

The Architecture of Programme Evaluation*

Giorgos Maggopoulos

(Hellenic Open University, Greece)

Abstract: The purpose of this paper is to highlight the structural elements of programme evaluation. In order to make the purpose of our research feasible, we focus on three main objectives: first, we attempt to deconstruct the most important evaluation models or approaches into their key elements, as those emerge in the relevant literature. Second, we outline the most important areas of decisions that the evaluator is required to make when designing the evaluation, as those emerge from the broader theoretical framework and the implementation of programme evaluation. Third, we venture a reconstruction of the above, in order to demonstrate, on the one hand, the architecture of the evaluation endeavour and, on the other hand, to compose a new proposal or evaluation model. The research is bibliographic and is based on the deconstruction of goal-based evaluation, responsive evaluation, theory-based evaluation, utilization-focused evaluation, empowerment evaluation, and constructivism evaluation. At the same time, emphasis is placed on seven areas of important decisions that the evaluator is required to make when designing the evaluation. These areas are called i) initial concerns, ii) intentions, iii) description of the object of evaluation, iv) questions, v) answers, vi) report, presentation and dissemination of findings and vii) meta-evaluation. Next, we attempt the dialectical synthesis of the above, making the basic structural elements of programme evaluation clear and recognizable. The above actions allow the formation and development of a holistic model of programme evaluation.

Key words: programme evaluation, programme evaluation models, holistic model of programme evaluation

1. Introduction

Programme evaluation is a complicated, complex and multidimensional task. A lot of scholars have made a critical contribution to the development of the theoretical and methodological frameworks of programme evaluation. The direction that each evaluation project follows, from the design phase to the dissemination of the study findings, seems to depend significantly on the theoretical and methodological background that the evaluator or evaluation team introduce and utilize in their study. The evaluators or evaluation teams *borrow from* or draw on conceptual tools within this background, in which their theoretical and methodological choices are grounded and legitimized. In short, the overall organization of the evaluation is based on conceptual tools provided by the evaluation theory. The choices made by the evaluator outline the “attitude” (Stake, 2004) of the evaluation, its basic theoretical and methodological assumptions and, of course, its epistemological orientation.

In the theory of programme evaluation, concepts such as *objectives, response to needs, programme theory,*

* The article is part of my doctoral dissertation that was supported at the Pedagogical Department of Primary Education of the University of Crete and has been published in Greek in 2015.

Georgios Maggopoulos, Dr., Collaborating Teaching Staff, Hellenic Open University; research areas: educational evaluation, programme evaluation, adult education. E-mail: maggopoulos@gmail.com.

negotiation, communication, participation, utility and empowerment play a key role. In fact, each of them constitutes the central concept in very popular evaluation models or approaches. Goal-based evaluation, responsive evaluation, theory-based evaluation, utilization-focused evaluation, empowerment evaluation, and constructivism evaluation contribute significantly to the formation of the theoretical framework of evaluation.

However, the evaluator is required to make more decisions that are very important. The theory of programme evaluation helps a lot in the identification and careful observation of relevant issues. Valuable help in this work is offered by the *standards* and *principles* that both scholars and evaluators' associations formulate, significantly enriching the relevant discussion. A characteristic example is Scriven's (1991, 2005) Key Evaluation Checklist (KEC), in which the evaluator is required to make important decisions on fifteen critical issues related to the evaluation. Another example is Stufflebeam's (2002) CIPP evaluation model checklist, which suggests that the evaluator and the participants take specific actions with respect to ten sub-dimensions of the evaluation. The positions of the Joint Committee (1994) are also widely recognized, as they seem to have been adopted to a considerable extent by the scientific community of evaluators in several European countries, such as Germany (Beywl, 2003) and Switzerland (Widmer et al., 2000). However, many other design suggestions appear in the existing literature. Suggestions made by Sanders (1992), Pawson and Tilley (2000), Knox (2002), Fitzpatrick et al. (2004), Owen (2007), and Mertens (2009) are useful and insightful. Each evaluator is asked to structure a comprehensive strategic proposal (Yin, 1994), in which the theory, the methods, the resources, the use, the participants' roles (Shaw, 1999; Robson, 2000; Marby, 2003), the implementing and/ or funding agent's policy and intentions, and their personal skills and values (Fitzpatrick et al., 2004) are connected.

2. Methodology

The purpose of this study is to reveal the architecture of programme evaluation, as it results from the synthetic utilization of the aforementioned concepts. The first goal of the research is to deconstruct each of the six evaluation models or approaches. The limited space at our disposal does not allow expanding to all sub-dimensions that each model inserts into programme evaluation. We will try, however, to answer the following research question: what does each of the evaluation models under analysis bring to programme evaluation at a conceptual level? The second goal is to make reference, in a brief but comprehensive manner, to the design decisions that the evaluator is required to make. Taking into account what was briefly mentioned in the previous section regarding design decisions, we will try to answer the following research question: how can we redefine these decisions in a new sequence of actions? The third goal is to restructure the above. The research question we ask is the following: how can we utilize, on the one hand, the conceptual deconstruction of the evaluation models under study and, on the other hand, the design decisions, in order to reveal the architecture of programme evaluation? Achieving the above objectives allows us to submit our own programme evaluation proposal, which we hope adequately represents its overall function. Although this endeavour does not claim exclusivity in the field of evaluation, it may be useful to anyone intending to design and implement a programme evaluation. The above suggests that the present research has a bibliographic character.

3. Review of Evaluation Approaches and Models

3.1 Goal-based Evaluation

Goal-based evaluation traditionally focuses on examining goals and trainees-students' performance. The key

question it raises is to investigate the extent to which learning performance, usually perceived as changes in behaviour, is consistent with the desired learning outcomes (objectives) of the programme. In other words, goal-based evaluation primarily measures the degree to which the objectives of the programme are achieved, as reflected in trainees-students' performance on normative or criterion-referenced tests (Hambleton, 2000), as well as other procedures that measure their performance (Madaus & Stufflebeam, 2000; Bonniol & Vial, 2007).

The rationale of the model is simple and easy to use. Objectives are transformed into quality-success indicators or criteria (Scriven, 1991), based on which the effectiveness of the programme is evaluated or, according to Bonniol and Vial (2007), skills and knowledge acquisition is checked. The comparison of intended and actual learning achievements determines the quality of the programme and consequently the agreement between the intended and the implemented (Alkin & Christie, 2004). A programme is considered successful when the actual learning achievements are as consistent with the objectives as possible.

Goal-based evaluation follows the following course: First, the goals and objectives of the programme are set and classified, and their content is clearly defined in terms of behaviour (Bonniol & Vial, 2007). Next, the situations in which the achievement of goals can be detected are identified. Then, measurement techniques are selected or developed and the data are collected. In fact, some claim that the emphasis placed on measurements provides increased rigour to the goal-based evaluation model. The evaluation is completed by comparing data (student performance) with formal behavioural goals (Fitzpatrick et al., 2004).

3.2 Responsive Evaluation

The criticism that the traditional (goal-based) evaluation model received, which is in agreement with the criticism *positivism* receives (Clarke, 1999; Stake, 2000; House, 2001; Abma, 2004; Knaap, 2006), resulted in the emergence of new models of evaluation, especially from the 1970s. Responsive evaluation, which was developed by Stake (1973), is one of the main alternative forms of evaluation that emerged at that time and it seems that it holds a significant position in the theoretical background of programme evaluation to date.

Responsive evaluation (Stake, 1973; Huebner & Betts, 1999; Abma, 2004; Knaap, 2006), attempting to overcome the weaknesses or shortcomings of traditional evaluation, attaches particular importance to the uniqueness and complexity of the programme and the programme as a whole. It involves the pluralistic representation of the concerns, views, experiences, values and needs of the people who are actively engaged in the implementation of the programme (Stake, 2004). Responsive evaluation is a research project documenting the quality of the programme, the main characteristic of which is the focus on the problems that are recognized as primary by the people directly involved. The fact that its course is directed by the interests and needs that the participants reveal is its dominant feature. The specific model is suitable when the aim of the evaluation is to explore and understand the quality of the processes within the programme (Stake, 2004).

Responsive evaluation seems to be inspired by a new philosophy. The question that one may ask is the following: what new can the shift of focus from examining the degree of goal achievement to studying the issues raised by the participants bring to evaluation?

The innovative element introduced by responsive evaluation lies in the emergence of a new concept that directs the process of the evaluation. The evaluation agenda now consists of *issues* raised by those directly involved. The topics usually reflect obstacles and the participants' concerns regarding the implementation of the programme. They originate from the theory of the programme, its context, the participants' needs (Stake, 2000), the implementation problems, the schedules, the input-output relationship, the goals of the evaluation and many

other issues, as well. In general, anything that inhibits the effective function of the programme is a potential element for examination (Stake, 2004).

The evaluator is responsible for the description of the *issues* under study; this responsibility essentially refers to the description of the programme, through which the evaluator makes the participants knowledgeable about the rationale of the programme. According to Stake (2004), the evaluator's observations should focus on the intentions or objectives of: i) the current situation ii) the intervention, and iii) the outcomes of the programme. At each of the three levels of the observation, the evaluation criteria, on which the evaluation judgments will be based, should also be developed.

3.3 Theory-based Evaluation Model

The evaluation that does not consider the transformational process which takes place between inputs and outputs has the following weaknesses: first, it provides little or no information about the reasons or causes that make a programme successful or unsuccessful, and, therefore, the findings of the study do not make a substantial contribution to the improvement of the programme (Chen, 1990; Rogers, 2000). Second, it fails to distinguish whether the failure of the programme is due to a theory of limited validity or because of inappropriate implementation. Third, it shows a weakness in distinguishing the effective from the ineffective or harmful elements of the programme (Donaldson, 2007). The so-called *black box* of the programme remains sealed. Simply providing information about the *input-output* relationship is not enough for important questions to be answered in a compelling manner (Stake, 2004; Alkin & Christie, 2004).

Several scholars believe that the above weaknesses can be addressed when the evaluation utilizes the theory in which the programme is grounded. In other words, they believe that programme theory enables the evaluator to identify the elements that can provide answers to two critical questions, *how* and *why* the programme succeeds or fails in its mission (Weiss, 2004; Robson, 2007).

The semantic content of the term *theory*, as perceived within this evaluation model, refers to a set of propositions that, on the one hand, guide the nature and function of the programme and, on the other hand, *account for* its outcomes (Rogers, 2000, p. 211). That means that the evaluation draws on two sets of suggestions from the theory of the programme. The first is about the structure of the programme, i.e., the elements that constitute the intervention, the implementation and the achievements; the second illustrates the mechanism that links the intervention, the implementation and the achievements.

Chen (1990) named the theory regarding how the structure of the programme should be *normative*. The normative theory establishes the structure of the activities, and develops or suggests strategic actions, through which the goals of the programme are expected to be met. Whether a programme ultimately achieves the expected results depends on: i) whether the structure of the normative theory is appropriate and efficient in relation to its intended outcomes, and ii) the degree to which the programme is implemented according to the design. The theory regarding the mechanism that connects the structural elements of the programme is called *causative*. The descriptive theory demonstrates the reasons why the programme works or not, by illustrating the causal relations between intervention and achievement, and providing information about the conditions that favour or limit the occurrence of causal relations.

The evaluation that is grounded in programme theory can be either *normative* or *causative*. The purpose of the evaluation that is grounded in normative theory is to examine the consistency between the theoretical structure

of the programme and its implementation. The information provided helps the evaluator and the participants determine whether the success/failure of the programme is due to a valid/invalid theory (Rogers, 2000), an appropriate/inappropriate environment and adequate/inadequate implementation. The findings of this investigation are used for feedback and monitoring, which lead to making relevant decisions; nevertheless, the findings may also contribute to the extension of the existing theory (Pawson & Tilley, 2000).

Examining the consistency between the intended and the implemented, however, is not enough for the success or failure of the programme to be justified. The quality of the interventions does not necessarily lead to the quality of the produced result. Therefore, the internal quality of the programme is not enough to determine its overall value. Thus, causative evaluation is necessary, as it provides information both about the effect that the programme has on the beneficiaries, and about the way in which it creates or produces this effect (Owen, 2007). The improvement and development of the programme structure, the understanding of its effectiveness and the production of knowledge are all mainly grounded in causative evaluation (Chen, 1990; Clarke, 1999; Stame, 2004; Donaldson, 2007).

3.4 Utilization-Focused Evaluation

In each evaluation, human and material resources are invested, in order to bring about changes that develop the programme and policies that improve the situation of the people. Therefore, the evaluation is implemented in order to be used not only by the contracting body, but also by those who show relevant interest. In fact, many scholars, who want to emphasize the crucial importance of the utilization of the study, express the view that evaluation as an end in itself does not make sense. The concept of “utilization” or “utility” is part of the theoretical background of evaluation and at the same time a criterion for its professional and responsible implementation (Joint Committee, 1994), especially when timely information and impact is its intention.

During recent decades, a broader evaluation approach focusing on utilization has emerged (utilization-focused evaluation). Its main feature is the shift of the study's interest to the utilization of the evaluation results and procedures. This does not imply that the evaluation of the quality of the programme is not of interest; evaluation is considered effective and successful when it is useful to people and society as a whole. The above view is largely grounded in the positions of *pragmatism*, which attributes a *practical interest* to evaluation (Habermas, 1990), a fact that attracts the attention of a number of evaluators and politicians. When evaluation is based on a variety of individual, collective and institutional expectations of utilization (Robson, 2000; Cousins, 2003) that aims at improvement, accountability and knowledge production, then it seems that the benefits are significant, giving the evaluation legitimacy and rigour. The concept of utilization is consistent with other concepts that are equally important, such as the concepts of *participation*, *collaboration*, and *response*, which, when combined together, give a flexible, dynamic and creative profile to evaluation.

Taking the above into account, it could possibly be argued that the utilization of the study becomes a key parameter for the evaluation. The relevant reflection is constantly enriched, inevitably bringing to the surface critical questions, such as: Who will use the evaluation? What aspects of evaluation are worth utilizing? What is the form of use? Which relationships among the stakeholders promote the use? What are the key factors that influence utilization (Sanders, 1992)? How does utilization relate to the other dimensions of evaluation (Clarke, 1999)? What are the restrictions on its use?

The discussion regarding who will use the evaluation directly refers to the groups of participants. Some groups are more interested in evaluation than others. The participants who care mostly about the evaluation and

expect benefits from the utilization of its findings are the *intended users* (Patton, 2004). Patton sees in this category of participants a direct, personal interest and/ or responsibility for the utilization of the evaluation, which he calls a *personal factor*. Groups that show personal interest, positive attitude, enthusiasm, care and initiatives, actively participate in the evaluation. They seek appropriate information (Patton, 2003) to enhance their effectiveness, their judgement and decisions. In this way, they try to limit as much as possible any uncertainties in their choices (Patton, 2000). Activating the personal factor, which undoubtedly involves political as well as cultural dimensions, may constitute a challenge for the evaluation.

Regarding the dimensions of the evaluation that are suitable for use, the relevant discussion refers to its results and procedures. The results are expected to support decision-making mainly by policy makers, programme executives, as well as educational or other qualified staff (Patton, 2003). Although the findings are the most anticipated useful aspect of the evaluation, it is not the only one. The possibility of utilizing the evaluation procedures is less obvious. Several researchers, driven by the above observation, argue that evaluation as a process involves a potentially metacognitive character, as the participants' and the body's learning about the theoretical assumptions, values and methodology of evaluation is not excluded. Patton (2004) perceives learning which results from the utilization of processes as a change in the thoughts and behaviours of those involved in the evaluation, and as a change in the organization and culture of the programme; changes, which contribute significantly to further, multileveled development of both individuals-groups and institutions-organizations (Patton, 2003; Abma, 2004). It should be highlighted, however, that the benefits of the procedures are not a natural consequence of participation, but are based on the personal characteristics of the users and the operation of the programme's implementing body as a learning organization.

3.5 Empowerment Evaluation

The expectations of those involved in the evaluation are particularly high today. Undoubtedly, the acquisition of information that leads to improvement, accountability, and, why not, production of knowledge constitutes the fundamental and most expected aspect of any evaluation. Fetterman (2000), however, believes that conducting an evaluation can lead to other forms of utilization. Particular emphasis is placed on the acquisition of additional knowledge and skills by the participants. In other words, he believes that conducting a study makes it possible for participants to learn and empower themselves, through enhancing their self-efficacy and self-determination (Fetterman, 2005), their active participation in decision-making, their ability to communicate and the ability to find alternative solutions.

Evaluation under specific conditions is able to integrate those involved in processes of *lifelong learning* or *adult education*. The question that reasonably arises is under which conditions the evaluation acquires an educational dimension. The answer is simple: when the evaluation attempts to enable the participants to improve the programme on their own, then the evaluation works educationally, as well. More specifically, when the people of the programme acquire additional skills and knowledge or improve the existing ones through their active involvement in the evaluation, then their empowerment leads to the improvement of the programme. Therefore, it seems that evaluation, learning and improvement are interconnected.

If the above connection is accepted, then further development of the programme is extremely difficult, if not unlikely, to be *inserted* into the programme, as it depends on the actions and decisions of the agents. Under these circumstances the task of evaluation cannot be externally guided. This view is strongly supported by Fetterman (2005), who states that empowerment evaluation should be designed in such a way that *helps people help*

themselves improve the programme through self-evaluation and reflection. It is now obvious that the evaluation takes place along with the participants and not in their absence. The question that would probably be asked is why the focus of the evaluation and improvement is transferred to the participants.

The participants know better than anyone the strengths, weaknesses and problems of the programme; the special position they hold in the programme essentially integrates problem solving, promotion of changes and improvement in their activities. For many researchers, participants' active involvement in the programme "imposes" in some way their active participation in the evaluation. More specifically, according to the empowerment evaluation, it is necessary for the participants to have an active participation in: i) the critical recording of important aspects of the programme, but also the evaluation of the value of the programme at the specific time, ii) the definition of the programme objectives, in order to outline its future orientation, iii) the development of strategies to achieve the selected objectives and iv) the methodology based on which the degree of achievement of the objectives will be examined and documented (Fetterman, 2000, 2005). The highly active role of the participants promoted by the empowerment evaluation stems from the belief that they can offer a deeper understanding of the programme, its mission, its theory, its implementation strategies, its problems and its context. The need for change, improvement, continuous learning and social justice (Fetterman, 1995, 2005) cannot be met if participants play second fiddle to evaluation.

Participants become protagonists of the evaluation. They are the ones who have to identify the problems and at the same time find real and applicable solutions. Many expectations are placed on the people of the programme and it may be extremely difficult, if not unlikely, to meet such a difficult and complex task. The transfer of initiatives and decisions to the participants may puzzle or even "frighten" both themselves and the evaluators. Undoubtedly, they need help and support. At this point the evaluator, whose responsibility is to transfer skills and knowledge of the evaluation procedures to the participants, intervenes (Preskill & Preskill, 1997). In order for the people of the programme to take responsibility, however, there is a need for trust, opportunities for reflection, independent, free and creative action (Fetterman, 2005), reinforcement and support, so that they become proficient in problem solving and decision making (Clarke, 1999).

3.6 Constructivism Evaluation

In constructivism evaluation (Lincoln, 2003; Guba & Lincoln, 2004), the role of the participants is crucial throughout the study. However, their active involvement in evaluation, an unquestionably scientific and political process, raises a variety of questions. Very likely, the most central one is about the reasons why the interest of evaluation is sometimes to a greater extent and sometimes to a lesser extent transferred to participation and negotiation.

The emergence of these concepts in evaluation is largely based on developments that have occurred at the *epistemological* and *ontological* levels in recent decades within the field of social research. If the philosophy of science highlights that there is no absolute and objective reality or it exists but, due to inherent weaknesses, it is not possible to fully conquer it, then evaluation has to face a very critical problem. How is it legitimate to make a value judgment of the processes and results when it has missed the opportunity to invoke an objective model of *truth*? Transforming evaluation into a *communicative act* (Habermas, 1997) seems to offer a solution to the impasse. Evaluation acquires *validity claims* or *objectivity*, which is grounded in the mutual commitment of those involved, as built in the context of rational dialogues and arguments (Habermas, 1990). *Validity claims*, on the one hand, transform experiences into events and, on the other hand, allow the formation of judgment, i.e., the

estimation of the success or failure of certain actions (Habermas, 1990). Drawing on arguments from Habermas's work, it seems that many evaluation theorists estimate that, if the objective and absolute truth or reality in the field of social sciences does not exist or we cannot fully approach it, then for the evaluation to take place it is necessary to conclude agreements regarding the content or form of the constructed objective *reality* or *truth*. It is basically a contract among those involved in evaluation, the legitimacy of which is based on consensus and rationality (Habermas, 1990). At the same time, it appears that the consensus is binding on both the evaluators and the participants. The evaluators draw on arguments, concepts, principles and tools to direct their work. The participants know what is expected, have the responsibility to implement the actions and, of course, are asked to account for them. Mutual agreements legitimize evaluation as a whole, as they constitute the basis on which drawing conclusions, forming value judgments, making decisions and taking actions can be achieved.

The necessity for participation, in addition to epistemological references, also has a political background. As nowadays modern societies are characterized by democracy and pluralism and at the same time *civil society* is becoming increasingly stronger, several social groups, disadvantaged or not, want to participate in decision-making and demand equal treatment. They demand to have the same opportunities, to submit their opinion, to have their needs taken into account and respected, to have access to the utilization of social goods. In the context of this debate, which has developed strongly in social sciences, evaluation is linked to the democratic principles of participation, equity, accessibility and justice (House & Howe, 1999; Lincoln, 2003; Fitzpatrick et al., 2004), support and response to the needs and values of all participants, strong or weak (House & Howe, 1999; Lincoln, 2003; Preskill, 2003; King, 2003), and the consideration of the local political, social and value context (Shaw, 1999, Cousins, 2003, Stake, 2004). It is now considered necessary to *listen to* the positions of the participants, and ensure participation and dialogue, in order for the *values* or *concepts* that will guide the evaluation to be revealed and agreed to by the participants. Those that take part in the process are perceived as an internal element of the study. They have valuable information and interpretations regarding the quality and weaknesses of the programme. They add knowledge and experience to the evaluation and at the same time enhance its value. Many researchers (Joint, 1994; Robson, 2000; Fitzpatrick et al., 2004) believe that participation enhances the design, the process of addressing daily issues (practical participatory evaluation) (Cousins, 2003), the implementation, the validity, the transformation of practices (transformative participatory evaluation) and the utilization of the study.

4. Design Decisions

4.1 Introduction

Critical decisions are made while designing the evaluation. Indicatively, it is stated that the evaluators usually have the main responsibility in choosing the appropriate questions, criteria and the methodological strategy in general (Stake, 2004). They are also responsible for identifying the available resources, ensuring the participants' cooperation (Knox, 2002), promoting negotiation, identifying political influences, and clarifying the direction of the evaluation. They take into account the priorities, the intentions, the point of development of the programme (Robson, 2000) and the practical obstacles posed by the environment. They ensure evaluability if deemed necessary, meta-evaluation, the work-task plan, the reduction of distractions and the development of interpersonal relationships (Sanders, 1992). They keep the procedures applied procedures simple, comprehensible and accessible.

In the following sections, we try to place the important decisions that the evaluators have to make in a sequence of actions. The proposed evaluation process consists of seven main stages; each part of it can either consist of component parts or be autonomous. The suggested course, although it can be considered linear, as the evaluators will encounter extreme difficulties in case they ignore any of the illustrated stages, is usually flexible and spiral in nature. Its flexibility lies in the ability of the evaluator and the contracting body to either abandon some of the subcomponents of the suggested stages, or choose among different priorities. For example, evaluability is not always necessary during the stage of initial concerns. The spiral implementation occurs when each new evaluation can either follow the choices made in the previous one, thus allowing, among other things, comparisons at different levels, or modify the basic conceptual tools it uses; that is, to be enriched with new questions, change the criteria and even change the methodology of the evaluation.

Outlining the evaluation process is a remarkably *delicate* endeavour, as the priorities, the abilities and the epistemological position of the evaluator are reflected on the stages of the study and their content. Of course, in no case should it be assumed that the proposed sequence of actions fully represents the possibilities or options that the evaluator has; it is simply a version of a multitude of alternatives. In the next sections, a brief reference is made to each of the stages and sub-stages of the proposed evaluation course.

4.2 Initial Concerns

The first stage, called *initial concerns*, represents the need to attend to the sub-dimensions regarding the programme, the participants and the evaluation itself. The distinctiveness of this stage lies in the fact that most of the individual aspects are able to lead to evaluations that sometimes constitute autonomous studies and sometimes are included as subsections in the overall evaluation. It should be made clear that these dimensions have the potential to significantly influence the course and nature of the evaluation. If not taken into account in a timely manner, they may adversely affect the overall course of the study, its quality, its conclusions and, of course, its utilization. More specifically, the *initial concerns* include:

- The *policy*: The nature of the programmes but also the evaluation introduces policy issues in any evaluation project (Clarke, 1999; Fitzpatrick et al., 2004; Preskill, 2003). The evaluators, in order to protect the independence and quality of their study and of course their professional and scientific prestige, must take them seriously into account (e.g., policy expectations, their explicit and implicit intentions etc.). Open communication and regular cooperation between the evaluator and both the policy makers and the other participants throughout the course of the study, as well as immediate notification to all stakeholders of any modifications that occur during the course of the evaluation, seem to offer the shield of protection the evaluators need to *resist* pressure and prevent misinterpretations and misunderstandings (Sanders, 1992; Robson, 2000).
- The *evaluability*: Evaluators must know if and to what extent the object of evaluation is ready to be evaluated. There are many cases, in which various uncertainties (Scriven, 1991; Shaw, 1999; Trevisan & Huang, 2003; Fitzpatrick et al., 2004) significantly limit the feasibility of the evaluation. If evaluators start the research in a timely manner, study the theory of the programme and examine its administrative and political context, then, to a significant extent, they will provide solutions to the aforementioned problems. Even if some people consider the evaluability unnecessary, especially in cases of well-designed programmes, the evaluators are advised to consider even briefly the possibility of uncertainties.

- The *needs of the beneficiaries*: The evaluation acquires the appropriate focus when it is known what the beneficiaries expect, under what conditions and for what reason, and when the evaluator has an overview of the needs of both the participants and the implementing body of the programme. Needs assessment provides valuable knowledge and provides evaluation with useful information (Robson, 2000; Knox, 2002; Alkin & Christie, 2004). More specifically, it helps significantly in identifying the reference population and diagnosing their educational needs; it contributes to the examination of the relevance of the programme and assists in its defence; it supports the selection of evaluation criteria and provides the opportunity to detect problems that need further consideration. The ultimate goal of the needs assessment is to promote appropriate changes in the programme, so that the services offered are more and more in line with the needs of the beneficiaries.
- *Ethics and biases*: Evaluation is not just a technical process. Many people with different roles, needs, views and interests are involved in the study, offering valuable information, knowledge and help. The need for dialogue, consensus, synthesis of different views and public information gives social and communicative content to the evaluation. Therefore, the management of the social and communicative dimension of the evaluation, which significantly concerns each of its stages, cannot ignore participants' basic political and professional rights. It is obvious that, when there is no institutional guarantees, data collection starts only after the consent of the participants. Important responsibilities of the evaluator (Scriven, 2000; Fitzpatrick et al., 2004; Owen, 2007) are the protection of the personal data and privacy of those involved, the non-“criminalization” of weak particularly participants (Shaw, 1999; House & Howe, 1999) and the promotion of the well-being of society as a whole.
- The *management*: Evaluation is a research activity that takes place either when provided for in the institutional regulation of the programme, or when it is decided by the sponsor or the executives. In any case, one or more persons of the contracting and/ or implementation body of the programme decide on its implementation and at the same time are asked to make important decisions regarding: i) who will be responsible for conducting the evaluation, ii) what resources will be available (Sanders, 1992; Fitzpatrick et al., 2004; Owen, 2007), iii) what will be the *deliverables* of the evaluation (Stufflebeam, 2002), iv) who will monitor the course of the study and e) the conclusion of a *contract*, when the parties agree on the above. While some argue that the money, human resources and time wasted on evaluation could be used directly to further the development of the programme, it is becoming increasingly clear that evaluation contributes significantly to the improvement of the programme. The gains from a well-organized and well-implemented evaluation seem to outweigh the costs to a significant degree.

4.3 Intentions

The second stage, called *intentions*, essentially reflects the position of the study, the expectations placed on the research and the people involved in it. More specifically, the *intentions* include:

- The *values*: The values that frame the evaluation are distinguished into two categories: the values regarding the quality of the study and the values regarding the quality of the programme. With respect to the values of the first category, they seem to be directly related to the epistemological direction of the study. However, it is good for the evaluation to be characterized as much as possible by validity, reliability, independence, transparency, pluralism, democracy and decency. The above values add prestige, trust and acceptance to the evaluation, as they protect the process, the results and of course the

evaluator from any misinterpretations and misunderstandings. Regarding the values that characterize the quality of the programme and on which value judgements are based to a significant extent, the evaluation can focus (Stake, 2004) on relevance, effectiveness, monitoring, efficiency, sustainability, utility and added value. Each of the above values is able to give evaluation its own direction, influencing the whole evaluation project. Each group of participants tries to promote their own values in the evaluation, and consequently they evaluate the quality of the programme according to their own criteria. The evaluators, based on their epistemological, methodological and theoretical assumptions, are invited to illustrate the complex form of the quality of the programme, identify the dominant values, examine their relevance to the programme, inform the participants and, when they deem it necessary, make a value judgement themselves.

- The *purpose*: Each evaluation is unique, resulting in a surprisingly heterogeneous goal-setting process (Patton, 1990; Joint Committee, 1994; Stufflebeam, 2001; McGuire, 2002). It is sometimes possible to focus on meeting a specific goal, and other times on meeting more than one goal. However, the situation becomes even more complex when the goal or goals are broken down into objectives. Naturally, each evaluation in a specific space-time context has its own objectives, which may show significant deviations even from relevant studies. Despite the strong differences among objectives in evaluation, goal-setting seems to focus mainly on (Stufflebeam, 2001; Robson, 2000; Owen, 2007): i) accountability, considering that the examination (final evaluation) focuses on the outcomes of the programme both in terms of achievements (effectiveness) and in terms of financial return (efficiency) ii) improvement, as the focus is mainly on the formative examination of processes (Patton, 1990; Fitzpatrick et al., 2004; Mertens, 2009) iii) knowledge production, considering that the focus is transferred to understanding the implementation of the programme, as well as exploring the relationships among the participants, the objectives, the processes and the context. When the production of knowledge takes place, it helps in the development of both the programme and the respective research field. Goals are a focal point for evaluation. Consistency among the goal, the methodological strategy, questions, reference criteria, and utilization has a determining role in the quality of the study. Goal setting in evaluation also reflects the developmental course of the object of evaluation and outlines the *vision* for the progress of the programme.
- The *role of the participants*: The participants (Robson, 2000, Stufflebeam, 2001, Knox, 2002, Abma, 2004) play an important role in the implementation of the evaluation as a whole. Their direct and indirect engagement in the programme makes them valuable partners. However, if the evaluators want to make the most of the participants, while at the same time maintaining the *monitoring* of the evaluation project, they should examine who participates, how they perceive their participation, what authority they have and what their expectations are (House & Howe, 1999). The participants' active involvement adds knowledge and experience (Shaw, 1999; Cousins, 2003; Stake, 2004; Fitzpatrick et al., 2004) to the evaluation; it enhances its utility and validity; it promotes democracy, learning, social justice and equity; it promotes bidirectional accountability and reciprocity in taking responsibility and, finally, it reduces any resistance while increasing the very value of the evaluation.
- The *role of the evaluator*: Evaluators (Clarke, 1999; Patton, 2003; Lincoln & Guba, 2004) play a key role in evaluation. They are the ones who will give shape to the quality of the programme, will present its *truth*, its strengths and weaknesses, will significantly specify its development prospects and will

define the cooperation with the participants. “*Good*” evaluators, in addition to the knowledge and skills (Patton, 1990; Sanders, 1992; House & Howe, 1999; McGuire, 2002; Fitzpatrick et al., 2004) that they must have, acknowledge the limitations of their potential; they know their predispositions, their strengths and *biases*; they investigate the way others see them and accept advice; they critically address the political and organizational context; they follow the codes of professional ethics; they give perspective to the development of the programme; they understand the expectations invested in the evaluation; they are trustworthy and fair, gain all stakeholders’ trust and respect them all; they are interested in public welfare, democracy and justice, rather than serving individual interests.

4.4 Description of the Object of Evaluation

In the third stage the *description of the object of evaluation* takes place, which is necessary in every evaluation (Sanders, 1992; Stufflebeam, 2001; Knox 2002; McGuire, 2002; Mertens, 2009). The evaluators must understand the various aspects of the programme and transfer them to the participants (Joint Committee, 1994). Evaluators will be particularly assisted in their work if they obtain information from the participants, critically examine the operating regulation, the administrative and other documents related to the programme, study the theory of the programme and the relevant literature, and communicate with experts on the programme or researchers. The key components of a programme are the broader social, cultural and political context, the inputs, the processes, the outputs and the feedback of the programme (Stufflebeam, 2001; Chen, 2004). The more detailed and *deep* the description is, the more the evaluators’ ability to understand the programme as a whole is strengthened, adequately substantiating the study and their value judgments.

4.5 Questions

The fourth stage, called *questions*, is central to evaluation (Robson, 2000; Pawson & Tilley, 2000; Stufflebeam, 2001; Knox, 2002; Levin & McEwan, 2003). It is the stage that connects the *initial concerns*, the *intentions* and the *description of the programme* with the purely inquiring part of the evaluation, the conclusions and the communication of the findings. In essence, at this stage, worries, concerns, expectations and intentions are transfigured into specific questions and/ or assumptions, criteria and standards, which direct the collection, analysis and interpretation of the data. More specifically, this stage includes:

- The *questions*: The uniqueness of each evaluation object, the level at which it is implemented (Owen, 2007), the impact of the epistemological discussion and the differences among the participants with respect to their priorities and interests, make it extremely difficult, if not impossible, to give a common answer to the question *what is worth learning about the programme* (Patton, 1990). The questions reflect, on the one hand, the values and, on the other hand, the criteria on the basis of which value judgments will be formed (Scriven, 2005). *Initial concerns* and *intentions* seem to have a determining role in the final choice of the questions under study. However, the evaluator is the one who will ensure that their final number is realistic, applicable and functional, reflecting the positions of all participants. It will be of great help in this work if the evaluator ensures that the questions under examination, on the one hand, cover as many aspects of the programme as possible (framework, inputs, processes, results, feedback) and, on the other hand, provide information about the *values* that characterize the quality of the programme. Each question should be linked to the rest of the questions and all of them should be connected with the goal of the evaluation, so that the formation of value judgments is based on the overall perception of the quality of the programme.

- The *criteria*: The criteria are closely related to the *values* that specify the quality of the programme and the *questions* of the evaluation. The value or values that will play the most important role in determining the quality of the programme (Stake, 2004) and consequently the questions that will be asked for consideration (House & Howe, 1999; Sanders & Davidson, 2003; Fitzpatrick et al., 2004) determine the nature of the evaluation criteria. The particularly critical role of the criteria lies in the fact that the value judgement is based on them. Still, it should not be overlooked that it is almost impossible to develop commonly accepted criteria for any evaluation. However, if the values are taken into account, then, to some extent, it is possible to develop a set of criteria for each of the values that relate to the quality of the programme. The limited space we have at our disposal, nevertheless, does not allow us to expand further.

4.6 Answers

Given the fact that, during the evaluation, questions are asked, it is only natural that *answers* follow. This stage includes the study methodology (Patton, 1990; Clarke, 1999; House & Howe, 1999; Robson, 2000; Fitzpatrick et al., 2004) and, more specifically, the sources of information, methods, techniques and tools that can be used in the process of data collection, analysis and interpretation. Today, the complexity of the programmes, the diversity of the questions under consideration, as well as the need for quantitative and qualitative analyses are increasingly leading to evaluations, in which data are collected from various *defensible sources of information* (Joint Committee, 1994), and various methods of data collection, analysis and interpretation are used. When a combination of quantitative and qualitative methods is used in evaluation, it is called a *mixed method evaluation*. In this case, emphasis is placed on the best elements or advantages of each method. In other words, through triangulation (Clarke, 1999), and the complementary and dialectical nature of the quantitative and qualitative methods, it is attempted to maintain the fullest possible representation of the object of evaluation, the reduction of biases, the enhancement of the *values of the study*, the enrichment of the interpretations and judgements regarding the quality (Fitzpatrick et al., 2004), and the provision of more sufficient evidence for both the processes and the results (Patton, 1990).

4.7 Reporting, Presentation and Dissemination of Findings

In the sixth stage, the *report, presentation and dissemination* of the findings to the stakeholders take place (Joint Committee, 1994; Clarke, 1999; Knox, 2002). The disclosure of findings serves a dual purpose. On the one hand, the participants are supported in their personal interpretations regarding the quality of the programme. On the other hand, considering that the report is not a neutral text, the equal, fair, impartial, documented and independent presentation of the strengths and weaknesses of the programme protects the evaluator from the implication of manipulating those who are interested in the evaluation report. Perhaps, an additional way to avoid conflicts and misinterpretations that may be caused by what is stated in the evaluation report is to notify the report to the stakeholders, before its final presentation. This provides them with the opportunity to study the report and at the same time express their views on the text. In fact, it is not unlikely that their substantiated observations (Sanders, 1992) may prove to be particularly useful.

4.8 Meta-evaluation

The above stage is followed by *meta-evaluation* (Scriven, 1991; Uusikyla & Virtanen, 2000; Stufflebeam, 2000; Scriven, 2005; Owen, 2007), which attempts to ensure the quality of the evaluation. Its priorities and

methodology are closely related to the *intentions*, *questions* and *answers* of the evaluation. Meta-evaluation is an autonomous research process that is either provided for in the initial design of the evaluation or is carried out independently. It has a supervisory role, as it identifies both strengths and weaknesses. If the evaluators are responsible for meta-evaluation, then their work is significantly facilitated if, during the study, they maintain evidence that clearly demonstrates the adoption of specific procedures and the presence of achievements (Stufflebeam, 2002). In fact, when the evaluators' data are confirmed by others, then meta-evaluation has increased validity.

5. The Architecture of Evaluation and the Model of Holistic Programme Evaluation

The theory, in which goal-based evaluation, responsive evaluation, theory-based evaluation, utilization-focused evaluation, empowerment evaluation, and constructivism evaluation are grounded, helps the evaluator make many critical decisions. But if we wanted to group them, they would fall into three broad categories.

The first category includes decisions regarding the issues under consideration, the questions and the evaluation criteria. In fact, they are decisions that determine the content of the research part of the study and of course play a key role in formulating the value judgment. More specifically, the objectives of the programme provide the evaluation (goal-based evaluation model) with a set of positions, which enable the evaluator to examine primarily two key issues or *values* of the quality of the programme. First, the evaluator examines the *relevance* between the objectives of the programme and beneficiaries' needs-problems, i.e., the extent to which the objectives of the programme are appropriate and relevant to the needs of the beneficiaries. Second, the evaluator can examine the *effectiveness* of the programme, which is based primarily on the examination of the relationship between results and objectives. The more the results of the programme, which are mainly reflected in performance and achievements, are in line with the objectives of the intervention (intended intentions), the more effective the programme is.

Programme theory offers evaluation (theory-based evaluation model) a set of positions that address more issues, rather than only the objectives, based on which in-depth exploration of practices is possible. It is essentially the study of the *transformational* process that takes place between inputs and outputs. This examination allows two critical questions to be answered: *how* and *why* the programme succeeds or fails in its mission (Chen, 1990; Weiss, 2004; Robson, 2007). When, during the evaluation, the first question is examined, the focus is placed on exploring the *internal quality* of the programme; that is, the consistency between the theoretical structure of the programme (intended programme) and its implementation (implemented programme) is examined. When evaluation focuses on the why question, then it studies the *causative* sequence that mediates between intervention and achievements (Chen, 1990; Pawson & Tilley, 2000; Donaldson, 2007). Programme theory allows *monitoring* the mechanism that either produces or does not produce the intended results and effects of the programme.

When the evaluation seeks to respond to the participants' needs (responsive evaluation), then the potential *issues* for examination are indeed many. Problems, uncertainties and, overall, the participants' information needs play a key role in the evaluation (Stake, 2004). In practice, this suggests that responsive evaluation, after a detailed description of the programme, can focus on either the objectives, the programme supervision or anything that is of interest to the participants. Examples include the examination of operational problems or uncertainties,

the *efficiency* of the programme, its *utility*, as well as the *sustainability* of the benefits gained by the beneficiaries. The responsive approach provides the evaluator, the contracting and funding body of the study, as well as the participants, with the flexibility to explore those aspects of the programme on which they believe its quality is primarily based (Stake, 2000, 2004).

The *objectives* of the programme, its *theory* and the *needs of the participants* provide evaluation with the required theoretical framework, in which, on the one hand, the issues under consideration and consequently the questions and evaluation criteria, and, on the other hand, the formation of the evaluation judgement are grounded. However, it is possible that the issues, questions and criteria draw on other factors, such as the mission of the organization or the implementing body of the programme. In these cases, social theory usually provides the adequate theoretical framework. Of course, it is also possible that there may be cases of evaluations, in which the issues under consideration are not sufficiently supported by theory. In these cases, the evaluation is very likely to contribute to the production of new knowledge and further theoretical grounding of the programme.

At this point, one might wonder why it is necessary for the issues under consideration, the questions and the evaluation criteria to be grounded in theory. What does their theoretical grounding offer to the evaluation? The answer is twofold. First, the issues under consideration, the questions and the evaluation criteria are directly related to the methodology of the study and the formation of the value judgment, thus determining the overall course of the evaluation. Second, the existence of the theoretical framework gives the evaluation scientific rigour and trustworthiness, by removing the possibility of *accusations*, which in some cases are very easily, intentionally and even unsubstantially made against evaluation.

The second category includes decisions regarding the possibilities of utilizing the study. The theory in which utilization evaluation and empowerment evaluation are grounded has a lot to offer. Among other things, they provide positions that shape the framework of utilization, which is one of the most crucial dimensions of evaluation. It may be unnecessary, but it should be noted once again that evaluation is not an end in itself; it is implemented in order to serve a specific purpose (e.g., accountability, improvement, knowledge production). It is important for the participants, being aware of various aspects of the subject matter, its strengths and weaknesses, to be assisted in making *defensible* decisions. But the *learnings* that come from both the results and the procedures of the evaluation, and which should be disseminated to everyone, are also very useful to the participants. They enable them to either modify personal thoughts and behaviours, acquire additional knowledge and skills, change the organization of the programme, or change the culture of the implementing body (Patton, 2004).

Although evaluators do not have the ability to impose or lead to a de facto use of the study, they will significantly increase its opportunities for use if they identify (Sanders, 1992; Clarke, 1999): i) which participants are the intended users, ii) which findings and which processes are most likely to be utilized, that is, what the content of the intended utility is, iii) which factors positively affect and which negatively affect the intended utility, iv) how the intended utility relates to the other parts of the evaluation, such as the choice of the model and methodology. The crucial role of utility is increasingly highlighted by evaluators, who believe that evaluation can gain acceptance, respect, rigour and even legitimacy, when it is deemed useful and utilized by the participants. The third category includes decisions that focus on the communication-collaboration between the evaluator and participants. Communication and collaboration among the people involved is now an integral part of most evaluations. Epistemological and socio-political developments have brought about significant changes in programme evaluation during recent decades. Both the field of drawing issues, questions and criteria, which is based on the first three models, and the framework of utilization, which is reflected by the next two models, are grounded

in a communicative, participatory and negotiating universe, so that, in the one hand, the knowledge and skills of the people involved in the implementation of the evaluation are utilized, and, on the other hand, commitment, acceptance, respect and, above all, trust among the participants are ensured. It should not be overlooked that evaluation is to a large extent a process of interaction and communication between the evaluator and the participants. It is becoming more and more necessary to create conditions that allow the evaluator and the participants to express their positions and views freely and naturally, to *speak the same language* at least to a minimal extent, to participate and to understand the conditions, the difficulties and also the necessity of the evaluation project. Communication, participation and negotiation between those involved are now necessary, regardless of the theoretical framework, the methodology and the utility of the evaluation.

Some may think that communication and cooperation are about and, at the same time, limited only to the field of data collection. That is, the evaluators want to communicate with the participants only when they are at the stage of data collection. Although their cooperation at this level is a prerequisite for every evaluation, the evaluators should not treat the participants simply as a source of data. The more they engage the participants in the processes of design, implementation and even interpretation of the findings, the more they enhance the dynamics of their study. Communication and collaboration need evaluators' special attention, as critical issues are almost always raised, such as: i) which participants will be involved, ii) who will be in control of the study, iii) what the specific characteristics of participation are, and iv) which the political implications of the cooperation are.

Of course, in no case should the above three major categories of decisions be perceived independently of each other; on the contrary, they have many close and reciprocal links with each other. For example, if participants are interested in examining the utility of the programme, then the utilization of the study and communication with the participants have a different content from the case in which the evaluation focuses on the objectives of the programme.

But how can we connect the conceptual tools offered by the evaluation models, the values on the basis of which the quality of the programme is evaluated, and the suggested course of implementation of the evaluation? An attempt to answer this question is provided in the following part of the section, while Figure 1 presents the schematic illustration of this effort.

It is suggested that the basis of the whole evaluation structure is the creation of the conditions for negotiation, communication and participation among all stakeholders. Although the reasons that promote the participatory nature of evaluation, as well as the difficulties involved in this choice have already been mentioned, a critical dimension that should not escape our attention is that evaluation is an interactive and communicative process. It almost always invites, urges, demands or forces participants to invest a significant portion of their valuable time in evaluation processes. Participants must be treated with seriousness, respect, honesty and above all trust. Those participants who voluntarily wish to be engaged in the evaluation should understand that, throughout the study, the participation and utilization of their knowledge and experience are required. Then again, those who are forced to engage in the evaluation should realize that it is not an opportunistic, standardized and *inhuman* process, but a research activity, in which their positions and work are taken seriously into consideration. Perhaps, in this way, their trust is gained, as it is a critical parameter of the whole endeavour.

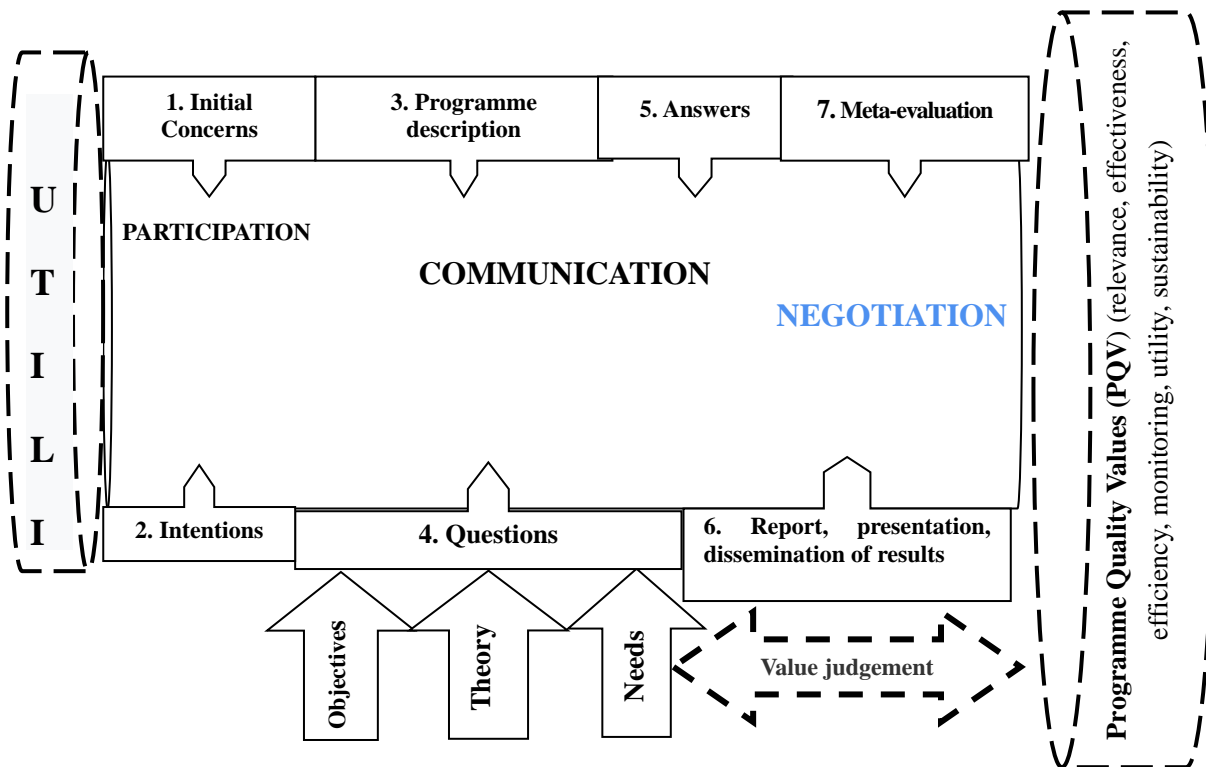


Figure 1 The Architecture of the Holistic Programme Evaluation Model

Moreover, utilization plays a particularly crucial role. Although the utility of the evaluation becomes feasible either when the evaluation is in progress or after its completion, the decision to utilize it should be included in the design decisions. The participants need to know that engaging in a process, which in addition to personal time and effort may be accompanied by *emotional turbulence*, will bring benefits to themselves, the beneficiaries, the programme, and, why not, the organization or the body implementing the programme. The participants, therefore, want the intentions of the people who are responsible for the programme to be clarified in a timely, honest and practical manner. If there is no consideration for the ways in which the evaluation may be useful, then a number of participants may consider that the study significantly loses its *ethical legitimacy*; if this happens, then the basis of the whole endeavour receives strong shocks with unpredictable consequences for the quality and effectiveness of the study itself.

The concepts of *objectives*, *programme theory*, and *response to the needs* of the participants play a particularly critical role. Their criticality lies in the following: first, the issues under consideration, the questions and the evaluation criteria are mainly derived from them. Second, the *values* that reflect the quality of the programme and on the basis of which the evaluation judgement is expressed, are grounded in them. This practically implies that the *intentions* of the participants and especially of the sponsors of the evaluation are a matter of particular importance. These determine the role that each of the *values* of relevance, monitoring, effectiveness, efficiency, utility, and sustainability will play, both in relation to the direction that the research itself will take and the formation of the judgement value.

6. Conclusions

The evaluator or anyone else undertaking an evaluation will be greatly assisted in their work if they clarify the levels at which they are asked to make decisions; i.e., on which elements they will focus their attention, whether these relate to the evaluation itself, the *object of the evaluation* or the roles that the evaluators and the participants will have. Perhaps, in this way, they can manage, on the one hand, to have alternative possibilities at their disposal, which allows them to be flexible at various levels, and, on the other, to attach coherence, consistency and grounded decisions to their study.

Without having any intention of providing *advisory instructions*, as the whole endeavour must be placed in its actual dimensions and perceived as a single version of the many that may appear in the field of programme evaluation, the following remarks are worth considering:

- The *initial concerns* (evaluability, policy, needs assessment, study management, ethics and biases) seem to be largely independent of the conceptual organizer (objectives, theory, needs) that determines the position of the study. Its subcomponents should be examined in any case.
- The *intentions* of the evaluation and especially the aims and values of the quality of the programme seem to significantly influence the choice of the conceptual organizer (objectives, theory, needs), on the basis of which the issues, questions and criteria of the evaluation are selected. On the contrary, the roles of both the evaluator and the participants do not seem to depend directly on the conceptual organizer that will direct the study.
- Although the *programme description* should be as complete as possible, the choice of the conceptual organizer also determines the importance that will be attached to the presentation of the sub-dimensions of the programme. For example, if theory is the basic conceptual organizer, then the description of the programme needs to consider more dimensions than in the case in which the objectives of the programme constitute the conceptual organizer.
- The relationship between *questions* and conceptual organizer is extremely close.
- The answers, which essentially refer to the methodology of the evaluation, do not seem to be particularly dependent on the conceptual organizer. All scholars in the field unequivocally suggest that the methodology to be followed should be the one that best serves the goal of the study. The answers are therefore closely related to the intentions.
- The report, presentation and dissemination of the findings, as well as meta-evaluation, do not seem to have a strong connection with the selection of the conceptual organizer. Communicating the findings and ensuring the quality of the study are closely related to the intentions.
- In every evaluation, there should be an intention for utilization. Without it, evaluation loses a key axis of its legitimacy.
- Participation, communication and negotiation among stakeholders should encompass as many decisions as possible at each stage of the evaluation.
- Usually, the conceptual organizer, and consequently the values that relate to the quality of the programme and guide the evaluation, is not necessarily just one. Lately, there is a growing tendency for evaluation to examine issues and ask questions that are grounded in more than one conceptual organizers and values that characterize the quality of the programme. For example, it is increasingly common in today's evaluations to consider both the effectiveness and the utility of the programme.

- The more complete theoretical grounding of the evaluation and the programme contributes, among other things, to the conceptual clarification of the questions and especially of the evaluation criteria and standards, a fact that significantly facilitates the formation of the value judgment.
- The value judgement is grounded in the relationships that develop among the conceptual organizer of the study (objectives, theory, needs), the values on the basis of which the quality of the programme is understood (relevance, effectiveness, monitoring, efficiency, sustainability and utility), the issues under examination, the questions and the evaluation criteria and of course the resulting answers.
- A comprehensive synthesis has a flexible character; it is more of an open framework of actions and options than a set of strictly defined rules and instructions for use.

The nature and technical specifications of the programme, the policy priorities, the intentions of the participants and funding bodies, the mission of the implementing body, the questions under consideration, as well as the knowledge, skills and experience of the evaluator or the evaluation team make each evaluation unique. However, we hope that the theoretical framework we used, the architecture of the evaluation we presented and the evaluation proposal we synthesized are useful in making decisions regarding: i) the evaluation itself, ii) the object of evaluation, iii) the methodology, and iv) the roles that the evaluator will assign to themselves and the participants. We have tried to outline the gestalt of programme evaluation without neglecting that our effort is subject to significant limitations. Indicatively, it should be mentioned that we have used only a part of the programme evaluation theory. Above all, however, we hope that we have contributed to a more comprehensive understanding of the evaluation endeavour.

References

- Abma T. (2004). "Responsive evaluation: The meaning and special contribution to public administration", *Public Administration*, Vol. 82, No. 4, pp. 993–1012.
- Alkin M. and Christie C. (2004). "An evaluation theory tree", in: M. Alkin (Ed.), *Evaluation roots: Tracing Theorists' Views and Influences*, Sage, pp. 12–66.
- Beywl W. (2003). *Selected Comments to the Standards for Evaluation of the German Evaluation Society*, German Evaluation Society.
- Bonnio J. and Vial M. (2007). *Evaluation Models: Fundamental Texts With Interpretive Comments*, I. Papadimitriou, Z. Polymeropoulou, G. Stergiou, Trans., Athens: Metaichmio.
- Chen H. (1990). *Theory-driven Evaluation*, Sage.
- Chen H. (2004). "The roots of theory-driven evaluation: Current views and origins", in: M. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences*, Sage, pp. 132–152.
- Clarke A. (1999). *Evaluation Research*, London: Sage.
- Cousins J. (2003). "Utilization effects of participatory evaluation", in: T. Kellaghan & D. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*, Kluwer Academic Publishers, pp. 245–268.
- Donaldson S. (2007). *Program Theory-driven Evaluation science: Strategies and Applications*, Lawrence Erlbaum Associates.
- Fetterman D. (1995). "A response to Stufflebeam's review of empowerment evaluation", *Evaluation Practice*, Vol. 16, No. 2, pp. 179–199.
- Fetterman D. (2000). "Steps of empowerment evaluation: From California to Cape Town", in: D. Stufflebeam, G. Madaus, T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Kluwer Academic Publishers, pp. 395–408.
- Fetterman D. (2005). *Empowerment Evaluation Principles in Practice: Assessing Levels of Commitment*, Guilford Publications.
- Fitzpatrick J., Sanders J. and Worthen B. (2004). *Program Evaluation: Alternative Approaches and Practical Guidelines*, Pearson.
- Habermas J. (1984). *The Theory of Communication Action: Reason and the Rationalization of Society*, Heinemann.
- Habermas J. (1990). *Cognitive Theory and Social Critique Texts*, A. Economou Trans., Athens: Plethron.
- Habermas J. (1997). *The Ethics of Communication*, K. Kavoulakos Trans., Athens: Enallaktikes Ekdoseis.

- Habermas J. (1997). "Positivism, pragmatism, historicism", in: G. Kouzelis (Ed.), *Epistemology: Texts*, L. Anagnostou Trans., Athens: Nisos, pp. 480–484
- Hambleton R. (2000). "Criterion-referenced measurement", in: J. Keeves (Eds.), *Educational Research, Methodology, and Measurement: An International Handbook*, Pergamon, pp. 719–725.
- House E. and Howe K. (1999). *Values in Evaluation and Social Research*, Sage.
- Huebner A. and Betts S. (1999). "Examining fourth generation evaluation: Application to positive youth development", *Evaluation*, Vol.5, No. 3, pp. 340–358.
- King J. (2003). "Evaluating educational programs and projects in the USA", in: T. Kellaghan & D. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*, Kluwer Academic Publishers, pp.721–732.
- Knaap P. (2006). "Responsive evaluation and performance management: Overcoming the downsides of policy objectives and performance indicators", *Evaluation*, Vol. 12, No. 3, pp. 278–293.
- Knox A. (2002). *Evaluation for Continuing Education: A Comprehensive Guide to Success*, Jossey-Bass.
- Levin H. and MacEwan P. (2003). "Cost-effectiveness analysis as an evaluation tool", in: T. Kellaghan & D. Stufflebeam (Ed.), *International Handbook of Educational Evaluation*, Kluwer Academic Publishers, pp. 125–152.
- Lincoln Y. (2003). "Constructivist knowing, participatory ethics and responsive evaluation: A model for the 21st Century", in: T. Kellaghan & D. Stufflebeam (Ed.), *International Handbook of Educational Evaluation*, Kluwer Academic Publishers, pp. 63–78.
- Lincoln A. and Guba E. (2004). "The roots of fourth generation evaluation: theoretical and methodological origins", in: M. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences*, Sage, pp. 225–242.
- Marby L. (2003). "In living color: Qualitative methods in educational evaluation", in: T. Kellaghan & D. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*, Kluwer Academic Publishers, pp. 167–188.
- Madaus G. and Stufflebeam D. (2000). "Program evaluation: A historical overview", in: D. Stufflebeam, G. Madaus, T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Kluwer Academic Publishers, pp. 3–18.
- Mcguire M. (2002). *Canadian Evaluation Society Project in Support of Advocacy and Professional Development: Literature Review*, Zorzi & Associates.
- Mertens D. (2009). *Research and Evaluation in Education and Psychology*, S. Kyranakis, M. Mavraki, & P. Bithara Trans, Athens: Metaichmio.
- Owen J. (2007). *Program Evaluation: Forms and Approaches*, The Guilford Press.
- Patton M. (1990). *Qualitative Evaluation and Research Methods*, Sage.
- Patton M. (2000). "Utilization-focused evaluation", in: D. Stufflebeam, G. Madaus, T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Kluwer Academic Publishers, pp. 425–438.
- Patton M. (2003). "Qualitative evaluation checklist", available online at: <http://www.wmich.edu/evalctr/checklists>.
- Patton M. (2004). "The roots of utilization-focused evaluation", in: M. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences*, Sage, pp. 276–292.
- Pawson R. and Tilley N. (2000). *Realistic Evaluation*, Sage.
- Preskill H. (2003). "The evaluation profession as a sustainable learning community", in: T. Kellaghan & D. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*, Kluwer Academic Publishers, pp. 361–372.
- Preskill S. and Preskill H. (1997). "Meeting the postmodern challenge: Pragmatism and evaluative inquiry for organizational learning", in: R. Stake, L. Mabry (Eds.), *Advances in Program Evaluation: Evaluation and the Postmodern Dilemma*, Jai press Inc., pp. 155–170.
- Robson C. (2000). *Small-scale Evaluation: Principles and Practice*, Sage.
- Robson C. (2007). *Real-world Research: A Resource for Social Scientists and Professional Researchers*, V. Dalakou & K. Vasilikou Trans., Athens: Gutenberg.
- Rogers P. (2000). "Program theory: Not whether programs work, but how they work", in: D. Stufflebeam, G. Madaus, T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Kluwer Academic Publishers, pp. 209–232.
- Sanders J. (1992). *Evaluating School Programs*, Corwin Press.
- Sanders J., Davidson J. (2003). "A model for school evaluation", in: T. Kellaghan & D. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*, Kluwer Academic Publishers, pp. 807–826.
- Scriven M. (1991). *Evaluation Thesaurus*, Sage.

- Scriven M. (2000). "Evaluation Ideologies", in: D. Stufflebeam, G. Madaus, T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Kluwer Academic Publishers, pp. 249–278.
- Scriven M. (2005). "Key evaluation checklist (KEC) intended for use in designing, evaluating, and writing evaluation reports on programs, plans, and policies; and for evaluating evaluations of them", available online at: <http://www.wmich.edu/evalctr/checklists>.
- Shaw I. (1999). *Qualitative Evaluation*, London.
- Stake R. (1973). "Program evaluation particularly responsive evaluation", in: *Conference on "New Trends in Evaluation"*, Goteborg, Sweden.
- Stake R. (2000). "Program evaluation, particularly responsive evaluation", in: D. Stufflebeam, G. Madaus, T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Kluwer Academic Publishers, pp. 343–362.
- Stake R. (2004). *Standard-based and Responsive Evaluation*, Sage.
- Stame N. (2004). "Theory-based evaluation and types of complexity", *Evaluation*, Vol.10, No. 1, pp. 58–76.
- Stufflebeam D (2000). "Foundational models for 21st century program evaluation", in: D. Stufflebeam, G. Madaus, T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Kluwer Academic Publishers, pp. 33–84.
- Stufflebeam D. (2000). "The CIPP Model for evaluation", in: D. Stufflebeam, G. Madaus, T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Kluwer Academic Publishers, pp. 279–318.
- Stufflebeam D. (2000). "The methodology of metaevaluation", in: D. Stufflebeam, G. Madaus, T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Kluwer Academic Publishers, pp. 457–472.
- Stufflebeam D. (2001). *Evaluation Models*, Jossey-Bass.
- Stufflebeam D. (2002). "CIPP evaluation model checklist: A tool for applying the fifth installment of the CIPP model to assess long-term enterprises intended for use by evaluators and evaluation clients/stakeholders", *Evaluation Checklists Project*, available online at: <http://www.wmich.edu/evalctr/checklists>.
- The Joint Committee on Standards for Educational Evaluation (1994). *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs* (2nd ed.), Sage.
- Trevisan M. and Huang Y. (2003). "Evaluability assessment: A primer", *Practical Assessment, Research & Evaluation*, Vol. 8, No. 20, pp. 1–5.
- Uusikyla P. and Virtanen P. (2000). "Meta-evaluation as a tool for learning: A case study of the European structural fund. Evaluations in Finland", *Evaluation*, Vol. 6, No. 1, pp. 50–65.
- Weiss C. (2004). "Rooting for evaluation: A cliff notes version of my work", in: M. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences*, Sage, pp. 153–168.
- Widmer T., Landert C. and Bachmann N. (2000). *Evaluation Standards*, Swiss Evaluation Society (several standards).
- Yin R. (1994). *Case Study Research*, Sage.