Sanja Španja, Ph.D., Postdoctoral Fellow
J. J. Storossmayer University in Osijek
Faculty of Humanities and Social Sciences
Department of Pedagogy
sspanja@ffos.hr

EVALUATION AS A PEDAGOGICAL PHENOMENON

Summary: Evaluation is the result of a continuous, recursive and divergent development process for all stakeholders in the education system which involves a number of interaction effects. Throughout the history, the evaluation reflected what the teachers believed to be the education of a particular moment in given history.

In professional and scientific literature, numerous definitions of evaluation can be found, and it is difficult to provide a unique definition of the concept which could be applied to different social and scientific surroundings. However, in this paper, the evaluation will be defined as an assessment of certain education activities and results, the effectiveness of interventions as well as the answer to whether the desired results have been reached and to what extent. Furthermore, the paper will discuss the evaluation of the education system and the stakeholders in the evaluation as well as provide an overview of the evaluation models which determine the differences in the evaluation processes.

Since the education is simultaneously both the process and the result of the process, the question arises as to whether the evaluation is to be predetermined or to be open to unexpected results and outcomes as well.

Key words*: stakeholders, evaluation, teaching process, education, pedagogical phenomenon*

DEFINITIONS OR CHARACTERISTICS OF EVALUATION

Evaluation is investing in people, but it is also the progress. (Guba and Lincoln, 1989)

The very notion of evaluation has its roots in the French word évaluation, which means determining the value of something, to grade, to assess, and the French word *evaluer –* the coin is derived from two Latin words *ex –* outside, out, and *valuer* – value (Hrvatski jezični portal/Croatian Language Portal, http: /hjp.novi-liber.hr/ index.php?show=search). Evaluation and valorization are sometimes used as synonyms. "Evaluation is an extremely important activity in education, and as a phenomenon and pedagogical concept it has always attracted the attention of experts from various fields of pedagogy, psychology, and dokimology (assessment of scholastic achievement), but research and studies differed in philosophy and practices of the upbringing used as postulates and in dokimology solutions they recommended." (Matijević, 2005: 280)

Furthermore, the focus on predetermined evaluation objectives assumes they are formally identified in the planning phase of the intervention and that clear and specific indicators of success have been developed. An additional problem in defining

the evaluation is the interpretation of the results which are to help decide on future modes of action – whether about the development of specific programs or the development of databases of evidence in general. Contrary to this, supervision, often erroneously referred to as evaluation, is defined as *systematic and continuous monitoring* or keeping track of the activities to ensure everything is being carried out according to the plan (Feuerstein, Falik, and Feuerstein, 1998: 184). The emphasis is therefore placed on monitoring and recording the situation during the program implementation process, while the evaluation is concerned with assessing what has been achieved and the manner in which the changes have taken place.

When it comes to differences, the answer can be found in pedagogical pluralism and tolerance. The key to understanding the complexity of the evaluation processes can be found in the most prominent features of the evaluation. Most definitions refer to several key features:

- response to specific, predetermined questions;
- collecting information;
- passing judgments;
- making decisions;
- commenting on policy in a given situation.

Evaluation can be defined as "gathering information on certain issues which help make judgments from which decisions on specific action are derived" (Morrison, 1993: 2). Numerous scientists have developed different theories which try to explain the essence of the determinism of the events in the evaluation process. For the purpose of this paper, we will rely on several evaluation theories and models with an emphasis on the fourth generation evaluation model (Guba and Lincoln, 1989).

Evaluation is an essential component of the theory of curriculum and an indispensable part of each national curriculum. Each school should constantly assess all school's activities and students' activities and achievements. That is why the schools undergo self-assessment. It is important that schools use quality assessment models which will enable teachers, as well as other stakeholders in the evaluation, the tools to develop the practice and improve continuous students' learning (Norris et al., 2017: 758). This process includes, besides teachers, students, parents and local school management as well as other members of the local community (Matijević, 2007).

The evaluation consists of procedures aimed at determining the effects of conducted research interventions or activities planned to achieve an individual or social goal (Posavec and Carey, 1992). It is an objective, critical review starting from the level at which we can verify if predetermined goals have been met, whether by the programs in whole or simply by some of their constituent parts (St. Leger et al., 1992: 1). The purpose of the evaluation is to measure the effects of the programs against the objectives set, and it is organized as an instrument which contributes to the improvement of the present programs and the improvement of future programs and decisions (Weiss, 1972: 4; Kaneko, 1999: 433). Therefore, the purpose of the evaluation is to find answers to questions about value and effectiveness, outcomes and effects, implementation of evaluation according to the plan and functional aspect of the theory and design in the

background of the program as well as the needs of the program (Rossi, Lipsey and Freeman, 2004, cited in Thistlethwaite, 2015: 292).

Weiss (1998) defines the evaluation as a systematic application of research procedures used to verify the concept, design, implementation, and utility of social interventions. The author also describes the evaluation as a process of organizational learning necessary for the progress and survival of each organization in a dynamic external environment of economic, political and cultural changes. Evaluation is a dynamic interdisciplinary and interactive process of asking questions, discussing and assembling the opinions of all parties involved in some of the activities for a greater understanding of one's own work and thus the future improvement. Evaluation is, therefore, the assessment and review of a particular activity (Weiss, 1998: 12). The purpose of the evaluation is to contribute to solving practical problems by clearly defining the rules and responsibilities of the person who is doing the work, of what is being done and why it is being done in a particular manner. The collected data is used as the source of information for future activities – performance (Springett, 2001: 144).

Evaluation of the education process involves the collection and use of information for the purpose of making decisions in education and ultimately for the future development and progress. It is an integral part of the professional practice of each teacher. Teachers rely on the evaluation for personal decision making and/or improvement of practices (Norris et al., 2017: 758). Teachers evaluate the teaching process in order to assess their achievements in their own professional practice, to determine their strengths and weaknesses, and to compare themselves with other fellow teachers (Hounsell, 2003). In short, we could define the evaluation as the establishing of the extent to which certain values and goals are achieved (Tones, 1997: 52; Green and South, 2006). Today, it has become a priority to evaluate or demonstrate evidence of students' achievements as well as teacher's effectiveness (Norris et al., 2017: 758). However, there are numerous discussions on how to measure the success and efficiency of the process changes and the evaluation itself. Different stakeholders in the evaluation process will discuss different background postulates as well as different types of evidence to confirm their success. An increasing emphasis on the evidence of success can be associated with *economic rationalism* as well as the need to convince the public authorities that the funds focused on the development of education are utilized to the best possible extent (Rapfhael, 2000).

Chen (1990) suggests that the evaluation should reflect the needs of different stakeholders, be objective, reliable and try to generalize the results obtained. Furthermore, the question is whether the emphasis should be on assessing what has been achieved or understood as being achieved, i. e. the assessment of the results or the process underwent in reaching those results. In addition, evaluating the quality of performance denotes that the one who evaluates has a clear idea of what needs to be measured and how it is measured (Kennedy, 2008, cited in: Norris et al., 2017: 759). Numerous evaluation systems we know of today developed in the mid-70s of the last century and were strongly focused on the documentation of small numbers of behavioral observations (Danielson and McGreal, 2000: 3).

It can be concluded that the evaluation (projects and/or programs evaluation) periodically evaluates the efficiency, effectiveness, performance, sustainability and program/project relevance in the context of the objectives set. Furthermore, evaluation is inherently more comprehensive and more detailed than the monitoring implementation, focusing on a broader range of program management issues and its effects. It is conducted as an independent analysis of the environment, goals, results, activities, and resources invested to reach conclusions which could later be used for future decisions. It must help develop, clarify, improve, monitor/oversee and justify what is being assessed. The evaluation must also be guided by changes in practice or increased understanding of practice and must be separated from the tasks of stakeholders who make only one part of the evaluation process (Thistlethwaite et al., 2015: 292). Thus, we are dealing with the quality of what has been done.

With regard to the evaluation cycle, we distinguish the following evaluations: ex-ante evaluation before program/project implementation, on-going evaluation or interim evaluation or mid-term evaluation during program/project, ex-post evaluation after the program/project implementation. In the program, in addition to performance evaluation, there are other evaluations as well (e.g. thematic evaluations, strategic evaluations, policy evaluation of common interest, theory-based evaluation, evaluation against factual impact, meta-evaluation/evaluation of evaluation) (Tyler, 1989).

Since the evaluation in the education system is an important issue, the next chapter will focus on internal and external indicators of education system evaluation which can effectively affect changes in the education system.

INTERNAL AND EXTERNAL INDICATORS OF THE EDUCATION SYSTEM EVALUATION

The fundamental task of the education system is to meet the social needs of individuals and entire communities. It is difficult to pass judgment on how much was the evaluation of the education system able to meet different social needs through its historical development. Expectations of all stakeholders involved in the evaluation of education processes often differed significantly from the practice. In the education system, people such as educational advisors, principals, and associates (usually pedagogues), most often evaluate education processes. In addition to attending classes for monitoring (regular and exceptional visits) and optional activities, they analyze lesson plans, pedagogical documentation (curriculum, programs, directories, daily logs, registry books...), students' success, and conduct structured and unstructured interviews.

It is important to emphasize that the problem of evaluation of education systems posits an entirely novel issue and challenge to numerous experts and scientists despite the millennial continuity of education.

The first evaluations of education systems in Europe were carried out after the Second World War and they have been conducted with various frequencies since then, but it is still not possible to name a single clear and unique model with adequate, reliable and measurable indicators which could be used to evaluate the education system as a whole (Freideburg, 1978: 5). Some indicators, Freideburg (1978) states,

have been more than well-developed as a possible approach to the evaluation, although their quality has in numerous cases been challenged by the scientific and professional public. However, everyone agrees that the issue of evaluation of the education system is an important problem and that it is only through it that the changes in the education system can be efficiently implemented. Such an evaluation should be conducted according to the following indicators:

1. internal evaluation indicators of the education system;
2. external evaluation indicators of the education system.

Internal evaluation indicators of a part or entire education system are, as a rule, straightforward, and by measuring external indicators it is necessary to combine intermediate and immediate indicators and sometimes, long-term estimates, which can be quite unreliable. As a result, we can conclude that for now there is no reliable guide to the evaluation of the issue of evaluation and the existing experience should be synthesized and, above all, taken as an incentive for further exploration of this controversial problem (Naisbitt 1984: 68).

There are certain objections stating that different individuals within the system have goals, not the system itself. According to such an extreme approach, individual goals can easily become average, local and targeted at one or more stakeholders in the education process. The other stance claims that the goals of the education system cannot be determined because such a comparison between individuals is possible only if we are willing to pass judgments of essentially ethical nature. Such a comparison can be summarized as a function of social protection in which the social status of a person has been in some way added to the social status of another person (Mijatović, 1994: 196).

An alternative approach could determine the minimum requirements for social existence. We are aware of the individual needs (the lack of those needs could cause physical and psychological damage) or social needs, the lack of which would cause society to regress. Of course, it is not easy to define such needs. A further possibility is to maintain a strictly sociological attitude – goals could only be the property of the organization (Runciman, 1975: 57).

In addition to the aforementioned internal and external evaluation indicators of the education process, numerous authors emphasize the need for better categorization. All previous papers on evaluation indicate issues of precisely defining the indicators (internal and external), as well as the fact that most of the issues which can be identified as indicators are interconnected and mutually conditioned. The fundamental internal indicators (those that can be tracked in the education system) are:

a) the adoption of knowledge (skills, habits, abilities);
b) educational results and contradictions.

Among external indicators, those which are used to value education system performance in a wider community, basically explore the link between the education system and the society. They are as follows:

a) education and labor market situation;
b) equality of educational opportunities;

c) educational opportunities for individual needs;
d) education and quality of life (Mijatović, 1994).

Each of these indicators is divided into a large number of groups and subgroups. A clear classification has a practical advantage because evaluation research is often directed to respond to political issues in given areas. There is also an area of values, tradition and social justice and social well-being. In doing so, it is important to bear in mind the fact that the order of global indicators actually reflects how they are perceived and determined by policy makers.

Evaluation experts should be informed about the basic measures of education opportunities available to the population. The offer of the desired target per average inhabitant or the average number of years of education could be the initial parameters in measuring the social well-being of a society. If they were to be applied to a school in the function of the social service, they could evaluate the level of use provided to each individual in society, which is the qualification structure of employees with the number of years of education and knowledge configurations. It would be desirable to evaluate the quality of that knowledge and be aware at all times of the influence of current political programs which have their own interests in education (Mijatović, 1994).

Although the aims of the education system are sometimes contradictory, once chosen, further activity is objective because of the manner in which the problem-solving depends on the paradigms of the education process. Moreover, the very notion of a rational argument depends in part on paradigms explaining the system functions. We must carefully distinguish education policy aimed at achieving certain goals and achieving those goals in the long run. On the other hand, education policy, once set to meet certain goals, can become self-serving. It is, therefore, necessary to consider equality of approach both as a final goal and as a final stage to one such goal as the equality of results (Weber, 1949: 43). Speaking about the evaluation of the education system, Madelin (1984) emphasizes that we do not have to go back to Cicero, who was very critical of the students of his time, for there has been poor students and unsatisfactory schools and there always will be. Libraries are filled with books from different historical periods of the education system warning the public of the catastrophic position of education and schooling system (Madelin, 1984: 13).

Vizek-Vidović et al. (2003) note that the evaluation of the education system should be useful to teachers. In this respect, all active participants in the education process would be able to compare the evaluation data of what they are really doing with what they think they are doing; to plan lessons, to develop curriculum, etc. The authors, besides the formative and summative evaluation, which will be discussed later in this paper, introduce "diagnostic evaluation", which records the current state of teaching, especially before the introduction of any changes – so that the evaluation experts would be able to attribute certain changes to their possible effects. The authors believe that this form of evaluation in our schools is usually seen as a controlled observation and has a negative connotation in terms of inspection and supervision (Vizek-Vidović et al., 2003: 444).

It can be concluded that without the internal and external evaluation indicators of the education system, no matter their categorization, it is certainly not possible to determine the objective of the evaluation, let alone conduct the evaluation process itself. In addition, for a better understanding of the evaluation and its implementation in practice, it is necessary to differentiate evaluation models and types of evaluations. Each education institution should therefore recognize and respond to their own needs.

EVALUATION MODELS AND TYPES OF EVALUATION

Scott (1998), Herman et al., (1987) identified and categorized seven evaluation models – goal-oriented evaluation, decision-oriented evaluation, evaluation research, participant-oriented evaluation, independent assessment evaluation, evaluation oriented towards alternative explanations and stakeholder-oriented evaluation (Table 1). The classification into these evaluation models, in theory, helps to better understand the evaluation process itself. However, in practice, we often realize they are all intertwined. It is important to emphasize that this classification can be useful to the scientific and professional public in raising awareness of the complexity of the evaluation, taking into account the different needs many stakeholders may have in the evaluation process.

Table 1. Seven evaluation models in relation to their focus in the evaluation process (cf. Scott, 1998; Herman et al., 1987)

EVALUATION MODEL	FOCUS
GOAL-ORIENTED EVALUATION	The assessment of efficiency, effectiveness and the cost-effectiveness of interventions in the implementation of the goals of the education process.
DECISION-ORIENTED EVALUATION	Enhancement of the decision-making.
EVALUATION RESEARCH	Explaining the correlation of the evaluation with learning outcomes.
PARTICIPANT-ORIENTED EVALUATION	Evaluation process from participants' perspective.
INDEPENDENT ASSESSMENT EVALUATION	Openness to the achievements of others unburdened by the predetermined goal and aims.
EVALUATION ORIENTED TOWARDS ALTERNATIVE EXPLANATIONS	Possible alternative acceptance and understanding of what is happening in the evaluation process.
STAKEHOLDER-ORIENTED EVALUATION	Usefulness of evaluation for different stakeholders.

According to Table 1, regardless of the model of evaluation, there are differences in evaluation processes and their orientation as well as the level of responsibility and "ownership and stakeholders" in decision-making, research, alternative explanations, etc., relative to results in the implementation of a predetermined objective and

aims. Seven evaluation models help in understanding evaluation processes, changes assessment, and processes involving openness to unexpected results and effects. Models also differ in relation to the level they are aimed at in meeting the needs of particular stakeholder groups. Osuna, Cruz-Casto, and Sanz-Menendez (2011: 588, cited in: De Rijcke et al., 2016: 163) state that the introduction of a system or a model of evaluation could influence the standardization of patterns and norms researchers should implement. Thus, comprehending the effects of the introduction of evaluation systems, i.e. models, would be used in order to improve both the more advanced quantitative analysis and the profoundly qualitative research on how these models are implemented in institutions and practice.

In addition to the seven models of evaluation in education practice, we differentiate formative and summative evaluation. The concept of the summative and formative evaluation was first introduced by Michael Scriven in 1966., in his book *The Methodology of Evaluation*.

The summative evaluation is carried out at the end of the intervention to assess the achievements. Although the outcomes and products are emphasized, process elements could also be included. Tones and Tilford (2001: 114) highlight the importance of summative evaluation for two reasons – firstly: to determine whether all the necessary components of the initiative are included and to ensure quality assurance to help explain the level of success or failure and secondly: to clarify the processes involved in the summative evaluation. This kind of clarification and assessment explanation has the ability and potential to offer an explanation for achievements and identify key aspects in designing and implementing interventions of great importance for wider implementation. In the process of summative evaluation, we can ask the following questions: "What have we achieved?" By providing information for questions "How do we achieve this?", "What did we learn?" and "Will we do it the same way and next time?" (Green and South, 2006: 15). Summative evaluation, according to Patton (2002), examines the overall program, policy or product effectiveness to make a decision on the continuation and assess the possibility of enhancing it in other situations. Summative evaluation rarely relies heavily on data collected via qualitative research. However, the qualitative approach in summative data evaluation gives a certain depth and detail, while in quantitative data refinement is achieved.

Formative evaluation implies the evaluation of the education process. It may include all aspects of the education process, but most often it includes teachers and students and refers to the evaluation of the education process. Analyzing and evaluating the teaching process enables improvement of its quality, which makes it a prerequisite for the success of education (Bognar and Matijević, 2002). Formative evaluation helps improve programs. Its goal is to shape the focus of its analysis. Formative evaluation can be conducted by both students and teachers, and the purpose is to analyze and evaluate the education process. It relies primarily on a qualitative approach to research: evaluation implementation, case studies, and evaluation programs.

While summative evaluation provides retrospective analysis, formative evaluation is carried out during the process as well as the initiative being developed and im-

plemented. It, therefore, has greater immediacy and relevance for those responsible for action and intervention, as it can provide key information and feedback to guide future trends. So, the summative evaluation is used for reliability and confirmation, while formative evaluation is used to improve and develop educational programs or interventions (Nevo, 1983, cited in: Thistlethwaite, 2015: 293). Formative assessment often deals with processes. Formative assessment, therefore, has a lot in common with action research, as Cohen and Manion (1996: 192) defined it – an "on-site" work with a concrete problem, and in the immediate evaluation situation. Patton (2002) and Scriven (1966) emphasize the importance of using both types of evaluation during research in education.

Formative and summative evaluations are observed through a unique integrative process, which includes four basic phases consisting of several mutually different tasks, depending on the type of evaluation. The phases are as follows:

Setting the boundaries or framework of the evaluation – since it is a formative evaluation, it includes parts of the program;

Selection of appropriate evaluation methods – includes a selection of different methodological options, with appropriate measurement instruments and evaluation methods which include sample selection strategies as well as data processing plan;

Collecting and analyzing information – this is the operational stage for all plans and drafts of the evaluation study;

Reporting findings – finding the mode of reporting most appropriate at the end of the education process; (Herman, Morris, Taylor, 1987, cited in: Halmi, 2008).

In discussing different types of evaluation, we need to distinguish between external and internal evaluation. In the internal evaluation, we start by answering the question *Who* is the client and the provider of the evaluation. Then the evaluation is carried out at school, and the organizers are teachers and students. In a contemporary approach to management, leadership, self-assessment and independent learning, this form of evaluation is extremely important. Often the subjects outside the school (the stakeholders in the evaluation process) show interest in education outcomes – for example, government and other governmental and non-governmental institutions, international institutions and organizations as well. That is the external evaluation, as the organizers of the evaluation are subjects outside the school, i.e. subjects not involved in the education and education process which takes place in the school (Matijević, 2005: 286, Matijević and Radovanović, 2011: 223). External evaluation can help the institution or school to discover the difficulties and/or the advantages which most likely wouldn't be perceived in other circumstances. It can also help increase transparency and provide a new space for reaching consensus on new approaches and ideas at school (Biersteker, 2016: 343). The advice provided by external evaluation should certainly be approached critically, accepting and achieving the quality and useful feedback which could encourage an innovative and reflexive approach to education and the teaching process as well.

EVALUATION OF THE TEACHING PROCESS

In psychological literature dealing with education processes, we find two dominant models of evaluation – developmental and differentiated. Glickman (1981) defines the developmental evaluation of the education system and the education process starting from the following assumptions: teachers are adults and evaluation needs to take into account the nature of their personal developmental process, i. e. the individual differences between them. The essence of developmental evaluation is the belief that the ultimate effect on outcome has, on the one hand, the underlying attitudes, values and behaviors of the evaluation expert, and on the other hand, the traits of the teacher himself. According to Glickman (1981) the characteristics of teachers are defined by two dimensions: on the one hand, there is the level of motivation and dedication to work, and on the other hand, the level of abstraction. The abstraction level refers to the level of cognitive development and the flexibility of abstract thinking with the emphasis on the direct work with students. Those teachers who have a lower abstraction level are more confused, find it harder to employ different approaches to solve the problem, teachers with a higher level of abstraction are better off in unexpected education settings, they are more successful in defining the problems that might arise, and can suggest more solutions but have difficulties in implementing the chosen approach to solving the problem, especially if the problem is complex. Finally, teachers with a high level of abstraction are able to analyze the problem from different angles. For example, they are able to understand the problem from the students' point of view, but can also perceive it from the standpoint of their fellow teachers. As a result of a high abstraction level, teachers are able to find numerous alternatives and solutions to even the most complex of problems and carefully plan their implementation. Concerning the dedication to work, research shows (Gould, 1972, cited in: Vizek Vidović, 2003: 453) that teachers in their professional career undergo several specific periods.

Low level of dedication to work is expressed through an indifferent relationship with students and colleagues, as well as the minimum energy and time invested to do the necessary work. In fact, energy is only spent on job retention. At the same time, in schools there are teachers who show a high degree of care for students and other teachers, who are motivated, always ready and interested in investing more time in school-related activities, and are always struggling to do more and perform better. Evaluation studies of these two dimensions of dedication to work and the abstraction level lead to four categories of teachers (Table 2).

Table 2. Categories of the teacher with regard to dedication to work and flexibility of thinking

		DEDICATION TO PROFESSION	
		high	**low**
FLEXIBILITY OF THINKING – ABSTRACTION LEVEL	**high**	professional	analytical type
	low	poor teacher	person who chose the wrong profession

By analyzing Table 2, we can conclude that the characteristic of the right experts (professionals) is the desire for systematic lifelong education and advancement. If

they are confronted with the problem, they feel competent to provide solutions which are both in the interest of students and in the interest of the school. Analytical observers are able to carefully consider all aspects of the situation, but they do not have the willpower to do anything. Poor teachers would be those who would gladly do everything, but they are unable to determine strategies and actions which would lead them to the desired goal. It is the desire for advancement and continuous professional development that encourages teachers to self-assessment, as well as the assessment of their work and students' achievements.

When it comes to attitudes and actions of evaluating experts, Glickman (1981) distinguishes three general orientations toward evaluation and a total of ten different behaviors and actions:

Non-directive evaluation is based on the assumption that teachers are competent to evaluate their own work and improve after analyzing it. It is expressed through the following behavior:

1. *Listening*: during the observation of the teacher's work, the evaluating expert does not talk; he non-verbally encourages the teacher to speak;
2. *Clarification*: an evaluation expert raises questions to the extent necessary to better understand the problem or activity of the teacher;
3. *Encouragement*: the teacher is encouraged to talk about the possible causes of the problem as well as about his own strengths and weaknesses;
4. *Presentation*: if the teacher raises the issue or the difficulties he encounters in his work, the evaluating expert presents his observations and opinions.

Collaborative orientation focuses on the initiative in discussing the perceived problem which can be taken either by a teacher or an evaluation expert. If there is a disagreement about the solution, it would be good to introduce a third party, a mediator who would mediate in finding a satisfactory solution.

5. *Discussion*: an evaluation expert encourages and leads the discussion on possible directions for solving teacher's problems;
6. *Troubleshooting*: the teacher is encouraged to solve the problem immediately;
7. *Demonstration*: the evaluation expert directs the teacher to new patterns in problem-solving, which can be used in similar circumstances, and at the same time addresses the current problem.

Directive evaluation – the evaluation expert is expected to have a high level of control over a relationship with a teacher. He does not necessarily have to be authoritarian, but a person with more knowledge, experience, and competence, and is open to establishing standards that he expects to be achieved.

8. *Routing*: the evaluation expert in detail instructs the teacher how to deal with a problem;
9. *Scoring*: it is explained to the teacher what he must do to make his actions consistent with the behavior of other colleagues at school;
10. *Strengthening*: the evaluation expert specifies in detail the conditions and consequences of improving teacher's work.

Differentiated evaluation - According to Glatthorn (1984), differentiated evalu-

ation starts from the fact that different circumstances require different approaches. The undeniable fact is that teachers object any assessment and evaluation of any segment of their work and they perceive it as a control, an inflexible monitoring, rigid, incompatible with the complex processes of education practice. As a result of this, Glatthorn (1984) differentiates the evaluation of the education system and the teachers' work into four forms of evaluation: clinical evaluation, collaborative professional development, self-directing development and administrative review.

Clinical evaluation as part of the differentiated evaluation is developed to improve the in-class work of the teacher. The main objectives of the clinical evaluation, according to Acheson and Gall (1987), are to provide the teacher with objective feedback on the current state of his teaching skills; to identify and solve problems which arise during the education process, to identify and solve problems in the education process, to develop with the teacher the key competencies of teaching skills and strategies, to recognize, praise and highlight the work of teachers in order to promote him and help him to adopt a positive attitude to lifelong learning. Consequently, the clinical evaluation covers several stages. In the first stage, by working together and agreeing, the evaluation expert, as well as the teacher, organizes and plans to attend teacher's classes. After the visit, they analyze it and according to the results of the analysis plan for education activities. After a joint agreement on the terms of the visit and the activities that will be evaluated, the expert should be acquainted with the contents and objectives of the teaching unit. It is also necessary to agree in advance where the evaluation expert will sit and how his presence will be explained to students (Vizek-Vidović, 2003: 455).

Collaborative professional development – implies two or more experts/teachers who, for the purposes of their personal professional development and exchange of the experience attend each other's classes. Organized visits are not informal. Teachers provide feedback to each other and discuss the problems that might arise. This type of class attendance is organized within the school's expert team (individual department) and does not imply the standard evaluation.

Self-directing development – has two key characteristics: the teacher independently develops his skills without external evaluation, and the teacher plans independently what and when he will do according to his personal needs and abilities (for example: attending different courses, seminars, training, etc.). Teachers are motivated to join and attend education programs of the Ministry of Education because the teacher is awarded points, which are taken into account in the promotion process. One of the most effective forms of this evaluation model is recording the work in the classroom, analyzing it, and revising the plan after the analysis.

Administrative review is an evaluation model closest to the inspection, control, and similar unpopular terms whose negative connotation is conveyed to all other forms of evaluation work, although it should be avoided. The aim of the administrative review is actually to record the work of the teacher in the usual circumstances, but not to "catch them red-handed". There are certain rules as well. The review should be announced and public, in the sense that the teacher knows who is coming

to the school and what is their objective, what to expect, what will particularly be reviewed, and whether the collected data will be used in the formal evaluation – teacher assessment. This evaluation model is extremely useful if the data obtained is used to improve and plan future education work, and not simply to give immediate feedback to the teacher.

Evaluation of the teacher's work is a complex and demanding job, accomplished by an assessment of performance (Chen and Mathies, 2016: 85), but it is also generally appreciated profession as there is a consensus in the scientific and expert public that evaluation is necessary for the further improvement of the education processes. However, practical application is burdened with a series of difficulties. We will briefly explain typical evaluation problems – unclear purpose of evaluation, inadequate teacher involvement, insufficient self-evaluation, inadequate evaluation criteria, inexistence of the evaluation process, and insufficient qualification for evaluations – as follows:

Unclear purpose of evaluation – This encompasses superficial knowledge or even lack of knowledge of fundamental evaluation models – deliberate or coincidental interwoven diagnostic and formative evaluation. Namely, a teacher is expected to let someone with confidence in the classroom, expecting help in the improvement of his work. Later he realizes that the real purpose of the visit was, in fact, a completely diagnostic nature, for example – a follow-up on a parent's complaint.

Insufficient teacher involvement – The most significant teachers' complaint is certainly not having much influence on the evaluation process itself, especially if someone is outside of his or her education institution and even when his formal function is mere counseling.

Inadequate self-evaluation – Teachers' self-assessment and self-evaluation in the education practice are at the very least underutilized, unsystematic and unclear. Despite the fact that the researchers emphasize even the smallest children can profit from self-evaluation as well as adults (Mott, 2017: 26).

Inadequate evaluation criteria – The evaluation criteria are often based on personal preferences. Thus, some estimate the effectiveness of the teacher's work as to how they themselves would organize and perform a lesson instead of trying to be objective and focus on observing the results of the teacher's activities.

Inexistence of the evaluation process – Numerous attempts to raise the evaluation to a higher level have failed because the teachers themselves do not have crucial information and are not familiar with the conceptual-categorical apparatus which would allow them to understand the essence of the evaluation process in the classroom. On rare occasions when rules are established, experts believe that all teachers are aware of them and they do not need to be revised. The reality is different, of course. It is necessary to ensure that the purpose, criteria and evaluation process of the teaching work are regularly and clearly revised.

Inadequate qualification for evaluation – The school principal is responsible for evaluating the education process. When attending teachers' classes, he is often accompanied by an expert associate, pedagogue or psychologist. It sometimes happens

that the principal delegates the task of attending teachers' classes to them, without a clear agreement on the rules and the purpose of the evaluation process. It is important to emphasize that principals do not have sufficiently developed competencies needed to carry out a quality process of evaluation of the education process (Vizek-Vidović et al 2003, 459).

Despite these difficulties, it is necessary to encourage teachers to carry out evaluation and self-evaluation as well as to influence the perception of evaluation as truly desirable and indispensable part of the modern education process. Each of the stakeholders in the evaluation process – the state, students, teachers, parents, employers, the public – can contribute to its implementation and improvement. Because of the important role of each of the stakeholders, the next chapter will discuss their roles in the evaluation process.

STAKEHOLDERS IN THE EVALUATION PROCESS

It is important to emphasize, and it is often forgotten, that national authorities allocate significant financial and material resources to the education system. Evaluation and quality have become one of the major political issues related to education across the globe (Gurova, Piattoeva, and Takala, 2015: 346). Some national authorities, aware of the need for quality investment in the development of the education system, have developed in detail the evaluation methods. Some of the factors for evaluating the education system are the students' success in standardized tests. The problem arises when the existing knowledge needs to be evaluated in order to justify that the financial and material resources invested contributed to the development of high-quality knowledge and knowledge for the future (Mijatović, 2002).

Approaches to the evaluation of the education system are established on the belief that we are all able to learn both from our own successes and from our own mistakes. In fact, understanding our own successes and failures allows us to choose the activities appropriate for the changes we want to achieve. There is a natural aspiration of every education institution for better quality work and a stronger social impact in the community. Patton (1987) set out key questions that may help in defining alternative positions and approaches in evaluation:
   • Who is evaluation being conducted for? What are the expected results? (What do we expect from the evaluation?)
   • What do I want to achieve/reach through the evaluation process?
   • When will the results of the evaluation be utilized?
   • Where do we need to gather information?
   • How will the evaluation results be used?

Essentially, there are several very important criteria if the quality evaluation is to be achieved. First of all, the evaluation needs to be useful, and this implies that the evaluation collects credible, useful and timely information. Furthermore, the purpose of the evaluation is to identify successful and unsuccessful aspects of the functionality of the education process and to use the collected and analyzed information for better decision-making and future activities. Also, the feasibility of the evaluation is

very important, so it needs to be practical, cost-effective and sustainable. In addition, in order to avoid subjective and biased evaluations, it is necessary for the evaluator to be trained and highly educated for the area he is evaluating (Norris et al., 2017: 783). Also, there is a need for professional development of those who conduct the evaluation – knowledge of new evaluation tools such as questionnaires and tests, knowledge of statistical methods, technology-related assessment, accreditation and quality assurance standards, laws, policies and rules governing education (Chen and Mathies, 2016: 90). When evaluating, particular attention must be paid to ethical and legal standards regulating the evaluation process by taking into account responsible behavior towards those who are subject to evaluation as well as those who may be affected by the evaluation. Henneman et al. (2006, cited in: Norris et al., 2017: 759) suggests five elements which could help the credibility of the evaluation: evidence-based teaching standards, valid instruments, through education and training of those who evaluate or assess, the participation of several stakeholders in the evaluation, and well-established feedback process and targeted support. It is important to have clearly defined expectations and feedback.

Finally, the evaluation needs to produce accurate data, be useful to all stakeholders and to ensure credibility in the subject matter of the evaluation. The accuracy of the evaluation is linked to the strength of the evaluation plan, the data collection methods and the willingness to report the poor results of the evaluation just as satisfactory evaluation results would be presented. The evaluation deals with an assessment of whether the interventions are effective. In the history of education systems, we can identify numerous research and evaluation approaches and discuss how students' efficiency and academic success could be valued in relation to different stakeholders who have different understandings and seek for different evidence in the evaluation process. There are different perspectives on what is the primary function of the evaluation (Thistlethwaite et al., 2015: 292), so this paper discusses the aforementioned stakeholders – state, students, teachers, parents, employers, the public - and their perspectives in the following several paragraphs.

The state, or the Government, which finances most of the education institutions of each country, is interested in the relationship between enrolled students and those who are eligible to continue their regular education, i. e. the proportion of students who, for some reason, drop out, and time devoted to learning, i. e. the time necessary to fulfill all student and teacher obligations. Consequently, the evaluation of the education process from the point of view of the Government and the education authorities is focused on how many students complete the schooling in the foreseen time and how successful they are on the national academic achievement evaluation tests (State graduation exam in Croatia – Državna matura) with the reduction of costs and application of the international standards, at least partially.

Students evaluate the success of the education process in a completely different manner. For them, the evaluation is related to assessing the contribution of education to personal development and the preparation for taking responsibility in the world of work and in society. Therefore, the process of education must be linked to the students' personal interests. Of course, the process of education must be organized so

that every interested student can successfully complete the education in the foreseen time.

Teachers, though, are perceived as an education training ground established on good knowledge transfer, good learning conditions, and good correlation between academic achievement and evaluation.

Parents in the education process are perceived as preparing children for the world of work with the emphasis on academic achievements and positive and stimulative relationships between teachers and their children.

Employers in planning the assessment of the effectiveness of the education system emphasize that the learner is a product which acquires the necessary knowledge, skills, and work and moral practices required for effective acceptance, performance and improvement of work obligations for the benefit of the individual, future family, enterprise, and society.

The public perceives the education as the production of skilled labor or as training for a research career.

CONCLUSION

Defining the concept of evaluation is neither simple nor uniform, which makes it impossible to conclusively choose a single evaluation model and type to be the best possible for all stakeholders at the same time. However, taking into account all stakeholders in the evaluation process, we can conclude that evaluation as a pedagogical phenomenon is a complex concept which needs groups of professionals such as teachers, parents, students, administrative staff, employers and other stakeholders in order to be successfully carried out and completed.

Since the evaluation is the most commonly used to assess the quality and efficiency of the education system, and thus the teaching process, it is necessary to care for the adequate training and education of the evaluators as well as those involved in the assessment. It is important to emphasize that the very difference in stakeholders' perspective in the evaluation process highlights the complexity and unpredictability of the outcomes of the evaluation. Namely, the quality of successful evaluation is also the process of overcoming the observed weaknesses (Mencer, 2003: 72). In all systems which create new values, as well as in the education system, evaluation is necessary. Therefore, we can conclude that the quality evaluation has a different meaning for different people in the process and that the evaluation is closely linked to the processes themselves and the results of those processes. Quality assurance is a continuous process of proving, i. e. providing the warrantee that the education system will be ready to fulfill all objectives presented in its mission (Vroeijenstijn, 1995: 13). In addition, if we really care about the quality of education, then we need to be proactive in evaluating and providing incentive conditions for the full development of students (Chen and Mathies, 2016: 91).

This paper tried to emphasize the fundamental theoretical background related to the evaluation and the richness of different definitions and explanations of evaluation, models, and types of evaluation as well as the importance of evaluation of the

teaching process and the education system. Taking into account the different perspectives of all stakeholders in the evaluation process, we tried to outline the complexity of the process. Due to the complexity of the evaluation process, the abundance of different perspectives, we consider it is important to further explore this pedagogical phenomenon and thus contribute to both scientific and professional literature. With the critical approach to the aforementioned theories, we believe that those who are encouraged to evaluate their work, but also their students' achievements, will be innovative in their education practice and thereby contribute to the improvement of the education in the 21st century.

REFERENCES

1. Acheson, K. & Gall, M. (1987). Techniques in the clinical supervision of teachers. New York: Longman.
2. Biersteker, T. (2016). Self-Assessment Reviews and External Evaluation Exercises in IR Education. International Studies Review, 341-345.
3. Bognar, L. & Matijević, M. (2002). Didaktika. Zagreb: Školska knjiga.
4. Chen, H.-T. (1990). Theory Driven Evaluation. Newbury Park, CA: Sage.
5. Chen, P. D., Mathies, C. (2016). Assessment, Evaluation, and Research. In: D. C. Roberts, S. R. Komives (eds.), New Directions for Higher Education (pp. 85-92). San Francisco: Wiley Periodicals, Inc.
6. Cohen, L. & Manion, L. (1996). Research Methods in Education. London: Routledge.
7. Danielson, C. & Mcgreal, T. L. (2000). Teacher Evaluation to Enchance professional Practice. Alexandria: Association for Supervision and Curriculum Development.
8. De Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use – a literature review. Research Evaluation, 25(2): 161-169.
9. Feuerstein, R., Falik, L. H., & Feuerstein, R. (1998). The learning potential assessment device: An alternative approach to the assessment of learning potential. In: R. J. Samuda, R. Feuerstein, A. S. Kaufman, J. E. Lewis, R. J. Sternberg et al. (eds.) Advances in cross-cultural assessment (pp. 100-161). Thousand Oaks, CA: Sage.
10. Freidburg, V. (1978). Bilanz der Bidungspoliting. Pädagogische Beitrage.
11. Glatthorn, A. A. (1984). Diferentiated supervision. Alexandria: Association for Supervision and Curriculum Development.
12. Glickman, C. D. (1981). Developmental Supervision: Alternative Practice for Helping Teachers Improve Instruction. Alexandria: Association for Supervision and Curriculum Development.
13. Gould, R. (1989). Reflective Learning for Social Work Practice, Social Work Education, 8(2): 38-49.
14. Green, J. & South, J. (2006). Evaluation. Maidenhead, England: Open University Press.
15. Guba, E.G. & Lincoln, Y.S. (1989). Fourth Generation Evaluation. San Francisco: JosseyBass.
16. Gurova, G., Piattoeva, N., & Takala, T. (2015). Quality of Education and Its Evaluation: An Analysis of the Russian Academic Discussion. European Education, 47: 346-364.
17. Halmi, A. (2008). Programi evaluacije i evaluacijska istraživanja u društvenim znanostima. Zagreb: Naklada Slap.
18. Herman, L. J., Morris, L. L. & Fitz-Gibbon, C. T. (1987). Evaluation Handbook. Newbury

Park, CA: Sage.

19. Hounsell, D. (2003). The evaluation of teaching. In: Fry et al. (eds.) A handbook for teaching & learning in higher education (pp. 200-209). Second edition. London: Kogan page.

20. Kaneko, M. (1999). A methodological inquiry into the evaluation of smoking cessation programmes. Health Edu. Res., 14(3): 433-441.

21. Madelin, A. (1984). Osloboditi školu – Obrazovanje a la carte. Zagreb: Educa.

22. Matijević, M. (2005). Evaluacija u odgoju i obrazovanju. Pedagogijska istraživanja, 2(2): 279-298.

23. Matijević, M. (2007). Evaluacija u nastavnom kurikulumu škole. In: Previšić, V. (ed.). Kurikulum, teorije-metodologija-sadržaj-struktura (pp. 267-303). Zagreb: Zavod za pedagogiju.

24. Matijević, M. & Radovanović, D. (2011). Nastava usmjerena na učenika. Zagreb: Školske novine.

25. Mencer, I. (2003). «ISO 9001:2000 Protocol of Standards – Maybe the way in our next steps and/or the dream route to excellence», Workshop II: Development and Implementation of Quality Assurance System(s) in Higher Education (str. 65-79). Dubrovnik.

26. Mijatović, A. (1994). Osnove ustroja školskog sustava. Zagreb: Školske novine.

27. Mijatović, A. (2002). Obrazovna revolucija i promjene hrvatskoga školstva. Zagreb: Hrvatski zemljopis.

28. Mott, (2017). Leaving Room for Self-evaluation: Shaping our Dialogue in Early Childhood Education. International Journal of Choice Theory and Reality Therapy, XXXVII (1): 26-30.

29. Morrison, K. R. B. (1993). Planning and Accomplishing School-centred Evaluation. Norfolk: Peter Francis Publisher.

30. Naisbitt, J. (1984). Megatrends: 10 Perspektiven, die unser Leben verandern werde. Bayreuth: Hestia.

31. Norris, J. M., van der Mars, H., Kulinna, P., Kwon, J. & Amrein-Beardsley, A. (2017). A Document Analysis of Teacher Evaluation Systems Specific to Physical Education. The Physical Educator, 74: 757-788.

32. Patton, M. Q. (2002). Qualitative Research & Evaluation Methods. California: Sage Publications.

33. Patton, M.Q. (1987). How to Use Qualitative Methods in Evaluation. London: Sage.

34. Patton, M.Q. (1997). Utilisation Focused Evaluation. Thousand Oaks, CA: Sage.

35. Posavec, E. J. & Carey, R. G. (1992). Program evaluation: Methods and case studies. New Jersey: Prentice Hall.

36. Raphael, D. (2000). The question of evidence in health promotion. Health Promotion International, 15(4): 355–67.

37. Runciman, W., G. (1975). Social Science and Political Theory. Cambridge: University Press.

38. Scott, D. (1998). Qualitative approaches to evaluations of health-promoting activities. In: D. Scott and R. Weston (eds.). Evaluating Health Promotion (pp. 31-49). Cheltenham: Stanley Thornes.

39. Scriven, M. (1966). The Methodology of Evaluation. Indiana: Prude University.

40. Springett, J. (2001). Participatory approaches to evaluation in health promotion. In: I. Rootman, M.S Goodstadt, B. Hyndman, D.V. McQueen, L. Potvin, J. Springett and E. Ziglio (eds.). Evaluation in Health Promotion. Principles and Perspectives (pp. 83-88).

Copenhagen: WHO Europe.
41. St Leger, A. S., Schnieden, H. & Walsworth-Bell, J.P. (1992). Evaluating Health Services' Effectiveness: A Guide for Health Professionals, Service Managers and Policy Makers. Milton Keynes: Open University Press.
42. Thistlethwaite, J., Kumar, K., Moran, M., Saunders, R. & Carr, S. (2015) An exploratory review of pre-qualification interprofessional education evaluations. Journal of Interprofessional Care, 29(4): 292-297.
43. Tones, K. (1997). Beyond the randomised controlled trial: a case for 'judicial review'. Health Education Research, 12(2): 1–4.
44. Tones, K. &Tilford, S. (2001). Health Promotion: Effectiveness, Efficiency and Equity. Nelson Thornes: Leeds.
45. Tyler, R. W. (1989). Educational evaluation: Classic works. Boston: Kluwer.
46. Vizek-Vidović, V., Vlahović -Štetić, V., Rijavec, M., & Miljković, D. (2003). Psihologija obrazovanja. Zagreb: IEP-VERN.
47. Vroeijenstijn, A. I. (1995). Improvement and Accountability: Navigating Between Scylla and Charybda. London and Bristol, Pennsylvania: Jessica Kingsley Publishers.
48. Weber, M. (1949). The methodology of the Social Sciences. London: Glencoe Free Press.
49. Weiss, C. H. (1972). Evaluation Research: Methods of Assessing Program Effectiveness. Englewood Cliffs, NJ: Prentice Hall.
50. Weiss, H. C. (1998). Evaluation. New Jersey: Prentice Hall.