influenced by their philosophy and psychology. Those who take a behavioristic, prescriptive, or sequenced approach to evaluation tend to specify specific behaviors or content learned as a result of curriculum and instruction. They like clearly stated objectives, precise indicators of whether their students have achieved the program's intended outcomes. Those who take a humanistic approach are more interested in whether the planned situations have enabled students to improve their self-concepts.

- In general, evaluation enables educators to (1) decide whether to maintain, revise, or replace the existing curriculum; (2) assess individuals (primarily teachers and students) in terms of instruction and learning; and (3) decide whether the existing managerial organization of the school and its program should be maintained or reformed. Also, part of the evaluation focuses on the school environment and the community environment within which the school exists.
Scientific versus Humanistic Approach

- Lee Cronbach has placed scientific and humanistic approaches at opposite ends of the evaluation continuum. **Scientific evaluators favor an experimental approach.** Scientific evaluators tend to concentrate on the learners. They use data, frequently in the form of test scores, to compare students’ achievement in different situations. Data are quantitative, so they can be analyzed statistically. Program decisions are based on the comparative information gathered.

- There is growing interest in **humanistic approaches.** Nontraditional evaluation procedures to obtain more complete pictures of curricula. **The humanistic approach emphasizes** naturalistic case studies and rejects experiments. Humanists prefer to study programs already in place, not programs imposed by evaluators. Naturalistic evaluators ask different questions of different programs. Benefits are described, not quantified. Observations are opportunistic and responsive to the local scene, not prestructured. Humanistic evaluators analyze qualitative data, such as impressions of what they observed. They describe actual incidents. Data gained from interviews and discussions with participants are included in the evaluation. Analysis uncovers patterns among many observations. **Five major humanistic approaches have been identified: interpretive, artistic, systematic, theory driven, and critical-emancipatory.**
1. **The interpretive approach**, the evaluator considers the educational scene and interprets the meaning and significance of people’s actions. Attention to social context is essential. The evaluators are people directly involved with the curriculum, especially teachers and students.

2. **In the artistic approach**, the evaluator engages in aesthetic inquiry, observing classes and other enactments of curricula and then publicly announcing what is good and bad about the curriculum. This approach relies on individual intuition honed by experience. The evaluator focuses on the quality of the relationships between teacher and students.

3. **The systematic approach** is most familiar. Evaluators try to be as objective as possible in their descriptions, employ logical analysis and base their judgments on fact.

4. **Theory-driven approach**. These evaluators apply philosophical, political, and/or social theories when judging the quality of curricula.

5. **Critical-emancipatory** evaluators tend to be the most radical. They judge a curriculum’s quality and effectiveness, according to how well the curriculum counters social forces that impede individual development and fulfillment. These evaluators draw heavily on Jurgen Habermas’s work on the construction of knowledge and meaning. They also draw on **critical theory**, especially Marxist theory.
Utilitarian versus intuitionist Approach

- Evaluation can be classified as either utilitarian or intuitionist. The utilitarian approach is closely linked to the scientific approach, whereas the intuitionist approach is tied to the humanistic approach. Utilitarian evaluation operates according to the premise that the greatest good is that which benefits the greatest number of individuals. Attention is on total group performances. Programs are judged by how they affect the school's overall student population.

- Intuitionist evaluators gather data to judge the program’s impact on individuals or small groups. Program participants, not outside evaluators, consider the program's quality. Everyone affected by the program can make judgments about it.
Intrinsic versus pay-off approach

- Intrinsic evaluators study the curriculum plan separately.
- The evaluators are merely trying to answer the question "how good is the curriculum? Intrinsic evaluators study the particular content included, the way it is sequenced, its accuracy, the types of experiences suggested for dealing with the content, and the types of materials to be employed. They determine if the curriculum has value.
- Once a curriculum's basic worth has been assessed, evaluators must examine the effects of the delivered curriculum. This is pay-off evaluation. Pay-off evaluation receives the most attention from educators because it indicates curriculum's effects on learners in terms of stated objectives. This evaluation approach may involve judgments regarding the differences between pre- and posttests. However, the results reported in pay-off evaluation studies are usually short-term results of a curriculum. Little attention is given to a program long-term outcome.
Formative and Summative Evaluation

- **Formative evaluation** encompasses activities undertaken to improve an intended program—that is, **optimize student learning carried out during program development and implementation**. They gather data, often in classrooms, that inform their decisions as to how to modify these program elements before they are fully implemented. During a curriculum's developmental and early piloting stages, formative evaluation provides frequent, detailed, specific information. **Formative evaluation allows educators to modify, reject, or accept the program as it is evolving.** It can be used to judge the effectiveness of teachers' pedagogical approaches and the productiveness of students’ learning processes. Primarily, formative evaluation focuses on the degree of student learning.

- Various procedures can be used in formative evaluation. Walter Dick, Lou Carey, and James Carey have developed one, especially useful in curriculum development. The procedure has **three** basic phases; one-on-one, or clinical; small groups; and field trial. **In the first phase,** the purpose of this phase is to challenge assumptions regarding what goals and objectives to address, what content to include, and what instructional strategies and educational materials to incorporate into the new curriculum being considered. **In phase two,** curriculum evaluators enact various procedures to determine if the selected (and the tried) curriculum components actually have value, worth and merit. **The final phase** of Dick, Carey, and Carey's approach is field-trial evaluation. Evaluates test the new curriculum **as it is actually taught.** They evaluate the data gathered in the previous phase to determine if results attained with a small sample of teachers and students actually work with the total teacher and student communities. **Teachers are the key evaluators.**
Summative evaluation is aimed at assessing the overall quality of a product and then taught curriculum. Data are gathered to ascertain the new program's worth and effectiveness. Such summative evaluation informs educators that students have met the school's or state's educational standards. It also indicates that teachers have met the minimum accountability standards.

Kirkpatrick's approach can be applied to curriculum evaluation. Kirkpatrick delineates four levels of summative evaluation: 1) reactions, 2) learning, 3) transfer, and 4) results.

Level 1, reactions, focuses on gathering data on how students reacted to the new program.

At level 2, evaluators gather data on whether students have gained new knowledge, skills, and techniques implicit in the new Program's goals and objectives.

At level 3, evaluators pose questions as to whether the individuals who experienced the new program can effectively employ newly acquired skills and knowledge and whether their attitudes have changed for the better.

Level 4, results, is a major challenge for evaluators. The results of a newly developed curriculum may not be evident immediately, or ever
Evaluation Models

▶ Scientific Models

Stake's Congruence-Contingency Model

Robert Stake distinguishes between formal and informal evaluation procedures. He notes that educators should strive to establish formal evaluation procedures. Formal procedures are objective and supply data that enable descriptions and judgments regarding the program being evaluated. According to Stake, educational outcomes are immediate and long-range, cognitive and affective, personal and community wide.

Stake delineates three data categories: antecedents, transactions, and outcomes.

An antecedent is any condition that exists prior to teaching and learning that may influence outcomes.

Transactions occur between students and teachers, students and students, and students and resource people. Transactions also include students, interactions with curriculum materials and classroom environments.

Outcomes are the program's results, including student achievement and, sometimes, attitudes and motor skills; impact on teachers' perceptions of their competence; and influence of administrators' actions.
CIPP Model of Evaluation

Mallory Buzun-Miller & Rachel Mainero
Humanistic models

- Humanistic evaluators recognize that individuals have different values, abilities, and experiences and, therefore, different perceptions of "reality." They argue for a more holistic approach to evaluation, one that provides detailed portraits of the situations being evaluated.

- The approach focuses more on human interactions than on outcomes, and more on the quality than the quantity of classroom or school life. Humanistic evaluators delve into the why, behind the what of performance. The merit, and interpretative understanding rather than objective explanation.
Eisner's Connoisseurship and Criticism Model

Who is Elliot Eisner?
Elliot Eisner has recommended two humanistic evaluation models—connoisseurship and criticism—that draw heavily from the arts.

Eisner describes connoisseurship as a private act engaged to personally appreciate the qualities that constitute some objects, situation, or event. Connoisseurship has essentially five dimensions: (1) intentional, (2) structural, (3) curricular (4) pedagogical, and (5) evaluative.

Intentional evaluation refers to a personal assessment of a curriculum's value, merit, and worth. Structural evaluation assesses the curriculum's design and the school's organization. Curricular evaluation assesses a curriculum's specific contents and how they are organized and sequenced. Pedagogical evaluation assesses instructional design and teaching strategies.

Eisner makes the point that evaluation should include reporting to the public (parents, school boards, local or state agencies, and so on). Evaluators must communicate the educational scene.
Illuminative Evaluation Model

- **Illuminative evaluation, sometimes called explication.**

  Developed by Malcolm Parlett and David Hamilton, this approach illuminates an educational program's specific problems and unique features.

- Illuminative evaluation allows evaluators to discern the total program as it exists and functions and to gather data about its particular workings. The evaluator determines the results of the taught curriculum and identifies assumptions evident in its delivery; the attitudes and dispositions of teachers, students, and the public and the personal and material factors that facilitate or impede the program.

- **Illuminative evaluation has three steps: observation, further inquiry, and explanation**

  The illuminative approach is holistic and subjective. Observed interactions are not broken down into discrete categories for measurement, but considered within the context of their environment.
1. **Observation.** Evaluators get an overview of the program and describe the context within which the curriculum is being delivered, considering all factors that might influence the program.

2. **Further inquiry.** Evaluators separate the significant from the trivial and seek to determine whether the program works and why or why not.

3. **Explanation,** Evaluators' explanations are presented to the people affected by the program, who then make decisions.
Action research is an evaluative approach that blends the scientific and humanistic. It is concerned with continual modification of the educational experience so that every educational event is fresh.

Action-research evaluation is distinguished by direct participation in the curriculum. Teachers are the key players in action-research evaluation.
Evaluation Steps

1. **Focus** on the curricular phenomena to be evaluated. Evaluators determine what they are going to evaluate and what design they will use. One particular subject area within a school for example, the total school system, or one particular subject area within a school.

2. **Collect** the information. Evaluators identify the necessary information sources and the means by which they can collect the information.

3. **Organize** the information. Evaluators organize the information so that the intended audience will be able to interpret and use it.

4. **Analyze** the information. Evaluators select and employ analysis techniques appropriate to the evaluation's focus.

5. **Report** the information. Evaluators decide the nature of the reporting, keeping in mind the report's audience. They might engage in informal reporting, such as giving opinions and making judgments based on general perceptions.

5. **Recycle** the information. The need for current information requires continuous re-evaluation.
Testing

- High-Stakes Tests

- Education is expensive. The public is increasingly concerned with getting the most for their money. Often, teachers and administrators feel that test results could jeopardize their jobs and their schools' reputations. In such cases, teachers may teach to the test, (often at administrators' urging).

- Ideally, however, tests should supply data by which educators can judge student performance, the curriculum appropriately and the effectiveness of the instructional delivery system. Although this approach usually boosts test scores, one wonders if such scores are evidence of higher-quality learning. What do the resulting scores actually tell you? Tests fail to tell us if a student is a good learner, let alone a good person, especially standardized tests.
Norm-Referenced Tests

- **Norm-referenced tests (NRIs) are the most commonly used.** A student performance on a particular test is compared with that of other students.

- However, such tests have questionable value for measuring the quality of a curriculum or the instruction. They do not relate specifically to a particular curriculum's goal and content. They do not effectively measure what has been taught. They do not indicate what a student can or cannot do. They do not even provide evidence that a student knows or does not know specific content. Nor do such tests furnish evidence of what a student or teacher needs to do to improve the student's level of knowledge and competence.
James Popham indicates that standardized achievement tests cannot detect the differences between students taught effectively and students taught ineffectively. Research indicates that standardized achievement tests highly correlate with students' socioeconomic status. Despite these limitations, educators continue to employ standardized tests to determine the curriculum's success and evaluate teaching effectiveness.

**Criterion-Referenced Tests**

The CRT is designed to indicate how a student performs a skill or task, or understands a concept, with respect to a fixed criterion or standard. This specificity enables educators to determine precisely what a student does or does not know, or can or cannot do, in relation to a particular curriculum. Criterion-referenced tests indicate changes in learning over time (in contrast, norm-referenced tests measure learning at a specific time). Perhaps the most serious criticism of criterion-referenced tests is that most lack information regarding their reliability.
Subjective Tests

- This essentially means that the test questions have one correct answer.

- However, curriculum evaluators also have access to subjective (constructed-response) tests. These tests have many correct responses to each question. For this reason, they are much more challenging to score than objective tests. Often, it is the depth or creativity of the response that determines the evaluative ranking.
Alternative Assessment

There many new forms of assessing open-ended tasks; students are required to use their knowledge and skills to create a product or solve a problem. Such evaluation events are called performance assessment. Performance assessment and authentic assessment, both examples of alternative assessment, as they employ methods other than multiple choice or like-developed objective tests. Authentic assessment includes real problem-solving, designing and conducting experiments on real problems. Authentic assessment employs strategies and approaches that present students with real-life situations and conditions. Dennis Wolf and Sean Reardon caution, "If new forms of assessment are to work, Educators must conceptualize intelligence, rethink what it means to know something, redefine excellence, and rethink their measurement habits.

- William Glasser has proposed seven features of optimal assessment.
- **First**, should foster student growth.
- **Second**, it should allow us to see the consequences of instructional effects.
- **Third**, the assessment should illuminate the processes and products of learning.
- **Fourth**, it should involve student self-assessment.
- **Fifth**, assessment should be an integral part of group activities.
- **Sixth**, assessment should entail meaningful tasks that tie in to overall learning and the curriculum's knowledge goals.
- **Seventh**, assessment should be comprehensive, addressing a broad range of information and skills rather than centering on narrow understandings of a particular content. Alternative assessment should be an ongoing activity integral to curriculum enactment,
New assessment methods require **new assessment criteria**. George Hein would include a moral standard among indicators of effective schooling. A school curriculum that **meets a moral standard provides students with skills and knowledge requisite for contributing to the general social good**. As Hein indicates, moral purpose was central to the progressive educational philosophy. The portfolio is perhaps the most popular method of alternative assessment.
Human issues of evaluation

- Many evaluation reports, valid in all technical details, have failed because of interpersonal in sensitivity ethnic or racial bias, or political naiveté, **the manner of presentation determines whether the evaluation results will be misused, or misinterpreted, or simply ignored.**

- Many aspects of curriculum evaluation are social, political, or ethical.

- Often schools release test results not to improve programs, but to please various power groups within the community or demonstrate to legislators that an educational program is effective. Sometimes test results are broadcast to convince various minority groups that their children are experiencing equity within the school system. Many evaluators are reluctant to confront issues of social justice within the educational system. **However, fairness is crucial.**

- According to James Pellegrino, Naomi Chudowsky, and Robert Glaser, the idea of comparable validity is at the core of fairness. **A fair test furnishes data from which one can draw valid inferences across individuals and groups.**
Evaluators and test designers realize that certain test items produce different results among students from different groups, even when all students have been matched in ability regarding the attribute or knowledge being assessed. For example, Students raised on farms are more likely than inner-city students to answer a question about agriculture correctly.

Also, is it fair to hold students with disabilities to the same standards as other students? Should students with reading and writing disabilities have to meet school standards in order to advance to the next grade or graduate? The issue of fairness also affects evaluation of students classified as gifted.

Several evaluators and assessment experts suggest that to really address, the issue of fairness, we must consider student’s backgrounds when we engage in evaluation. If we do this, we will be able to make conditional inferences from the data analyzed.
Evaluation should encourage, not intimidate, students. It should foster cooperation and a sense of community among students rather than feelings of tense. "Aggressive competition. Teachers should present tests as learning experiences, not as a means of reward and punishment.

Evaluation needs to address more than students' cognitive dimension. Tests and other means of evaluation need to gather data on the total person: cognitive, affective, emotional, and moral, we also need to realize that students and teachers function within various contexts and milieus. They interact with others inside and outside the school. We need to assess the amount and quality of such interactions and how they affect the curriculum, teaching, and learning.
Key Players in Evaluation

- Evaluation is a cooperative activity involving teachers, administrators, evaluators, students, and parents. Often it involves members of organizations outside of the school district.

- **Student** participation in evaluation empowers students to manage their own lives in socially acceptable, personally meaningful ways. Student self-assessment provides students and teachers with valuable data and nurtures autonomy in students.

- **Teachers** evaluate the success of their teaching and the curriculum within their classrooms. They may create journals in which they record the successes and failures of various curricula being attempted.

- **The evaluator** can be a member of the school system. Such a person knows the system and its goals. Disadvantages to having an insider as the key evaluator. An insider may not be willing to issue an evaluation report critical of the system. The evaluator designs the means of gathering data so that knowledge can be supplied to decision makers. The evaluator does not determine how the data will be used; instead, the evaluator helps the decision makers clarify their values.
Consultants

It is sometimes wise for a school district to hire an outside consultant to conceptualize the evaluation approach and coordinate the evaluation effort. Often small schools do not have any staff persons specifically trained in evaluation.

Parents and Community Members

Those affected by education should be involved in decisions about education, including curriculum. Parents and even community members without children should be involved in curriculum evaluation.
THANK YOU FOR LISTENING