
The Historical Of Evaluation Program And Philosophy Assessment

Nurhayati¹⁾, Ambiyar^{2*)}, Ishak Aziz³⁾

1) Program Studi Doktor Ilmu Pendidikan Pascasarjana Universitas Negeri Padang

2,3) Universitas Negeri Padang

*Corresponding Author

Email : nurhayati@uin-suska.ac.id

Abstract

This study aims to identify and understand the historical of evaluation program and philosophy assessment. The method used in this study uses a library method or approach (library research). Data obtained through observation and literature study. The results showed that the historical of evaluation program called Age of Reform in 1900; the second, from 1900 until 1930, we call the Age of Efficiency and Testing; the third, from 1930 to 1945, may be called the Tylerian Age; the fourth, from 1946 to about 1957, it called the Age of Innocence; the fifth, from 1958 to 1972, is the Age of Development; the sixth, from 1973 to 1983, the Age of Professionalization; and finally the seventh from 1983 to 2000 the Age of Expansion and Integration. The philosophical foundations of evaluation and assessment include ontological, epistemological, and axiological studies. The general philosophy that forms the basis of the assessment includes idealism, realism, pragmatism, and existentialism. Meanwhile, the schools of Philosophy of Education that underlie the assessment are perennialism, essentialism, progressivism, and reconstructivism..

Keywords: *The Historical of Evaluation Program, Philosophy Assessment*

INTRODUCTION

Evaluation is not a new thing in human life. Actually evaluation, checking, and judging activities have been carried out since the beginning of human history. Practitioners, managers, and policy makers make judgments about students, clients, personnel, programs, policies and so on. This assessment leads to choices and decisions. Thus, it can be said that evaluation activities are a natural part of human life. Quality control of education is essentially an effort to control the quality of human resources in the system. To find out the extent of the effectiveness of the control, information is needed about how the condition of students is changing, whether teachers are carrying out their functions, whether the school supports the implementation of educational programs so that the results can be achieved optimally. Information about this, one of which can be obtained through evaluation, assessment, measurement, and valid testing (Majid, 2015: 1). Evaluation is a systematic assessment of the value or benefits of an object (Sanders and Chair, 1994: 3). Fitzpatrick et al. (2011) provided a comprehensive definition of evaluation where evaluation is the identification, clarification, and application of defensible criteria to determine the value of the object of evaluation (value or achievement) in relation to these criteria. Evaluation is a judgment on learning outcomes based on data obtained from the assessment.

S. Jerome Arcaro (2007) Assessment is a process to obtain data/information from the learning process which aims to monitor the development of the learning process and provide

feedback to teachers and students. There are various types of assessment and one of them is a test. A test is an assessment in which participants are asked to answer a series of written/oral questions or demonstrate an ordered skill. Generally carried out at the end of a learning stage and the answers given have the possibility of being right or wrong. Assessments, including tests, produce information on achievement relative to learning objectives. Assessment can be divided based on its use, namely summative assessment and formative assessment. Summative assessment is used to obtain information about student achievement after completing an activity or module or unit. Assessment can provide continuous feedback about students to improve learning. To understand more deeply about these two things, the author will explore them from the historical aspects of program evaluation and assessment philosophy.

RESEARCH METHODS

The research method is literature review or literature study, which contains theories relevant to research problems. The problem in this research is to find out philosophy, the historical of evaluation program and philosophy assessment. In this section, an assessment of the concepts and theories used is carried out based on the available literature, especially from articles published in various scientific journals. Literature review serves to build concepts or theories that form the basis of studies in research. Literature review or literature study is an activity that is required in research, especially academic research whose main purpose is to develop theoretical aspects as well as aspects of practical benefits

RESULTS AND DISCUSSION

The Age of Reform 1792-1900

William Farish invented the quantitative mark to score examinations. Replacing qualitative assessments of student performance with a mark for a “correct” answer permitted the ranking of examinees and the averaging and aggregating of scores. This was the first development in the field of psychometrics as we know it today). In fact, Farish revolutionized testing, a technology that plays an important role in the history of program evaluation to the present. The 19th century was the era of the Industrial Revolution with all of its attendant economic and technological changes. The very structure of society was transformed. Major social changes occurred. There was drastic change in physical and mental health and outlook, in social life and social conscience, and in the structures of social agencies. There was the laissez-faire philosophy of Bentham and the humanitarian philosophy of the philanthropists. There were continued but often drawn out attempts to reform educational and social programs and agencies in both Great Britain and the United States. The effectiveness of a school or instructional program. Then, at the urging of Samuel Gridley Howe, written essay examinations were introduced into the Boston grammar schools by Horace Mann and the Board of Education. Ostensibly the essay exam, modeled after those used in Europe at the time, was introduced to replace the viva voce or oral examinations. The latter mode of examination had become administratively awkward with increased numbers of pupils and was also seen as unfair because it could not be standardized for all pupils. The interesting point in terms of program evaluation was the hidden policy agenda behind the move to written examinations; namely, it was the gathering of data for inter-school comparisons that could be used in decisions concerning the annual appointment of headmasters. Howe and Mann attempted to establish differential school effects and used these data to eliminate headmasters who opposed them on

the abolition of corporal punishment. This is an interesting early example of politicization of evaluation data.

The Age of Efficiency and Testing 1900-1930

During the early part of the twentieth century the seminal work by Fredrick Taylor launched the scientific management movement, an early form of personnel evaluation. Taylorism continues to affect almost all aspects of American life to this day. By 1915, thirty to forty large school systems had completed or were working on comprehensive surveys on all phases of educational life. A number of these surveys employed the newly developed “objective” tests in arithmetic, spelling, handwriting, and English composition to determine the quality of teaching. These tests were often developed in large districts by a bureau or department set up specifically to improve the efficiency of the district.

It is important to point out that studies of efficiency and testing were for the most part initiated by, and confined to, local school districts. In contrast to the national curriculum development projects of the late 1950s and early 1960s, curriculum development before the 1930s was largely in the hands of a teacher or committee of teachers. It was natural, therefore, that evaluations of that period were addressed to localized questions. This focus or emphasis on local evaluation questions continued into the 1960s despite the fact that the audience for the evaluations was state-wide or nation-wide; this resulted in many useless educational evaluations being carried out during the 1960s. And, it wasn't until the 90s with the advent of standards based reform that the focus shifted from local to state level control over many aspects of the curriculum.

The Tylerian Age 1930–1945

Ralph W. Tyler has had enormous influence on education in general and educational evaluation and testing in particular. He is often referred to, quite properly we feel, as the father of educational evaluation. Tyler began by conceptualizing a broad and innovative view of both curriculum and evaluation. Curriculum as a set of broadly planned school experiences designed and implemented to help students achieve specified behavioral outcomes. Tyler coined the term “educational evaluation” which meant assessing the extent that valued objectives had been achieved as part of an instructional program. During the early and mid-1930s, he applied his conceptualization of evaluation to helping instructors at Ohio State University improve their courses and the tests that they used in their courses. During the depths of the Great Depression, schools, as well as other public institutions, had stagnated from a lack of resources and, perhaps just as importantly, from a lack of optimism. Just as Roosevelt tried through his New Deal programs to lead the economy out of the abyss, so too John Dewey and others tried to renew education. The renewal in education came to be known as the Progressive Education Movement, and it reflected the philosophy of pragmatism and employed tools from behaviorist psychology.

The Age Of Innocence 1946–1957

During this period there was considerable development of some of the technical aspects of evaluation; this was consistent with the then-prevalent expansion of all sorts of technologies. Chief among these developments was the growth in standardized testing. Many new nationally standardized tests were published during this period. Schools purchased these tests by the thousands and also subscribed heavily to machine scoring and analysis services that the new technology made available. The testing movement received another boost in 1947 with the establishment of the Educational Testing Service.

By the 1950s, the standardized testing business had expanded tremendously, and the professional organizations concerned with testing initiated a series of steps designed to regulate

the test-related activities of their members. In 1954, a committee of the American Psychological Association prepared *Technical Recommendations for Psychological Tests and Diagnostic Techniques* (APA, 1954). In 1955, committees of the American Educational Research Association and the National Council on Measurements Used in Education prepared *Technical Recommendations for Achievement Tests* (AERA and NCMUE, 1955). These two reports provided the basis for the 1966 edition of the joint AERA/APA/NCME *Standards for Educational and Psychological Tests and Manuals* (APA, 1966) and the 1974 revision entitled, *Standards for Educational and Psychological Tests* (APA, 1974). The latter report recognized the need for separate standards dealing with program evaluation. A revision of the Standards in 1985 contained a chapter on the use of tests in program evaluation, as did a further revision in 2000. During the 1950s and early 1960s there was also considerable technical development related to the evaluation. Since implementing the Tyler approach in an evaluation required that objectives be stated explicitly, there was a need to help educators and other professionals to do a better job articulating their objectives.

The Age of Development 1958–1972

The age of innocence in evaluation came to an abrupt end with the call in the late 1950s and early 1960s for evaluations of large-scale curriculum development projects funded by federal monies. This marked the end of an era in evaluation and the beginning of profound changes that would see evaluation expand as an industry and into a profession, focused on helping meet society's needs and dependent on taxpayer monies for support.

As a result of the Russian launch of Sputnik in 1957, the federal government enacted the National Defense Education Act of 1958. Among other things, this act provided for new educational programs in mathematics, science, and foreign language; and expanded counseling and guidance services and testing programs in school districts. A number of new national curriculum development projects, especially in the areas of science and mathematics, were established. Eventually funds were made available to evaluate these curriculum development efforts. The best and the brightest of the educational evaluation community were involved in efforts to evaluate these new curricula; they were adequately financed, and they carefully applied the technology that had been developed during the past decade or more. Nonetheless, by the early 1960s it became apparent to some leaders in educational evaluation that their work and their results were neither particularly helpful to curriculum developers nor responsive to the questions being raised by those who wanted to know about the programs "effectiveness." The late 1960s and early 1970s were vibrant with descriptions, discussions, and debates concerning how evaluation should be conceived; however, this period in the history of program evaluation ended on a down note.

The Age of Professionalization 1973–1983

Beginning about 1973 the field of evaluation began to crystallize and emerge as a profession related to, but quite distinct from, its forebears of research and testing. While the field of evaluation has advanced considerably as a profession, it is instructive to consider this development in the context of the field in the previous period. Before this period evaluators faced an identity crisis. They were not sure whether they should try to be researchers, testers, administrators, teachers, philosophers, or iconoclasts. It was unclear what special qualifications, if any, they should possess. There was no professional organization dedicated to evaluation as a field, nor were there specialized journals through which evaluators could exchange information about their work. There was essentially no literature about program evaluation except unpublished papers that circulated through an underground network of practitioners. There was a paucity of pre-service and in-service training opportunities in

evaluation. Articulated standards of good practice were confined to educational and psychological tests. The field of evaluation was amorphous and fragmented—many evaluations were carried out by untrained personnel, others by research methodologists who tried unsuccessfully to fit their methods to program evaluation. Evaluation studies were fraught with confusion, anxiety, and animosity. Evaluation as a field had little stature and no political clout. During this period, evaluators increasingly realized that the techniques of evaluation must achieve results previously seen as peripheral to serious research; serve the information needs of the clients of evaluation; address the central value issues; deal with situational realities; meet the requirements of probity; and satisfy needs for veracity. While the field had yet to develop a fully functional methodology that meets all these requirements, there were during this period some promising developments, including goal-free evaluation.

The Age of Expansion and Integration 1983-2001

The final period in our abbreviated history of program evaluation begins with the publication of the first edition of this book and ends with the publication of the second edition. A lot happened in the field during these eighteen years. Despite the positive title we chose for this last period it began on somewhat of a down note as there were considerable cut backs in the funding of evaluations during the Reagan years. Many evaluations during this period were geared toward cost cutting and cost benefit issues. However, these accountability systems need to be independently evaluated as many set unrealistic improvement goals based on student test performance. The standards based reform movement has now reached 49 states. Curriculum frameworks are developed, as are tests to measure progress on reaching the standards contained in the frameworks. A number of states have linked student performance on these state level tests to graduation and retention decisions. Kentucky and Vermont have had excellent evaluations of their state testing programs, which experimented with the substitution of performance measures for the traditional multiple-choice tests. As the economy grew however, evaluation as a field expanded and became considerably more integrated. The expansion is seen especially in the development of professional evaluation societies in more than twenty countries and in the coming together, communication, and collaboration of evaluators from various disciplines. In education the reform movement has had a profound effect on program evaluation. Proposed reforms—such as charter schools, vouchers, and privatization of schools—all predicated on the belief that introducing competition into the system will lead to improvement, are currently being evaluated.

Philosophy Assessment

A univocal philosophy of evaluation and assessment as well as a universally accepted philosophy of Science has yet to be found. Basically all knowledge, whether it is science, art, or any kind of knowledge (including evaluation) has three foundations, namely ontology, epistemology, and axiology (Suriasumantri, 1994: 35). *Ontology* comes from the Greek words *on*, *ontos* (being or being) and *logos* (study or knowledge of). So ontology can be interpreted as the study of the essential characteristics of what is in itself which is different from the study of things that exist in particular. The term epistemology also comes from the Greek words *episteme* (knowledge or knowledge) and *logos* (knowledge or information). Epistemology is said to be knowledge of knowledge or rather the theory of knowledge. While axiology comes from the Greek *axios* (worthy, appropriate) *logos* (study of science), so axiology can be interpreted as value analysis, namely limiting meaning.

According to Orstein and Hunkins (2013: 31) there are four general philosophical schools that are very influential in the world of education (including assessment) namely Idealism, realism, pragmatism, and existentialism. The following are:

1. Idealism: the main character of this school is Plato with his Republic. This school holds that something true is what exists in the realm of ideas. The highest goal of this school is to achieve eternal truth and values. Because ideas can be integrated into universal concepts and meaningful wholes. Idealistic educators always connect ideas with concepts. Knowing something is rethinking a hidden idea that once came to mind. The teacher's task is to awaken the hidden into knowledge that is truly realized, because learning involves, remembering, and working with ideas (Orstein and Hunkins, 2013). Idealism wants education directed at the inheritance of intellectual and cultural property that is rich in truth throughout the ages. This means that students must move closer to the ideal by imitating the teacher through discussion of ideas in subjects that relate ideas and concepts to each other. According to the idealism of synthesis between concept systems in the subjects of language, mathematics, and aesthetics, it represents the diversity of absolute dimensions (Ansary, 2015: 77).

2. Realism. The main character of realism is Aristotle. Other characters are Thomas Aquinas, Harry Broudy, and John Wild (Orstein and Hunkins, 2013). This flow is of the view that something that is true is a reel, which is tangible. The world for realists is objects and matter that can be perceived by the senses and reason. Everything for the rationalist comes from nature and is subject to the laws of nature. Human behavior is said to be rational if it is in accordance with natural laws and governed by physical and social laws. Realism emphasizes subjects that prioritize logic and thinking. Realists value science as much as art. Realism emphasizes causal relationships in the real world which implies that realism is more realistic than idealism, which is more focused on things as they are, not what they should be. The implication of realism in the world of education is to teach students to understand and adapt to nature. They must learn how to live in harmony with nature where nature shows very diverse symptoms (Ansary, 2015: 79-80).

3. Pragmatism: It assumes that knowledge is a process where reality is always changing, therefore learning occurs when someone is involved in problem solving. According to John Dewey, education is a process of improving, not accepting the condition of humanity. Therefore, the main emphasis is on problem solving using the scientific method (science), not collecting facts or views. So the subject is interdisciplinary (Orstein and Hunkins, 2013). This school is also known as experimentalism because it emphasizes experimentation, the process of change, and relativity. Reality for pragmatists is always changing because they reject the idea of universal and fixed truth. This flow holds that something true is something that has a use value or benefit. The emphasis of pragmatism is on problem solving and the scientific method. The only guide that people have when they interact with their social world or environment is a well-established generalization, a statement that is subject to further research and verification. For pragmatism, teaching must be centered on critical thinking, teaching is more exploratory than explanatory. The method is more important than the subject matter. There are three important characteristics of the knowledge version of the good according to pragmatism. First, pragmatic knowledge is only assumed to be true tentatively. Second, pragmatic knowledge generally has social references. Third, pragmatic knowledge is not accepted or discovered, but reconstructed (Zais, 1976).

4. Existentialism: The figures of this school are Soeran Kierkgaard (theistic) and Neitzche (atheistic). If Descartes bases his philosophy on the premise that "I think, therefore I exist" on the other hand existentialists start with the premise "I exist, therefore I think". In short, for existentialists existence precedes essence. According to existentialism, humans cannot free themselves from the responsibility to choose and define themselves. The existentialist ontology places the ultimate reality "inside" the individual self of every human being (Zais, 1976). Humans according to existentialism are unfinished creatures, so we must continue to update ourselves. Humans determine their identity, because humans determine their essence.

Existentialism suggests that students are given the freedom to choose what and how they want to learn. Existentialist critics believe that the most important knowledge is knowledge of the human condition. Education must develop an awareness of choice and its significance. Existentialists reject the imposition of group norms, authority, and established order. They recognize some standard, custom, or opinion as undeniable.

From several schools of general philosophy as described above, it becomes the basis for the flow of educational philosophy that influences educational decisions, curriculum, assessment, and learning.

The schools of education philosophy are as follows:

1. Perennialism. Perennial means eternal. The philosophical basis of perennialism is that humans are essentially the same and constant, even though they are in different times, times and places. Therefore, by studying past values about human nature, righteousness, and virtue, we can take lessons. Humans have the ability to understand and understand the universal truths of nature. The purpose of education is to develop human rationality and reveal universal truths by means of intellectual training (Orstein and Hunkins, 2013).
2. Essentialism. The originator of essentialism was William Bagley. Essentialism is supported by idealism and realism, but the two do not merge into one. Essentialism emphasizes on preserving the best knowledge, values, character, and customs of society from the past and present. Education aims at nurturing students, our future citizens, reaffirming their commitment to society and contributing to their cultural renewal. Essentialism pays more attention to contemporary issues. According to the essentialists, the school curriculum should be directed to its essential nature only; science, history, literature, mathematics and the arts. As for high school English, mathematics, science, history and foreign languages (Orstein and Hunkins, 2013).
3. Progressivism. Progressive means forward or towards progress. Progressivism emerged as a reaction to the traditionalist formalist perennialist education model (Nurhayati, 2002). Progressivism rejects the following: authoritarian teaching, over-reliance on textbook methods, memorizing factual data with constant practice, static goals and material that fail to account for a changing world, intimidation or corporal punishment as a form of discipline, and attempts to segregate education from individual experience and social reality. However, according to Lawrence Cremin, the movement's inability to reach consensus on the school's goals, or even establish a set of pedagogical principles, led to its downfall. For progressivism, skills and tools for learning include problem solving methods and scientific inquiry. The learning experience must include cooperative behavior and self-discipline. Both are considered essential for a democratic life.
4. Reconstruction. It means to remodel the old order and replace it with a new one. This flow arises as a reaction to the loss of liberty and equality. With the loss of freedom or independence where humans are segmented and not free to actualize themselves, the loss of a sense of brotherhood in which the proletariat continues to exploit the workers, and social stratification occurs. schools have a very large role in this social reconstruction because schools must take on the role of agents of change (Nurhayati, 2002: 70-71).

CONCLUSION

Based on the description above, several conclusions are:

1. Evaluation, assessment, measurement, and tests have a role in the assessment of a series of educational activities. Each plays a role and there is a wedge between one another.
2. Evaluators need to be aware of both contemporary and historical aspects of their emerging profession, including its philosophical underpinnings and conceptual orientations. Without this background, evaluators are doomed to repeat past mistakes and, equally debilitating, will fail to sustain and build on past successes. The history of evaluation activities has been going on since 1900, many evaluation practices have been found throughout history, but program evaluation is still informal. Experts differ on when to start a formal evaluation. However, there are similarities between one character and another.
3. The philosophical foundations of evaluation and assessment include ontological, epistemological, and axiological studies. The general philosophy that forms the basis of the assessment includes idealism, realism, pragmatism, and existentialism. Meanwhile, the schools of Philosophy of Education that underlie the assessment are perennialism, essentialism, progressivism, and reconstructionism.

REFERENCES

- Ansyar, Mohamad. 2015. *Kurikulum: Hakikat, Fondasi, Desain & Pengembangan*. Jakarta: Kencana Prenadamedia Group.
- Arikunto, S. and Jabar, C.S.A. 2009. *Evaluasi Program Pendidikan*. Jakarta: Bumi Aksara.
- Bagus, Lorens. 1996. *Kamus Filsafat*. Jakarta: Gramedia.
- Basuki, I. and Hariyanto. 2017. *Asesmen Pembelajaran*. Bandung: Remaja Rosdakarya.
- Fitzpatrick, J.L., Sanders, J.R., dan Worthen, B.R. (2011). *Program Evaluation, Alternative Approaches and Practical Guidelines*. Boston: Pearson.
- Madaus, G.F, and Stufflebeam, D.L. 2000. *Evaluation Models*, Boston: Kluwer Academic Publishers.
- Majid, A. 2015. *Penilaian Autentik: Prosedur dan Hasil Belajar*. Bandung: Remaja Rosdakarya.
- Masrus, R. dkk. 2016. *Educational Assesment and Evaluation*. Islamabad: Allama Iqbal Open University.
- Nurhayati 2002. *Filsafat Pendidikan Islam*. Pekanbaru: Benteng Media.
- Ornstein, Allan C. and Hunkins, Francis P. 2013. *Curriculum: Foundations, Principles, and Issues*. England: Pearson.
- Poerwadarminta, W.J.S. 2007. *Kamus Umum Bahasa Indonesia*. Jakarta: Balai Pustaka.

- Sanders, James R. and Chair. 1994. *The Program Evaluation Standards*. London: Sage a Publications.
- Stufflebeam, C.F. Madaus and T. Kellaghan (eds.). *Evaluation Models*. Copyright © 2000. Kluwer Academic Publishers. Boston.
- Suryasumantri, J. 1994. *Filsafat Ilmu Sebuah Pengantar Populer*. Jakarta: Pustaka Sinar Harapan.
- Travers, R.M.W. 1983. *How Research has Changed American Schools*. Kalamazoo: MI: Mythos Press.
- Worthen, Blaine R. and Sanders, James R. 1987. *Educational Evaluation, Alternative Approaches and Practical Guidelines*. London: Longman Group.
- Zais. Robert S. 1976. *Curriculum: Principles and Foundations*. New York: Harper & Row Publishers..