

Illuminating Diverse Curriculum Evaluation Paradigms: Implications for Educational Practice

Serkan BOYRAZ¹

Curriculum evaluation has become a critical component of educational improvement efforts worldwide, serving as a foundation for informed decision-making, accountability, and continuous development. This study offers a comprehensive analysis of program evaluation models frequently applied in educational contexts, with a particular focus on their theoretical foundations, methodological approaches, and practical implications. Drawing from an extensive review of seminal and contemporary literature, the paper categorizes evaluation models into expertise-oriented, collaborative and stakeholder-based, logic and analytical, adversarial, qualitative and client-centered, postmodern and constructivist, and traditional approaches. Each model is examined in terms of its defining features, strengths, and limitations, as well as its alignment with different educational settings and goals. The findings underscore that no single evaluation model sufficiently addresses the diverse and complex needs of educational programs on its own. Instead, the integration or hybridization of different approaches is often necessary to capture both quantitative and qualitative dimensions of educational effectiveness. Moreover, the study highlights the increasing importance of participatory and empowerment-focused models, which promote inclusivity and foster stakeholder ownership. This review contributes to the field by offering a synthesized framework that can guide educators, policymakers, and researchers in selecting or designing context-appropriate evaluation strategies. This comprehensive perspective advances the discourse on educational evaluation and encourages the adoption of models that not only measure outcomes but also support program improvement and transformative change.

© IJERE. All rights reserved

Keywords: curriculum development, curriculum evaluation, curriculum philosophy,

INTRODUCTION

Curriculum is not a static entity but a dynamic and evolving construct that changes, develops, or is completely redesigned in response to various factors, such as becoming outdated, inadequate, impractical, or the emergence of more effective alternatives. Consequently, curriculum and curriculum development are inherently interrelated processes that are carried out simultaneously (Karagöz, 2016). In literature, curriculum development has been defined in multiple ways. Demirel (2012) describes it as the set of dynamic relationships between the objectives, content, educational situations, and evaluation elements of the educational program. Similarly, Varış (1989) defines it as a field of study that encompasses both the theoretical principles and practical activities necessary for determining and achieving program objectives in an effective manner. Lunenberg (2011) offers a more straightforward definition, describing curriculum development as the process of planning, implementing, and evaluating a curriculum with the aim of producing a coherent curriculum plan.

Philosophy, in the process of curriculum development, provides educators—particularly curriculum specialists—with a foundational framework for organizing schools and classrooms. It helps them address fundamental questions such as the purpose of schooling, which subjects are of greatest importance, how students learn, and which methods and materials should be employed. In this regard, philosophy plays a critical role in determining and organizing the goals and subject matter of education, structuring the teaching-learning process in the classroom, and guiding the selection of appropriate learning experiences and activities (Ornstein, 1990). For an overview of educational philosophies, see Table 1 in Ornstein and Hunkins (1988):

¹Faculty of Education, Aksaray University, Aksaray, Türkiye, serkanboyraz@aksaray.edu.tr, orcid.org/0000-0001-6483-1397

Table 1. Overview of Educational Philosophies and Their Instructional Implications

Educational Philosophy	Philosophical Foundation	Instructional Goals	Nature of Knowledge	Role of the Teacher	Curriculum Focus	Related Curriculum Orientation
Perennialism	Realism	To cultivate rational thinkers; to promote intellectual reasoning	Focused on enduring and universal truths; acquisition of timeless, factual knowledge	Facilitates logical thinking; uses Socratic dialogue; direct teaching of traditional values	Classical subjects; literary analysis; fixed curriculum	Great books; Paideia (everyone learns)
Essentialism	Idealism, Realism	To foster intellectual development; to educate competent individuals	Emphasizes essential skills and academic knowledge; acquisition of core concepts and principles	Authoritative in subject matter; direct instruction of traditional values	Core skills (e.g., reading, arithmetic, science, history, foreign language); essential academic subjects	Back to basics; academic excellence
Progressivism	Pragmatism	To promote democratic and social living	Knowledge supports growth and maturation; learning is an active, interest-based process	Acts as a guide in problem solving and scientific inquiry	Student-centered; addresses real-life human problems and interactions; interdisciplinary, project-based learning	Humanistic education; integrated curriculum; alternative and free schooling
Reconstructionism	Pragmatism	To advance and reconstruct society; to use education as a tool for social change and reform	Knowledge comprises skills and content needed to address societal issues; learning is active and future-oriented	Change agent and reform facilitator; acts as project leader and researcher; raises awareness of global human issues	Emphasizes social sciences and research methods; focuses on social, economic, and political issues; includes national and global concerns as well as future trends	Social justice in education; cultural diversity; international education; futurism

Traditional and modern educational philosophies differ significantly in their foundational beliefs, values, and approaches to teaching and learning. Traditional philosophies, such as perennialism and essentialism, are grounded in realism and idealism and emphasize the superiority of the past, viewing education as a process of transmitting fixed, objective, and often universal values and knowledge (Dilekli & Karagöz, 2018). Learning is perceived as the acquisition of discipline-based content, and the teacher functions as an authoritative figure who directs activities and delivers content, often using methods such as Socratic dialogue. In this model, homogeneous grouping, emphasis on academic excellence, and a focus on cognitive development are prominent, and education is often oriented toward preserving cultural norms and ensuring societal conformity (Ornstein & Hunkins, 1988).

In contrast, modern philosophies such as progressivism and reconstructionism are rooted in pragmatism and emphasize the reconstruction of current experiences, future concerns, and the need for societal change. These perspectives view values as subjective and changeable, and education as a creative, learner-centered process where knowledge is actively reconstructed. The teacher's role becomes that of a facilitator and co-learner, encouraging inquiry, student-initiated activities, and dialogue. Heterogeneous grouping, respect for diversity, and the integration of social issues into learning are fundamental. Rather than merely preparing students to conform, modern philosophies aim to empower individuals to transform their social contexts. They stress moral and social development, active citizenship, and democratic participation. Freedom is not a product of obedience and discipline but emerges through creativity, critical thinking, and lived democratic experience. Moreover, while traditional philosophies favor merit-based academic excellence, modern approaches advocate for equity in education, proposing differentiated support for disadvantaged groups to achieve fairer outcomes (Ornstein & Hunkins, 1988).

Curriculum Approaches and Design Perspectives

According to Ornstein and Hunkins (2014), our approach to curriculum reflects our values, knowledge, and perspectives on school and society. Curriculum approaches are broadly categorized into technical/scientific and non-technical/non-scientific paradigms. The behavioral approach, as a technical model, relies on logic, structured models, and step-by-step strategies. The managerial approach emphasizes the planning and organization of personnel, materials, and schedules, focusing on the administrative and operational aspects of curriculum implementation. The systems approach treats the school as an interconnected unit, planning processes to influence behavior and emphasizing feedback and comprehensive planning. Non-technical approaches include academic, humanistic, and problem-centered designs. The academic approach is analytical, often historical or philosophical, and emphasizes conceptual understanding. It gained traction in the mid-20th century but later declined with the rise of discipline-specific and qualitative approaches. The humanistic approach, inspired by progressive and child-centered philosophies, values experiential, artistic, and student-interest-driven learning activities, especially in primary education.

Demirel (2012) classifies curriculum designs into three major categories: subject-centered, learner-centered, and problem-centered. Subject-centered designs, the most widely used, are further divided into subject design, discipline design, broad-field design (which integrates subjects like geography and history into social studies), and process design (focusing on thinking strategies like problem-solving and decision-making). Learner-centered designs emerged in the early 20th century, emphasizing the student as the core of the curriculum. These include child-centered, experience-centered, romantic (radical), and humanistic models, each prioritizing the learner's individuality, interests, and psychological complexity.

Problem-centered designs aim to address unmet societal needs and individual development within the social context. These include life-situation designs (encouraging real-world problem solving), core designs (integrating disciplines while focusing on societal issues), and social reconstruction designs (linking curriculum development to sociopolitical and economic reform). The systems approach, increasingly prominent with the advent of technology, integrates input-process-output-feedback mechanisms into curriculum design. Based on Wulf and Schave's (1984) model, this approach involves problem identification, development of materials and strategies, and ongoing evaluation with feedback loops. Additionally, Varış (1989) views curriculum development not as a static document but as a continuous process encompassing extra-curricular activities, guidance, epistemology, learning theories, methodology, teacher roles, and feedback for improvement, and emphasizes the need for a centrally coordinated structure with national curriculum committees and local experts, arguing that success depends on coordinated leadership among specialists and administrators.

Curriculum Theories: Perspectives and Functions

Setting aside the ongoing debates about whether program theories truly exist, there are three primary perspectives regarding the purpose and function of such theories. The first and most prominent group views theory as a guiding framework for applied curriculum development and research, serving both as a practical tool and as a means of evaluating curricular initiatives. A second, smaller and relatively younger group of

theorists adheres more closely to traditional conceptions of scientific theory, seeking to identify and explain variables and relationships within the curriculum. Their aim is inherently conceptual, emphasizing the empirical investigation of curriculum variables and their interrelationships to establish theoretical validity. A third group, however, approaches theorizing not as a prescriptive or experimentally testable endeavor, but as a creative intellectual activity. Their goal is to construct conceptual schemes that enable new ways of thinking and talking about curriculum (Project 2061, 2013).

According to Morris and Hamm (1976), program theory is characterized by three main features. First, program theory holds ontological assumptions; its ultimate concern is not how or what to teach or learn, but rather what education is and why it exists. Its primary focus lies in the nature of knowledge—whether oriented toward process or product—rather than instructional or learning strategies. Second, program theory explores alternative intellectual structures for organizing knowledge. Each theoretical framework implicitly seeks to answer the foundational question: “What constitutes the most valuable knowledge?” Third, unlike the microcosmic nature of classroom-level curricula, program theory operates at a macrocosmic level, aiming for abstraction and universality. Classroom practices, in this view, are considered localized applications or adaptations of broader, more abstract program theories to specific contexts such as schools or school districts.

Curriculum Evaluation Models

Education is fundamentally a scientific activity, and program evaluation models employ scientific methods. Therefore, the methods used in quantitative and qualitative research approaches are also utilized in educational contexts (House, 1978). The quantitative research method is based on the philosophy of positivist science, which assumes that there is an order in the universe that can be discovered, understood, and controlled by humans. This method relies on a realist philosophy, asserting that knowledge is obtained through experiments and observations according to the rules of logic. The dominant reasoning method is induction, which involves moving from individual facts to generalizations. Fundamental research aims to develop theories related to a field or to test and strengthen existing theories. Applied research focuses on solving current problems and is conducted by authorities, responsible individuals, or scientists. Laboratory research is conducted in artificial, controlled environments, whereas field research occurs in natural settings without artificial constraints. Historical research emphasizes identifying the location, time, people, and institutions associated with events and proving the cause-effect relationships between these factors. Descriptive research seeks to describe and explain what exists or what is currently occurring. Correlational research aims to identify relationships among variables. Causal-comparative research investigates the causes affecting the outcomes of already occurred events. Finally, experimental research involves determining the effect of an independent variable on a dependent variable within a controlled and organized environment, resulting in empirical knowledge.

The qualitative research method, on the other hand, is based on interpretivist and postmodern scientific philosophies. It incorporates post-positivist and postmodernist understandings of science and involves collecting, analyzing, and interpreting qualitative data to arrive at judgments. Intuitive insights and deep understanding are also considered forms of evidence within this approach. Qualitative research methods include ethnographic research, where the researcher becomes part of the community, group, individual, or culture being studied. Narrative research involves describing real or potential events by specifying location, person, and time. Grounded theory research aims to uncover theories embedded within data. Phenomenology, inspired by Husserl’s philosophy that absolute reality does not exist, seeks to reach the essence of phenomena by excluding the effects of the investigator and external influences. Case studies focus on investigating a complex, unique, or interesting phenomenon within its own context. Root analysis examines and connects multiple points in diverse ways. Document analysis entails the collection and examination of written and visual materials.

Types of Evaluation

Evaluation can be categorized based on the criterion used for comparison and the intended purpose. According to the comparison criterion, evaluations can be norm-referenced or criterion-referenced. Because norm-referenced evaluations involve comparing individuals to each other and making selections among them, criterion-referenced evaluations are generally considered more consistent in program evaluation.

Diagnostic evaluation is conducted before students begin a program to identify prerequisite cognitive behaviors, affective characteristics, and psychomotor skills. Formative evaluation, according to Demirel (2012) is implemented during the program to identify learning difficulties and make necessary adjustments. This type of evaluation provides continuous feedback to the program, creating a control system to implement corrective measures. Summative evaluation, conducted at the end of a program, measures the acquired behaviors, characteristics, and skills of students. This type of evaluation allows judgments to be made about whether the educational program effectively imparts the intended behaviors to students.

Formal evaluation refers to evaluations that are systematic, structured, and purpose-driven, where details such as the location, timing, and evaluators are predetermined, and the data collection tools used have high reliability and validity (Fitzpatrick et al., 2004). In contrast, informal evaluation lacks systematic procedures and uses irregular data collection techniques, often relying on personal judgment. Experience, intuition, generalizations, and logic significantly influence informal evaluation outcomes (Yüksel & Sağlam, 2014).

Formative evaluation aims to provide data necessary to improve and enhance the quality of a program. Conducted during program implementation, it gives administrators and practitioners the opportunity to take corrective action before the program is completed. This type of evaluation is essential for establishing objectives, determining priorities, implementing draft plans, and assessing interim internal evaluation results (Yüksel & Sağlam, 2014). Stake succinctly differentiates between formative and summative evaluations by stating: "When the cook tastes the soup, it is formative evaluation; when the guests taste the soup, it is summative evaluation." Summative evaluation is conducted after program completion, for the benefit of external stakeholders or decision-makers (such as sponsoring organizations or potential future users) and may involve internal, external, or both types of evaluators (Madaus & Kellaghan, 2002). Summative evaluation, also called "total evaluation," is performed after the program has been completed. The purpose is to use retrospective data to make decisions about the program's future, such as whether to continue or terminate it (Yüksel & Sağlam, 2014).

General Information on Evaluation Methods

Evaluation methods can be further understood through several models (Hansen, 2005). Outcome models include the goal-oriented model and the impact model. In the goal-based model, the primary question is the degree to which objectives have been achieved, while the impact model seeks to determine what effects can be identified. In both cases, results should be explicitly presented and derived from established objectives. The elucidative process model focuses on whether activity levels are sufficient and whether there are problems in implementation. The systems model evaluates overall program performance by comparing inputs, processes, structures, and outputs with goals or by direct comparison. Economic models, which include cost-adequacy, cost-effectiveness, and cost-benefit analyses, involve questions such as whether production, effectiveness, and the overall program are satisfactory. In these models, outputs, efficiency, and programs are evaluated in relation to costs.

Goal-Oriented Evaluation Approach

Several primary models are used in program evaluation, including the goal-oriented evaluation model developed by Ralph Tyler between 1933 and 1941. This model, based on objectives, learning experiences, and evaluation as core elements of program development, remains widely utilized. In this model, quantitative data are employed, and it is necessary to measure the degree of goal attainment at least twice—at the beginning and end of instruction. For cognitive objectives, achievement tests based on objectives can be used; for psychomotor skills, observation forms are employed; and for affective objectives, attitude scales can be utilized. In this model, the primary objective is to identify the discrepancy between the expected objectives and the actual student performance, in order to address deficiencies within the program and utilize the findings to inform the program development process. This model, which Tyler describes as evaluating the adequacy of the product resulting from the teaching-learning process and defined as student performance, is fundamentally based on the changes and developments observed in student behaviors in line with predetermined objectives (Yüksel & Sağlam, 2014). In outcome-based evaluations, quantitative research methods, particularly descriptive and experimental designs, are generally utilized as appropriate (Sönmez & Alacapınar, 2015).

The Metfessel-Michael Evaluation Model, which was developed under the influence of Tyler's model, differentiates itself by placing greater emphasis on the selection of alternative measurement tools (Kuo et al., 2012) and by comparing performance against established standards rather than predetermined objectives (Yüksel & Sağlam, 2014). Metfessel and Michael outlined the evaluation process in eight steps (Ornstein, 1990). First, teachers, administrators, students, and even ordinary citizens from the educational community should be involved directly or indirectly in the evaluation process. Next, a hierarchy of objectives should be developed, progressing systematically from general to specific, to create an intensive paradigm of goals. The specific objectives established in this stage should then be transformed into implementable forms within the program. Subsequently, measurement tools capable of assessing the program's effectiveness on individuals, in light of these objectives, should be developed. Throughout the program implementation, regular observations should be conducted using tests and other appropriate instruments. The information collected during this process should be analyzed systematically. Furthermore, standards and values that can be utilized to philosophically evaluate the program should be articulated clearly. Finally, based on the analyzed data, recommendations should be developed regarding the future applicability and implementation of the program. This model, based on systems theory, divides the evaluation process into five stages and four components, each of which involves comparing actual performance with established standards (Nyre & Rose, 1979). The components of this model include determining program standards, identifying program performance, comparing performance with standards, and identifying whether there is a discrepancy between performance and standards. Depending on these identified discrepancies, decision-makers can choose to move to the next stage, revise and reuse a previous stage, restart the program, reorganize both performance and standards, or terminate the program altogether.

Within the stages of program evaluation, the adequacy of the program is assessed against the predetermined standards. The first stage, design, involves comparing the program design in terms of the prepared criteria or standards. The second stage, construction, focuses on evaluating the program elements referred to as resources, methods, and student behaviors. The third stage, processes, involves evaluating student and staff activities, functions, and relationships. The fourth stage, product or outcome, entails an overall assessment of the program by taking the original objectives into account. Finally, in the fifth stage, the outputs of the program are compared with those of similar programs. This model distinguishes itself from others by focusing more on evaluating the program design during the program development process. Defined as the description of discrepancies, inconsistencies, and differences between the expected and the actual outcomes, the core procedure of this model is the comparison between performance and standards (Yüksel & Sağlam, 2014). Based on the results of this comparison, four potential decisions can be made: the program may be terminated, modified, repeated as it is, or the standards may be revised (Nyre & Rose, 1979).

The Goal-Free Evaluation Model, developed as an alternative to the goal-oriented program evaluation approach, possesses several key characteristics (Fitzpatrick et al., 2004). In this model, the evaluator consciously avoids becoming aware of the program's objectives. Predetermined goals are not allowed to narrow the focus of the evaluation. Rather than concentrating on expected outcomes, goal-free evaluation emphasizes actual outcomes. Furthermore, the evaluator maintains minimal communication with program administrators and staff, which increases the likelihood of identifying unexpected outcomes. This approach enhances the possibility of discovering the effectiveness of unanticipated goals.

The Hammond Program Evaluation Model not only addresses whether the implemented program has achieved its objectives but also explores why certain goals were not accomplished while others were successfully met. Aiming to assist evaluation specialists in investigating factors influencing the success or failure of educational activities, Hammond developed a three-dimensional "evaluation cube" designed to define educational programs and organize variables related to evaluation (Uşun, 2012). The first dimension of the cube, instruction, refers to the characteristics of the educational activity being evaluated. This includes organization (e.g., scheduling, sequencing of courses, school organization—whether vertical or horizontal structures), scope (the subjects to be covered), methods (instructional activities and types of interactions, such as teacher-student or media-student, as well as learning theories), services (space, equipment, consumable materials), and costs (funding for services, maintenance, and staff). The second dimension, institution, relates to the characteristics of individuals and groups involved in the educational activity. For students (the first column of the cube), factors include age, grade level, gender, family background, social class, health, mental

health, achievement, abilities, and interests. For teachers, administrators, and educational specialists (the second, third, and fourth columns), relevant variables include age, gender, race and religion, health, personality, educational background, work experience, and appropriate personal traits. For families (the fifth column), considerations include the degree of involvement in the activity, culture or language, family size, marital status, income, educational level, and affiliations (e.g., religion, political orientation, unions). The community (the sixth column) is evaluated in terms of geographic location, history, demographic, economic, social, and political characteristics. The third dimension, target behaviors, focuses on the objectives of the educational activity. These include cognitive objectives (knowledge and intellectual skills), affective objectives (interests, attitudes, emotions, and enthusiasm), and psychomotor objectives (physical skills and coordination).

The Bennett Program Evaluation Model is a goal-oriented model designed to evaluate the success of farmer education programs through data collected across seven stages (Uşun, 2012). First, program inputs, which are the resources identified in detail by the implementing organization, must be defined. Second, the activities to be included in the educational program for the target audience should be determined. Third, public participation involves selecting participants according to open and democratic principles, preventing negative attitudes toward the program and its organizers. Fourth, participants' reactions are assessed to measure their responses to both the educational activities and the experts delivering the program. Fifth, changes in knowledge, attitudes, skills, and aspirations (referred to as KASA) among participants are evaluated. Sixth, changes in practice are assessed by examining how participants apply the acquired knowledge, attitudes, skills, and behaviors to their work and daily lives. Finally, the outcomes stage investigates the potential changes that occur within the target group as participants begin to implement program content in their professional and daily activities.

The Participant-Oriented Evaluation Approach, exemplified by Stake's Responsive Evaluation Model, encourages evaluators to collect and process deeper types of data, consider the dynamics among people involved in the program, evaluate the various roles assumed by different stakeholders, allow greater participation in decision-making processes, and adopt a stance regarding the value of the program (Ornstein & Hunkins, 2014). According to Stake, evaluation-related data can be organized into three dimensions (Demirel, 2012). The input dimension refers to existing factors prior to the teaching-learning process that influence outcomes. The process dimension focuses on interactions among teachers and students, among students themselves, and between students and resource people. This dimension encompasses the learning-teaching process, including factors such as time allocation to different sections, the nature of teachers' questions, and students' activities in laboratories. Lastly, the output (or product) dimension pertains to evaluations of academic achievement, attitudes, and skill levels. Thus, data should be collected about pre-implementation conditions (e.g., student attitudes, levels of achievement, participation, and teacher attitudes and experience); the applications involving students, teachers, materials, and the environment during the learning-teaching process; and the cognitive, affective, personal, social, and both short- and long-term outcomes of the program (Bhattacharyya, 2011). The process itself concerns what happens directly in relation to the educational program, whether in the classroom, laboratory, library, home, or other settings. It involves considerations such as the time spent on different parts of the program, the questions posed by the teacher, and what students do during lab activities. In summary, this approach addresses not only what is taught but also how it is taught (Baumgart, 1972).

The purpose of the Illuminative Evaluation Model developed by Parlett and Hamilton is to investigate and reveal how a new program operates, the ways it affects different schools, the variations in implementation across schools, and both the positive and negative aspects of these differences. Additionally, the model examines how these factors influence students' academic, technical/vocational, and other developmental outcomes. In this model, the researcher utilizes three types of data collection tools and processes: observation, advanced research, and explanation phases (Uşun, 2012). In Stake's Responsive Evaluation model, on the other hand, evaluators focus primarily on assessing program activities and processes rather than merely examining outcomes. The evaluator describes the narrative and characteristics of the program, identifies the clients and staff, highlights important issues and problems, and reports on successes (Demirel, 2012). Stake emphasizes the importance of stakeholder decisions in the program evaluation process and asserts that stakeholder participation is an indispensable component of evaluation studies. According to Stake, there are two

fundamental actions in evaluation: description and judgment (Yüksel & Sağlam, 2014). Stake named his approach "responsive evaluation" to emphasize the dimension of being responsive to and aligned with the interests and concerns of stakeholders. In this context, he proposed managing the program by considering four dimensions: context, working area, output, and support. The stages of responsive evaluation, as described by Stake (1977), include talking with clients, program teams, and audiences; defining the scope of program effectiveness; presenting a general overview of program activities; exploring objectives and interests; conceptualizing issues and problems; specifying data requirements related to topics or problems; selecting observers, experts, and any available resources; observing predesigned preliminary findings, procedures, and results; making descriptions and preparing case studies; tailoring topics to the audience, critiquing, and eliminating undesired topics; shaping the overall purposes of the audience; and finally, collecting and compiling formal reports, if available.

System-Oriented Evaluation Approaches

Stufflebeam's Context, Input, Process, and Product (CIPP) Model

According to Stufflebeam, evaluation involves describing what kind of information will be collected, obtaining this information, and providing it to relevant groups (Ornstein & Hunkins, 2014). He views program evaluation as the process of defining, collecting, and applying descriptive and judgmental information concerning the program's objectives, design, implementation, and outcomes. The primary purpose of evaluation is not to prove the merit of a program but rather to continuously improve it (Yüksel & Sağlam, 2014). Stufflebeam considers evaluation as a continuous process. In the context of program development, decision-makers must make planning, structuring, implementing, and revising decisions, all of which require the collection of information based on the evaluation of four stages: context evaluation (collecting information that serves as a foundation for determining objectives), input evaluation (analyzing the program and its components at a micro level), process evaluation (examining the alignment between planned and actual activities during implementation), and product evaluation (comparing expected and actual outcomes) (Demirel, 2012).

Alkin's UCLA Evaluation Model

This model categorizes evaluation types as system evaluation (providing information about the current state of the system in which the program is implemented), program design (selecting programs that meet educational needs), program implementation (determining whether the program is delivered to the target audience as intended), program development (evaluating whether objectives are achieved and whether unexpected outcomes have emerged during implementation), and program certification (providing information on the program's value) (Fitzpatrick et al., 2004).

Stufflebeam's Total Evaluation Model

This model includes planning following context evaluation, structuring after input evaluation, implementation following process evaluation, and recycling decisions after product evaluation. Decisions vary depending on the type and amount of information: major changes based on little information indicate dynamic change; a series of small changes based on little information indicates incremental change; small changes based on extensive information reflect static change; and large changes strongly supported by extensive information represent transformational change (Demirel, 2012).

Dick and Carey's Instructional Design Model

Dick and Carey (2005) describe instructional design as a systematic process that involves components such as teachers, students, activities, and their implementation to ensure effective learning. This model includes defining instructional goals, conducting instructional analysis, identifying entry behaviors and learner characteristics, developing objectives, creating assessment tools, selecting instructional strategies, developing instructional materials, conducting formative evaluation, and performing summative evaluation. The model employs a systems approach, allows flexibility for environments and users, meticulously applies stages from objectives to final products, is effective especially in primary education, resembles approaches used in software engineering, and can be exemplified by one-way educational TV or video. In this approach, the teacher acts as both initiator and moderator of communication (Uşun, 2012).

Kirkpatrick's Training Evaluation Model

Widely used in adult and in-service training, this model consists of four stages: reaction (measuring participants' responses to the program), learning (assessing changes in knowledge, attitudes, and skills), behavior (examining behavioral changes resulting from the training), and results (determining concrete outcomes such as increased productivity, improved quality, reduced costs, or fewer accidents) (Fitzpatrick et al., 2004). Evaluation should start with the first stage and proceed to subsequent stages as conditions allow. Participants' feedback should be reviewed carefully, and suggestions for future improvement should be implemented. The principle "If the customer is not satisfied, it is my fault because it is my job to satisfy them" highlights the importance of participant satisfaction. While easy to apply at the individual level, measuring results at the organizational level is challenging due to numerous variables affecting long-term performance (Yüksel & Sağlam, 2014).

Realistic Evaluation Model

Mark et al. (2000) developed the theoretical foundation of this model, emphasizing realism to enhance validity. This model explores three research areas and related questions: functioning (what mechanisms lead to specific outcomes in a given context?), context (what conditions must be met for these mechanisms to be activated?), and sample outcomes (what practical benefits emerge from these mechanisms in specific contexts?).

Saylor, Alexander, and Lewis Model

This model adopts a systematic approach to curriculum development and consists of five components: goals and sub-goals, the educational program as a whole, specific program elements, instruction, and the evaluation program. The model relies on feedback systems, where each program element is evaluated separately and the findings are returned to the respective units. Initial stages involve defining general and specific goals, guiding curriculum areas, and addressing individual development, human relations, learning, and specialization. Following this, program designers focus on organizing content and designing appropriate learning experiences, develop instructional plans, and implement the program. Teachers select suitable methods and materials, and ultimately, comprehensive evaluation provides feedback on maintaining, revising, or discontinuing the program (Saylor et al., 1981).

Total Quality Evaluation Model

This model evaluates teaching processes, program components, teachers, students, methods, societal and environmental factors, theoretical foundations, and outputs. It examines how all elements interact, their relationships, principles, and qualities through a systems approach.

Walker's Evaluation Design Model

Walker's model (2003) consists of eight steps: defining the purpose of the evaluation, clarifying project objectives, developing a change model, selecting criteria and indicators, identifying data sources and measurement frequency, designing the evaluation study, observing and evaluating, and using and reporting results.

Expertise-Oriented Evaluation Approaches

Eisner's Educational Connoisseurship and Criticism Model (1976) is implemented by experts and requires the collection, interpretation, and evaluation of qualitative information about program outcomes after implementation. According to Woods (1988), the model comprises three stages: description, interpretation, and evaluation. The description stage focuses on elucidating the qualities and characteristics of the educational experience and is considered the simplest phase as it does not involve decision-making. In the interpretation stage, possible outcomes resulting from program implementation are anticipated. Finally, in the evaluation stage, judgments are made about the program based on the outcomes of the descriptive and interpretive stages. Eisner identified five dimensions of expertise: intentional (evaluating the value and contributions of the educational program), structural (evaluating program design and school organization), curricular (evaluating specific content and its organization and sequencing), pedagogical (evaluating instructional design and strategies), and evaluative (assessing the evaluation process itself).

In this model, the program evaluator is likened to an educational art connoisseur, and the evaluation process resembles art criticism (Eisner, 1976). Thus, the evaluator first describes what they observe, then interprets it, and finally makes evaluative judgments. The descriptive dimension provides information about, for example, the types and frequency of questions posed in class, the duration of discussions, and the lasting impressions left on students. The interpretive dimension relates to understanding the meaning and significance of social interactions within the classroom context. The evaluative dimension involves assessing the educational importance and impact of the interpreted experiences or activities. The Expert/Accreditation Model asserts that expert opinion is necessary to make judgments about a program's quality. Its primary goal is to evaluate the program from a professional perspective (Uşun, 2012).

Collaborative and Stakeholder-Based Evaluation Approaches

Stakeholder-Based Evaluation aims to provide feedback to stakeholders for program improvement. This model is typically used when stakeholder groups do not agree on program objectives and require the guidance of an evaluator to achieve consensus (Cousins & Earl, 1992). Here, stakeholder evaluators are experts in the program area and evaluation methods. They design the evaluation process, collect necessary data, and prepare reports based on the criteria and interpretations provided by stakeholders (Uşun, 2012).

Empowerment Evaluation is a flexible and collaborative model focused on program improvement. This approach emphasizes using findings, techniques, and evaluation concepts to promote development and strengthen autonomy. Rather than solely judging a program post hoc, the model seeks to facilitate learning and change, shifting the evaluator's role from content expert to facilitator. Participants develop their own internal evaluations, while external evaluators often act as coaches based on program capabilities. This model involves three main phases: mission formulation, situational analysis, and future planning. Its key concepts include fostering, facilitating, advocacy, illumination, and liberation (Fitzpatrick et al., 2004).

Democratic Evaluation focuses on both empowerment and the utilization of evaluation findings, aiming to inform the entire public about the characteristics of an educational program. Its key concepts include confidentiality, negotiation, and accessibility, with the primary guiding principle being the "right to know." According to MacDonald (1976), democratic evaluation contrasts with autocratic and bureaucratic models. Because participants hold control over interpreting and clarifying data and results, this model is described as democratic. On the other hand, Utilization-Focused Evaluation emphasizes meeting decision-makers' needs and explaining how to make evaluation more useful to them. It starts with the premise that evaluations should be assessed based on their utility and actual use. According to (1997), evaluators must facilitate the evaluation process and act conscientiously. In this model, evaluators engage stakeholders in decisions regarding the design, purpose, and methods of evaluation, thereby increasing the likelihood that findings will be used. Evaluators are also responsible for fostering stakeholders' capacity to engage with and apply evaluation information.

Logic and Analytical Approaches

The Logic Model, increasingly popular in recent years, is a systematic and visual method used to present the rationale behind a program, its resources, planned activities, and anticipated outcomes. It provides a logical and rational framework to show how a program operates under specific contextual conditions to solve particular problems. Its core components include resources (human, community, financial, organizational), activities (methods, tools, technologies, events, actions), outputs (intended changes), and short-, medium-, and long-term outcomes (changes in participants' behavior, awareness, skills, status, and effectiveness) (Uşun, 2012). Development involves six stages: gathering new information, defining the problem and context, specifying model elements, describing logic in diagrams and text, validating and implementing the model with stakeholders, and using the model for observation and evaluation.

Demirel's Analytical Evaluation Model is built on two fundamental dimensions: document analysis and stakeholder perspectives. The first dimension involves analyzing and comparing current and previous programs based on predetermined criteria (context, goals, content, process, and evaluation). The second dimension focuses on collecting feedback from program beneficiaries and stakeholders regarding program implementation (Demirel, 2012).

Adversarial Evaluation Approaches

Wolf's Legal Model aims to help educational decision-makers confirm program challenges by presenting opposing viewpoints. This model involves four stages: issue identification, issue selection, evidence preparation, and hearing announcement. Wolf (1990) emphasizes an evidence-based and guiding process that facilitates fair decision-making and serves public interest. The Adversarial Model facilitates the use of various data sources and requires more investigative effort than traditional approaches. It involves planned evaluations where different evaluators or teams adopt opposing perspectives. While one team may present a highly positive view defending the program, another highlights deficiencies. The model concludes with a simple "yes or no" decision on whether the program should continue. The core question is: "Should the program be maintained or terminated?"

Qualitative and Client-Centered Approaches

Patton's Qualitative Evaluation Model defines program evaluation as a systematic process for collecting information about program activities, characteristics, and outcomes to make decisions, improve effectiveness, and provide future guidance. Patton (1997) emphasizes three key elements: qualitative data, a holistic and inductive design based on natural inquiry, and content or situational analysis. Scriven's Goal-Free Evaluation Model, within the client-centered approach, posits that evaluators should not take program goals as the starting point but rather determine which outcomes to investigate. This model focuses on actual rather than intended outcomes. Evaluators deliberately enter the field without being informed of the program's explicit aims and derive insights inductively (Alkin & Christie, 2013). Key characteristics include a focus on actual outcomes, minimal interaction with program managers and staff, the potential to reveal unintended effects, and no obligation to define or measure specific objectives (Scriven, 1994). The model's strength lies in sensitizing evaluators to the magnitude and significance of program effects.

Traditional Evaluation Approaches

Traditional Evaluation, emerging in the 1960s with the application of scientific methods to program evaluation, emphasizes reliability and validity of collected data. Objectivity, neutrality, and outcome-focused evaluation are fundamental expectations. Core features include a search for objectivity and scientific rigor within standard processes, reliance on quantitative measurements, high administrative control, exclusion of stakeholder influence, sole control by evaluators (often academics and professionals), and a summative purpose (Uşun, 2012).

Postmodern and Constructivist Approaches

Postmodern Evaluation emphasizes the interpretability of textual meanings, attention to binary oppositions (e.g., good–bad, modern–traditional), the deconstruction of these binaries to reveal their contingency, and the identification of omissions and distortions—whether intentional or unintentional. Guba and Lincoln's (1989) Fourth-Generation Constructivist Evaluation Model consists of five stages: measurement (e.g., intelligence testing), description (e.g., formal evaluation of programs), judgment of appropriateness and value, inclusion of social reality (constructivism), and meta-evaluation (evaluation of the evaluation). The process involves identifying stakeholders, collecting information about their perspectives and interests, designing and applying an inclusive method for all groups, fostering consensus, addressing unresolved issues through agenda-setting, collecting desired information, holding expert meetings with stakeholder representatives, preparing reports reflecting shared interpretations and solutions, and revisiting unresolved comments through iterative evaluations.

RESULT and DISCUSSION

This review revealed a remarkable diversity in program evaluation models, each with distinct philosophical underpinnings, methodological approaches, and intended uses. This diversity underscores the complexity of educational contexts and the need for tailored evaluation strategies that align with specific program goals, stakeholder expectations, and societal demands. The expertise-oriented approaches, such as Eisner's Educational Connoisseurship and Criticism Model, emphasize qualitative judgments by experts, drawing parallels to art criticism. This model's descriptive, interpretive, and evaluative dimensions allow for nuanced understandings of educational practices and outcomes. It suggests that evaluations should not only measure objective outcomes but also capture the subtle, often intangible, qualities of educational experiences.

Such models are particularly valuable for programs that emphasize creativity, holistic development, and complex social interactions.

Collaborative and stakeholder-based approaches, including the Empowerment Evaluation Model and Stakeholder-Based Evaluation, highlight the importance of participation, dialogue, and shared ownership in the evaluation process. These models foster democratic values and promote program improvement through collective reflection. They align with modern educational philosophies that prioritize inclusivity and responsiveness to diverse learner needs. The emphasis on stakeholder engagement also enhances the relevance and utilization of evaluation findings, ensuring that they inform future program modifications effectively. Logic and analytical models, such as the Logic Model and Demirel's Analytical Evaluation Model, emphasize systematic planning and evidence-based decision-making. By clearly outlining resources, activities, outputs, and outcomes, these models facilitate transparency and accountability. They are particularly suited for large-scale educational interventions requiring structured implementation and measurable impacts.

Adversarial approaches, exemplified by Wolf's Legal Model and the broader adversarial model, introduce an element of debate and critical scrutiny. By presenting contrasting perspectives, these models help uncover both strengths and limitations of a program, encouraging more balanced and informed decision-making. This approach is especially relevant in contentious educational reforms or when programs face substantial public scrutiny. The qualitative and client-centered models, such as Patton's Qualitative Evaluation Model and Scriven's Goal-Free Evaluation Model, prioritize actual program effects over predefined objectives. These models are crucial in contexts where unexpected outcomes or unintended consequences are significant. Their focus on emergent data and inductive reasoning aligns with contemporary views of education as a dynamic and context-dependent process. Finally, traditional evaluation approaches maintain their relevance in contexts that demand rigorous, objective, and summative assessments. Their focus on reliability, validity, and measurable outcomes provides critical data for accountability and large-scale policy decisions.

Overall, this review highlights that no single evaluation model can address all educational needs comprehensively. Instead, selecting or combining models based on program characteristics, stakeholder expectations, and contextual realities is essential. Future research and practice should emphasize hybrid evaluation designs that integrate the strengths of different models to provide a more holistic and impactful understanding of educational programs.

Declarations

Conflict of Interest

No potential conflicts of interest were disclosed by the author with respect to the research, authorship, or publication of this article.

Ethics Approval

Not required as no human participant was included in the research.

Funding

No specific grant was given to this research by funding organizations in the public, commercial, or not-for-profit sectors.

Research and Publication Ethics Statement

- This material is the authors' own original work, which has not been previously published elsewhere.
- The paper reflects the authors' own research in a truthful and complete manner.
- The results are appropriately placed in the context of prior and existing research.
- All sources used are properly disclosed.

REFERENCES

- Alkin, M. C., & Christie, C. A. (2013). An evaluation theory tree. M. C. Alkin (Ed.). In *Evaluation Roots* (p. 12-66). SAGE Publications, Inc.
- Baumgart, N. (1972). A model for curriculum evaluation. *Research in Science Education*, 2(1), 1–12.
<https://link.springer.com/article/10.1007/BF02558535>
- Bhattacharyya, D. (2011). *Department of Education*. Retrieved from University of Kalyani:
<http://www.kudbhattacharyya.com/pdf/cur-2011eva.pdf>

- Cousins, J.B., & Earl, L.M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, 14(4), 397–418. <https://doi.org/10.3102/01623737014004397>
- Demirel, Ö. (2012). *Eğitimde program geliştirme* (18th ed.). Pegem.
- Dick, W., Carey, L., & Carey, J. O. (2005). *The systematic design of instruction*. Pearson/Allyn and Bacon.
- Dilekli, Y., & Karagöz, S. (2018). İhsan Sungu'nun program değerlendirme modelinin Avrupa ve Amerika'da yaygın kullanılan program değerlendirme modelleri ile karşılaştırılması. *OPUS International Journal of Society Researches*, 8(14), 465-496. <https://doi.org/10.26466/opus.391195>
- Eisner, E. W. (1976). Educational connoisseurship and criticism: Their form and functions in educational evaluation. *Journal of Aesthetic Education*, 10(3/4), 135-150. <https://doi.org/10.2307/3332067>
- Erişen, Y. (1998). Program geliştirme modelleri üzerine bir inceleme. *Kuram ve Uygulamada Eğitim Yönetimi*, 13(13), 79-97. <https://dergipark.org.tr/tr/pub/kuvey/issue/10335/126656>
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2004). *Program Evaluation-Alternative Approaches and Practical Guidelines* (3rd ed). Allyn & Bacon Publisher.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Sage Publications, Inc.
- Hansen, H. F. (2005). Choosing evaluation models. *Evaluation*, 11(4), 447-462. <https://doi.org/10.1177/1356389005060265>
- House, E. R. (1978). Assumptions underlying evaluation models. *Educational Researcher*, 7(3), 4-12. <https://doi.org/10.3102/0013189X007003004>
- Karagöz, S. (2016). Cumhuriyet öncesi dönem eğitim ve öğretim programları ile ilgili görüş ve öneriler. *OPUS – Uluslararası Toplum Araştırmaları Dergisi*, 6(11) 651-676. <https://dergipark.org.tr/tr/download/article-file/232508>
- Kuo, L.-H., Wei, H.-M., Hsueh-Chih, L., Ho, M.-K., & Yang, H.-J. (2012). A model of evaluating integrating emerging technology into formal technology curriculum. *Latest Advances in Educational Technologies*, 59-65. <https://naun.org/main/NAUN/educationinformation/16-288.pdf>
- Lunenburg, F. C. (2011). Curriculum development: Inductive models. *Schooling*, 2(1), 1-8. <https://www.nationalforum.com/Electronic%20Journal%20Volumes/Lunenburg,%20Fred%20C.%20Curriculum%20Development-Inductive%20Models-Schooling%20V2%20N1%202011.pdf>
- MacDonald B (1976) Evaluation and the control of education. In: Tawney D (ed.) *Curriculum Evaluation Today: Trends and Implications*. Macmillan, 125–136.
- Madaus, G. F., & Kellaghan, T. (2002). Evaluation models. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 19–32). Kluwer Academic Publishers.
- Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. Jossey Bass.
- Morris, R. C., & Hamm, R. (1976). *Toward a curriculum theory*. Retrieved from ASCD: http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_197601_morris.pdf
- Nyre, G. F., & Rose, C. (1979). The practice of evaluation. *Professional and Organizational Development Network in Higher Education*, 1(3), 189-196. <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1019&context=podqtrly>
- Ornstein, A. C. (1990). Philosophy as a basis for curriculum decisions. *The High School Journal*, 74(2), 102-109.
- Ornstein, A. C., & Hunkins, F. P. (1988). *Curriculum: Foundations, principles, and issues*. Prentice-Hall.
- Ornstein, A. C., & Hunkins, F. P. (2014). *Eğitim programı: Temeller, ilkeler ve sorunlar* (A. Doğutaş, Trans.). Eğitim Yayınevi. (Original work published 2014)
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Sage Publications, Inc.
- Project 2061. (2013). *Project 2061*. http://www.project2061.org/publications/designs/online/pdfs/reprints/2_macdld.pdf
- Scriven, M. (1994). Product evaluation—The state of the art. *Evaluation Practice*, 15(1), 45-62. <https://doi.org/10.1177/109821409401500105>
- Sönmez, V., & Alacapınar, F. G. (2015). *Örnekleriyle eğitimde program değerlendirme*. Anı Yayıncılık.
- Stake, R. E. (1977). The countenance of educational evaluation. In A. A. Bellack & H. M. Kliebard (Eds.), *Curriculum and evaluation* (pp. 372–390). McCutchan.
- Uşun, S. (2012). *Eğitimde program değerlendirme süreçler yaklaşımlar ve modeller*. Anı Yayıncılık.
- Varış, F. (1989). Eğitimde program geliştirmeye sistematik yaklaşım. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 11(2), 7-13. https://doi.org/10.1501/Egifak_0000000866
- Walker, F. (2003). *Fundamentals of Curriculum Passion and Professionalism*. Lawrence Erlbaum Associates.
- Wolf, R. L. (1975). Trial by jury: A new evaluation method. *Phi Delta Kappan*, 57, 185-187. <https://doi.org/10.3102/01623737001003019>
- Woods, J. D. (1988). Curriculum evaluation models : Practical applications for teachers. *Australian Journal of Teacher Education*, 13(1), 3-8. <https://doi.org/10.14221/ajte.1988v13n2.1>
- Yüksel, İ., & Sağlam, M. (2014). *Eğitimde Program Değerlendirme* (2nd ed.). Pegem.