Problemy Zarządzania – Management Issues, vol. 16, no. 6(80) part 2: 23–38
ISSN 1644-9584, © Wydział Zarządzania UW
DOI 10.7172/1644-9584.80.2

Evaluation as Reflective Practice

Paulina Kubera*

Submitted: 07.09.17 | Accepted: 26.10.18

Reflective practice has become an influential concept in the evaluation field. A greater use of reflective practice is advocated in reference to both evaluators' own professional development, and as a means to enhance dialogue, stakeholders' involvement and organisational learning in the evaluation process.
The aim of the paper is to examine the evaluation endeavour from the latter perspective, i.e. to present evaluation approaches which offer the opportunity for collaborative reflective practice. To this end, evaluation as reflective practice is discussed at three levels: (1) the organisational level – the model of single, double and triple-loop learning is discussed in reference to formative, summative and developmental evaluation, (2) the evaluator's level – different roles performed by the evaluator are considered from the point of view of promoting collaborative reflective practice, and (3) the broader socio-political level, in relation to the concept of civil society, as evaluation can contribute not only to a greater rigour and effectiveness of public spending but to social empowerment, to appreciating diversity or building trust (improving democratic policy-making).

Keywords: evaluation, reflective practice, organisational learning, evaluator's roles, civil society.

Ewaluacja jako refleksyjna praktyka

Nadesłany: 07.09.17 | Zaakceptowany do druku: 26.10.18

Zagadnienia refleksyjnej praktyki stanowią ważny obszar zainteresowań w dziedzinie ewaluacji. Zaleca się jej szersze wykorzystanie zarówno w odniesieniu do samodoskonalenia zawodowego ewaluatorów, jaki i jako środek wzmocnienia dialogu, zaangażowania interesariuszy i uczenia się organizacyjnego w procesie ewaluacji.
Celem artykułu jest ukazanie procesu ewaluacji właśnie z tej drugiej perspektywy, czyli zaprezentowanie takich podejść stosowanych w ewaluacji, które sprzyjają praktyce refleksyjnej opartej na współpracy. Ewaluacja jako praktyka refleksyjna jest rozważana na trzech poziomach: (1) organizacji – w tym celu model uczenia się w pojedynczej, podwójnej i potrójnej pętli odniesiono do trzech podstawowych typów ewaluacji: formatywnej, sumatywnej i rozwojowej, (2) ewaluatora – różne role odgrywane przez ewaluatora są rozpatrywane z punktu widzenia promowania refleksyjnej praktyki wśród głównych odbiorców ewaluacji (interesariuszy), oraz (3) szerszego socjopolitycznego poziomu – w odniesieniu do społeczeństwa obywatelskiego, ponieważ ewaluacja może przyczyniać się nie tylko do większej dyscypliny i efektywności

* **Paulina Kubera** – PhD, Poznań University of Technology, Faculty of Management Engineering, Chair of Entrepreneurship and Business Communication.

Correspondence address: Poznań University of Technology, Faculty of Management Engineering, 11 Strzelecka Street, 60-965 Poznań; e-mail: Paulina.Kubera@put.poznan.pl.

Ministry of Science and Higher Education
Republic of Poland

The creation of the English-language version of these publications is financed in the framework of contract No. 607/P-DUN/2018 by the Ministry of Science and Higher Education committed to activities aimed at the promotion of education.

wydatków publicznych, lecz także do wzmocnienia zaangażowania społecznego, docenienia różnorodności czy budowania zaufania (wsparcia demokratycznego procesu tworzenia polityki).

Słowa kluczowe: ewaluacja, praktyka refleksyjna, organizacyjne uczenie się, rola ewaluatora, społeczeństwo obywatelskie.

JEL: 038

1. Introduction

Given the complex issues with which decision makers in both private and public sector are confronted today, calling on professional advice has become almost indispensable, as it gives credibility to the solution of a problem and helps to secure social acceptance (Ulrich, 2008). Against this background, the concept of 'reflective practice' has attracted a lot of attention. In most general terms, it involves the integration of theory and practice (Thomson, 2000), understanding how knowledge is integrated into practice in a meaningful way, and how knowledge is generated from practice. However, there is a rich literature on reflective practice that proposes various interpretations of the concept and its application in different settings: starting from the perspective of an individual reflective practitioner (e.g. Schön, 1983: Reflection-in-Action and Reflection-on-Action), towards the organisational perspective (e.g. Argyris & Schön, 1978: Single and double-loop learning, Vince 2002: Organizing Reflection) and, more recently, a broader socio-political level (e.g. Ulrich, 2000: Reflective Practice in Civil Society, Thomson and Pascal 2012).

Reflective practice is also central to the evaluation process, which can be defined as 'a systematic process for gathering and interpreting information intended to answer questions about a program' (McDavid, Huse & Hawthorn, 2013, p. 491), a policy or an organization. The question, if well-crafted, can help build knowledge and critical thinking about specific public actions and improve ongoing or future public interventions.

Madus, Stufflebean and Scriven (2000) have distinguished seven periods in the history of programme evaluation dating as far back as the nineteenth century. However, according to Cruise (1999), it was not until the 1960s that the social and political climate allowed evaluation research to emerge as a formal endeavour, in particular in the field of education, health and antipoverty programmes. In the United States, this was a time when vast amounts of money were spent on social programmes implemented under the banners of 'War on Poverty' and 'the Great Society'. The tension between the magnitude and complexity of social needs to be addressed by public measures and limited public resources was so great that social programme spending and its benefits became subject of growing attention (Rossi, Freeman & Lipsey, 2004).

 DOI 10.7172/1644-9584.80.2

Evaluation can serve two main purposes: (1) accountability purposes – evaluation can provide a basis for the justification of actions taken, including the provision of information to the public; it implies an external rather than internal orientation – evaluation for others rather than evaluation for ourselves (decision makers), and (2) learning purposes – evaluation can improve existing and future policies, programmes or projects through feedback from lessons learned; it implies an internal orientation, i.e. evaluation for ourselves rather than for others.

Currently, a greater use of reflective practice is advocated in reference not only to the evaluators' own professional development, but as a means to enhance dialogue, stakeholders' involvement and organisational learning in the evaluation process. Thus, the aim of the paper is to examine, on the basis of a rich body of literature in the field of reflective practice and evaluation, evaluation approaches that offer the opportunity for collaborative reflective practice.

The article is structured in three sections discussing evaluation as a reflective practice at three different levels: (1) the organisational level – the model of single, double and triple-loop learning is discussed in reference to formative, summative and developmental evaluation, (2) the evaluator's level – different roles performed by an evaluator are considered from the point of view of promoting collaborative reflective practice, and (3) the broader socio-political level, in relation to the concept of civil society, as evaluation can contribute not only to the rigour and effectiveness of public spending, but also to social empowerment, appreciating diversity or building trust (improving democratic policy-making).

2. Organisational Perspective. Formative, Summative and Developmental Evaluation in the Context of Single, Double and Triple-loop Learning

As learning is a key term in the context of reflective practice, the focus of the paper is on evaluation as a tool to enhance organisational learning. Depending on the type of evaluation we can identify single, double or triple-loop learning model in the evaluation process.

In literature, learning is often described as a dichotomy. According to Argyris and Schön (1978, p. 2) learning involves the detection and correction of an error. 'Whenever an error is detected and corrected without questioning or altering the underlying values of the system' it is referred to as 'single-loop learning'. 'When mismatches are corrected by first examining and altering the governing variables, and then the actions', it is referred to as 'double-loop learning'. In a similar vein, Fiol and Lyles (1985, p. 807–808) distinguish low-level and higher-level learning. The former occurs 'within a given organizational structure, a given set of rules', while

the latter 'aims at adjusting overall rules and norms rather than specific activities or behaviours'. Only later was a further category of organisational learning conceived, referred to most often as 'triple-loop learning' (Tosey, Visser & Sounders, 2011).

One can identify these models in different types of evaluation, where formative evaluation represents single-loop learning, summative evaluation – double-loop learning and relatively recent phenomenon, developmental evaluation – triple-loop learning. Similarly to the advancements in the theory of organisational learning that go beyond the traditional dichotomy, a third type of evaluation has been proposed to answer the problem of how to deal with complex and dynamic realities.

Scriven (1967) was the first to suggest a distinction between formative and summative evaluation. Since then the formative/summative dichotomy has been considered fundamental for program evaluation and widely used in evaluation research and literature (Shadish, Cook & Leviton, 1991). However, not without criticism (see: Chen, 1996; Patton, 1996). Scriven (1991) formative evaluation is defined as '…evaluation designed, done, and intended to support the process of improvement, and normally commissioned or done by, and delivered to someone who can make improvements' (p. 20). Seen from this perspective, formative evaluation is a specific process evaluation, performed while a programme (or other public initiative) is still in progress, in order to work out implementation difficulties. Summative evaluation, in turn, serves the function of merit assessment, in particular to determine the overall effectiveness of a given public intervention. Using the analogy referred to by Robert Stake (cited in Scriven, 1991), the act of tasting the soup by the cook represents formative evaluation, as it occurs during the cooking process with a view to improving the soup. The act of tasting the soup by the guest represents summative evaluation. It provides a conclusive opinion about the soup (outcome assessment) and suggests that summative evaluation is performed *ex post*, i.e. after the completion of a given programme (or another public initiative). Evaluation findings can be used when deciding whether to continue, alter or discontinue a given intervention. However, it should be noted that both formative and summative evaluation can deal with processes and outcomes, and they can be conducted at various stages of programme implementation (e.g. mid-term evaluation can be summative) (Patton, 2012a; Chen, 1996). Moreover, it is argued that the formative/summative dichotomy does not encompass all evaluation designs (Patton 1996). Hence, the third category – developmental evaluation – has been introduced to complement the evaluation typology. Nevertheless, this distinction remains common and it is used to define evaluation intended to improve performance (formative evaluation) and evaluation intended to provide information about the worth of the programme (summative evaluation). It is the purpose of evaluation that distinguishes formative from summative evaluation, although in practice,

 DOI 10.7172/1644-9584.80.2

formative evaluation is often performed during the implementation phase of a project or a programme, and summative evaluation is conducted at the end of an intervention or an intervention phase (see: Glossary of Key Terms in Evaluation and Results Based Management 2010). A debatable issue is the area of overlap between formative evaluation and monitoring. The demarcation line lies principally in the regularity of the process and the depth of the analysis. Monitoring is a continuous process that entails systematic and routine collection of data on specific indicators with a view to informing the management and the main stakeholders about the progress of a specific intervention, while evaluation tends to be periodic and involve a deeper level of analysis (Markiewicz & Patrick, 2016).

From this perspective, it can be argued that formative evaluation represents single-loop learning while, in essence, it entails action-oriented changes made within a given set of assumptions (intervention logic). In simple-loop learning, individuals correct their actions taking into account the identified difference between anticipated and actual outcomes, and they introduce changes that contribute to the attainment of the desired outcomes (Patton, 2017). Its focus is on implementation difficulties. Summative evaluation, in turn, is more about testing the assumptions that are the foundation of the adopted intervention logic. Intervention logic is a path which leads from inputs to immediate effects (outputs) and, subsequently, to intermediate and long-term effects (impacts), identifying all along the way any important preconditions and providing justifications supporting the links between program inputs (activities) and the expected effects of different levels (see: Kubera, 2017). This is particularly true for theory-based evaluation (TBE), which is often described as the 'opening of a black box' of public intervention (Stame, 2004). (This type of evaluation has emerged as an answer to limitations of the previously dominant approach in the evaluation process, namely the method-oriented approach, which focuses on measuring the magnitude of effects rather than on defining why and how an intervention works). This simply means going beyond the single-loop learning, as the process entails gaining deeper understanding of how the system in which an intervention takes place functions, and which aspects of the system's dynamics led to the problem in the first place.

However, as public interventions are implemented in an increasingly complex and uncertain environment, a new approach, i.e. developmental evaluation has been developed in order to deal with the unexpected. Developmental evaluation is underpinned by theoretical premises that are different from those of the traditional approaches mentioned above; it departs from the classic modelling approach (Patton, 2012a). In formative and summative evaluation, efforts are made to find a model that works in a certain context, and to disseminate it. It is about prediction and control gained through the specification and measurement of the fidelity of implementation and the attainment of predetermined outcomes. It is suitable

when the problem is well recognised, there are clear boundaries and limited alternative solutions, one of which is likely to be optimal (Patton, 2017; Gamble, 2008). Developmental evaluation, however, is context-responsive, accepts and adopts to emergent and unexpected realities, is driven not by a model, but by principles. Its primary focus is on exploring possibilities and experimenting, rather than arriving at a fixed intervention (Dicks & Saunders, 2014). As Patton (2017), the founding father of developmental evaluation states, it 'tracks and attempts to make sense of what emerges under conditions of complexity, documenting and interpreting the dynamics, interactions, and interdependencies that occur as innovations and systems-change processes unfold. It can be argued that in developmental evaluation triple-loop learning is embedded' (p. 9).

Although inspired by Argyris and Schön, the concept of triple-loop learning does not appear explicitly in their published work. Tosey, Visser and Sounders (2011) identified three different conceptualisations of triple-loop learning which are presented in the literature. One of them views triple-loop learning as a higher or deeper level compared to single and double-loop learning. Following the logic of Argyris and Schön, if double-loop learning is about changing the governing variables, triple-loop learning should involve changes in what underpins those governing variables. Isaacs (1993, p. 38–39), for instance, writes about 'a dialogue', which 'involves learning about the context and the nature of the process by which people form their paradigms'. It has a potential to transform 'the grounds out of which all our thinking and acting emerges. An alternative conceptualisation of triple-loop learning is equivalent to Argyris and Schön's concept of 'deutero-learning', where emphasis is placed on change in the process of learning rather than in the purposes, principles or paradigms underpinning operations and the strategy. It is described as 'learn[ing] how to carry out single- and double-loop learning" (Argyris & Schön, 1978, p. 27), not a level further in the hierarchy of learning. Yuthas, Dillard and Rogers (2004), for instance, define this concept as 'continual reflection on the learning process, the contexts within which learning occurs, and the assumptions and values motivating the learning and influencing its outcomes which, in most general terms, can mean learning how to learn' (p. 239). Finally, the third conceptualisation of triple-loop learning has been inspired by Bateson's Learning III, which is defined as 'a corrective change in the system of sets of alternatives from which choice is made' (Bateson, 1973, p. 250). It differs from the two conceptualisations mentioned earlier since this type of learning is not to be achieved through conscious, instrumental means. In this perspective, the unconscious and aesthetic play an important role. Moreover, learning levels do not represent a hierarchy, but the idea of recursion. There are feedback loops between levels. Each subsequent level goes beyond and includes the previous one, but it also entails higher risk. Thus, efforts

 DOI 10.7172/1644-9584.80.2

made by managers to control and influence Learning III may result in undesired consequences (see: Tosey, Visser & Sounders, 2011).

However, one more conceptualisation of triple-loop learning is worth mentioning, developed by Flood and Romm (1996), who introduced their triple-loop learning by addressing the political dimension behind learning. From this point of view, the evaluation process represents triple-loop learning when it entails the reflection on relations of power that determine what particular purposes are privileged and valued.

3. An Evaluator as a Reflective Practitioner. Self-oriented and Collaborative Approaches

The quality of evaluation is, notably, dependent on the expertise of evaluators. Selecting the right evaluators with evaluation and subject-matter expertise is crucial for enhancing the use of evaluation findings and the evaluation process itself (de Laat, Williams, 2014).

Stevahn, King, Ghere and Minnema (2005) identified six essential competencies for programme evaluators. One of them is cultivating reflective practice, which implies self-awareness, reflecting on practice, pursuing professional development and building professional relationships. However, evaluators use reflective practice not only for personal development and improving their own practice, but also to improve many facets of the evaluation endeavour, seeing reflective practice as a process of communication with stakeholders and organisational learning, which enhances the use of evaluation (Preskill, Zuckerman & Matthews, 2003; Jones & Stubbe,, 2004; Patton, 2012b). Thus, evaluators can be engaged in a reflective practice, which is self-oriented, when it entails a critical and deliberate inquiry into their own practice, their biases, theoretical predispositions and preferences (Schwandt, 2001), as well as a collaborative reflective practice, when the inquiry pertains to a community of practice involving critical thinking with others to understand themselves, others, and meanings that are jointly constructed (Preskill & Torres, 1999).

As Patton (2012b, p. 55) rightly points out: ‘Reflectivity reminds the evaluator to be attentive to and conscious of the cultural, political, social, economic, linguistic, and ideological origins of one’s own perspective as well as perspectives and voices of those you gather data from and those to whom you present findings. The question is, however, what is the extent and the manner in which evaluators actually engage in reflective practice. This, in turn, depends on the role assumed by evaluators while performing their tasks. The traditional purpose of an evaluator is to make judgement about the merit and worth of an evaluand (e.g. a programme), assess its efficiency and make recommendations for improvement. However, there is evidence that evaluation professionals are more concerned with taking more engaging and encompassing view of the evaluation process to enhance

the impact of evaluation: promote the use of evaluation, organisational learning and buy-in to the evaluation process (Smith, 2014).

Luo (2010) conducted a comparative analysis of positions towards evaluator's role represented by five theorists. All theorists under investigation, i.e. Scriven, Campbell, Stake, Weiss and Rossi, are regarded as most influential in the evaluation field and recognised in Alkin's famous evaluation theory tree (2013), which serves as a framework of evaluation prescriptive theories. The evaluation tree classifies a vast array of evaluation approaches and highlights features that set them apart. In addition to the above five, one more evaluation theorist is worth mentioning, namely M. Patton. Now, the list includes representatives of all three branches of the tree, that is, methods, values and use (evaluation theorists have been classified by their primary focus on one of these three essential elements of evaluation).

Sriven and Stake are representatives of the 'value branch', as they consider the process of placing value on the evaluation as the most essential component of evaluators' work. In Scrivens's theory (1986), the evaluator is a 'judge'. His or her role is to determine the criteria of merit (based on the needs analysis), set comparative evaluation standards, assess programme performance and offer a final evaluative judgement supported by empirical facts and probative reasoning. In order to avoid bias, Scriven advocates 'goal-free evaluation', i.e. ignoring predetermined goals and objectives of a given programme and investigating all actual consequences without taking into account the programme's intentions. This means that, in Scriven's theory, the evaluator must be highly experienced, as there is a great risk of personal bias. It is the evaluator who decides whose needs should be considered first and which merits are of higher priority; thereby, the evaluator is given more authority than the programme stakeholders. Despite representing the same branch, Stake believes that the evaluator should play the role of a 'facilitator' in the process of evaluation through assisting different stakeholders in 'discover[ing] ideas, answers, and solutions within their own mind' (Stake & Trumbull, 1982, p. 1). As evaluators work with individuals who know a lot about the matter that is to be evaluated, their job is to help them add to that understanding. He points out that there are different constituencies, different stakeholders with different expectations and different values, and the evaluator must recognise them. The evaluator's main responsibility, in Stake's view, is to illustrate those differences, to portray the complexity, however, without pressing for a consensus.

Cambell, Rossi and Weiss are representatives of the 'method' branch, as they consider developing models for evaluation practice as a key component of the evaluator's work. In Cambell's theory (1984), the evaluator is a 'methodologist', who employs scientific methodologies, mostly derivatives of randomised control trials, to eliminate biases in evaluative research and track a causal interference about a given programme and its hypothesised

 DOI 10.7172/1644-9584.80.2

effects. In Campbell's view, the evaluator is less concerned about assessing the value of the evaluated programme or advocating specific uses of this evaluation. His or her main responsibility concerns methods of measuring the programme's outcome (Shadish, Cook & Leviton, 1991). In doing so, the evaluator is distanced from the programme's stakeholders and works independently to establish facts. In this perspective, the evaluator's role does not differ significantly from the role of other social science researchers (Luo, 2010). In the early theory of Weiss, one can identify the same great emphasis on experimental methodology employed to assess the outcome of the programme. However, due to political and organisational problems encountered in her evaluation practice, Weiss adopted a more realistic approach to evaluation. As she points out: 'What looks elementary in theory turns out in practice to be a demanding enterprise. Programmes are nowhere near as neat and accommodating as the evaluator expects. Nor outside circumstances as passive and unimportant as he might like. Whole platoons of unexpected problems spring up' (Weiss, 1972, p. 25 as cited in Shadish, Cook & Leviton, 1991, p.183). Consequently, she recognised the need to understand the programme under evaluation and to understand how it works, i.e. the mechanisms of change that differ in individual fields. Therefore, evaluators 'need to become savvy about the mechanisms of change in different fields. Such knowledge can come from specializing in a single field long enough to gain profound knowledge of the ways in which interventions produce effects and the conditions under which efforts are most likely to appear. If evaluators move from field to field, they need to be sure that they gain as much knowledge as possible, either from other team members, consultants, reading, or programme practitioners. Evaluators cannot rely solely on their expertise in research methodology any longer' (Weiss, 1998; cited in: Horsch, 1998, p. 5). Weiss also makes a case to increase the impact of evaluation on the decision-making process (Weiss, 1998). In her view, the evaluator is more like an educator (Luo, 2010), as she argues that evaluation 'should be continuing education for programme managers, planners and policy makers". (Weiss, 1988, p. 18). Rossi, in turn, believes that the role of the evaluator should vary depending on the stage of evaluation and the type of social programme (Luo 2010). This approach seems to be in line with Patton's stance.

Patton is a representative of the 'use' branch, the third major branch, most concerned with the use of evaluation itself and the information generated through evaluation. Patton (2007) suggests a situational view on evaluators' roles. He argues: 'the roles played by an evaluator in any given situation will depend on the evaluation's purpose, the unique constellation of the conditions with which the evaluator is faced, and the evaluators' own personal knowledge, skills, style, values and ethics' (p. 213). What seems to be most relevant from the point of view of the aim of the paper is the role that the evaluator takes in developmental evaluation, whose

founding father is Patton, as this type of evaluation involves a process of co-creation. Contrary to traditional evaluation approaches, in developmental evaluation, evaluator forms part of the intervention team. His or her primary function is to elucidate the innovation and adaptation process, to track their implications and results, and to facilitate the on-going, real-time, data-based decision making. In developmental evaluation, evaluators are described as being 'active-reactive-interactive-adaptive". Patton (2011) points out that the evaluator is 'one among many at the design negotiating table' (p. 314).

If we assume that dialogue, stakeholders' involvement and organisational learning are elements of collaborative reflective practice (Smith, 2014), the evaluation approach advocated by Patton, Stake or Weiss seems to be more conducive to reflective practice, whereas the approach suggested by Scriven and Cambell remains on the other end of the spectrum. In the first case, stakeholders' involvement in the evaluation process is suggested and a 'learning'-based role of evaluators, who encourage evaluative dialogue rather than pronounce unilateral judgements about the quality process. In the latter case, evaluators are distanced from stakeholders, work independently, are seen as 'judges of right and wrong' (Scriven) or scientific methodologists (Cambell).

4. Civil Society Perspective. Using Systems Concepts (Critical Systems Heuristics) in Evaluation

Ulrich (2000) makes an interesting point in his article entitled 'Reflective Practice and Civil Society'. He argues that the concept of reflective practice and the concept of civil society are inseparable. He advocates for a broader understanding of 'reflective practice', which goes beyond practitioner's personal experience in applying knowledge to practice, his or her intellectual, methodological, emotional and intuitive skills. The concept should also include societal, ethical and political considerations, as some pressing questions need to be addressed, *inter alia*, how to expect that professionals (e.g. evaluators) justify their finding and conclusions when there are controversial views on what constitutes relevant facts and values, what is meant by solving a problem rationally when parties concerned have different rationalities, and what is meant by formulating recommendations for improvement when improvement means something different for different people (Ulrich, 2008).

It is no longer legitimised that professionals – owing to their expertise in a field – are granted authority over ordinary citizens, considered incompetent in matters that affect their daily lives. In modern society, the process of solving problems ought to be open and accessible in order to allow those affected to voice their concerns. Thus, practitioners of reflective practice should not shy away from dealing overtly and reflectively with conflicts of views, values, and rationality (Ulrich, 2000; 2008). In the field of evaluation

 DOI 10.7172/1644-9584.80.2

it means the application of concepts of systems thinking: (1) understanding interrelationships and interdependencies between different entities associated with a complex situation (i.e. getting the 'bigger picture'), (2) engaging with multiple perspectives (appreciating different viewpoints), and (3) reflecting on boundaries (i.e. revising 'the big picture' and the viewpoints): exploring and reconciling power relations, potential conflicts between different entities and perspectives. Reynolds and Holwell (2010) point out: 'The aim here is not to provide yet another ready-to-hand matrix to offer clients through a consultancy, but rather to gently disrupt, unsettle and thereby provoke new systems thinking' (p. 17).

The critical systems heuristic (CSH) is, according to its author W. Ulrich, a framework for reflective practice based on practical philosophy and systems thinking. It is used, in particular, in equity-focused developmental evaluation (see, for example: Donaldson & Picciotto, 2016). M. Jackson (2003), a distinguished systems practitioner in the field of management sciences, acknowledged that Ulrich 'described for the first time an approach that takes as a major concern the need to counter possible unfairness in society, by ensuring that all those affected by decisions have a role in making them (cited in Reynolds & Williams, 2012 p. 115).

The basic premise of the critical systems heuristic is that the assessment of a given instrument depends on how we determine the relevant reference system, which reflects the facts we observe and the values we judge appropriate. One needs to be aware of this inevitable selectivity of how the problem is defined, what solution is proposed, what is suggested as measures of success or improvement, etc. All these claims represent only a fraction of the entire set of conceivable considerations and serve some groups better than other, as 'no proposal, no decision, no action can get it equally right for everyone' (Ulrich, 2005, p. 2). What is paramount is to make it clear to ourselves and to all concerned what the assumed reference system is and, thereby, promote mutual understanding and constructive dialogue. Critical systems heuristic can serve this purpose.

CSH is a set of twelve questions designed to expose key perspectives and boundary decisions which are subject to critique. These questions are associated with four sources of influence: (1) motivation (Where does a sense of purposefulness and value come from? / Who gets what?), (2) power/control (Who is in control of what is going on and is needed for success? / Who owns what?), (3) knowledge (What experience and expertise support the claim? Who does what?), and (4) legitimacy (Where does legitimacy lie? Who gets affected by what some people get?). Three categories are associated with each source of influence: (a) stakeholders (social roles), (b) stakes (specific concerns) and (c) stakeholding issues (difficulties, key problems), which correspond to the following questions: Who are (ought to be) the stakeholders? What is (ought to be) at stake? What are (ought to be) the opportunities for stakeholding development? (Fig. 1). Given the

dual nature of selectivity (empirical and normative), each boundary category is considered in two modes: what 'is' and what 'ought to be'. Differences between 'is' and 'ought to be' answers point to unresolved boundary issues.

SOURCES OF INFLUENCE	STAKEHOLDERS	STAKES	STAKEHOLDING ISSUES
SOURCES OF MOTIVATION	Q1 Beneficiary	Q2 Purpose	Q3 Measures of success
SOURCES OF CONTROL	Q4 Decision maker	Q5 Resources	Q6 Decision environment
SOURCES OF KNOWLEDGE	Q7 Expert	Q8 Expertise	Q9 Guarantor
SOURCES OF LEGITIMACY	Q10 Witness	Q11 Emancipation	Q12 Worldview

Fig. 1. Template for a CSH- based evaluation. Source: Reynolds 2007, p. 104.

For more information on critical systems heuristics and its practical application, see: Ulrich 2005, Reynolds 2007 and Reynolds and Williams 2012, as this article outlines only the main ideas underpinning CSH, which can serve as a framework for reflective practice, as well as emancipatory practice in the evaluation field.

5. Conclusions

Reflective practice should be at the heart of the evaluative endeavour. In the literature, reflective practice is regarded from different perspectives: personal perspective, organisational perspective and broader civil society perspective. These are not mutually exclusive, but rather intersect each other.

Whilst the traditional role of the evaluator involves making judgements about the merit and worth of an evaluand (e.g. a programme), assessing its efficiency and making recommendations for improvement, it can be argued that evaluators become increasingly concerned about enhancing the impact of evaluation through promoting the use of evaluation itself, the information generated in the evaluation process and organisational learning. Reflective practice is used not only in evaluators' own professional development, but also as a means to enhance dialogue, stakeholders involvement and organisational learning in the evaluation process (collaborative reflective practice).

In the organisational perspective, the theory of single, double and triple-loop learning is very frequently discussed in the context of evaluation. Given the three main types of evaluation – formative, summative and

DOI 10.7172/1644-9584.80.2

developmental – each of them can serve learning purposes. The first, i.e. formative evaluation, represents single-loop learning, since the evaluation process is structured around the following question: how can existing action be performed better? In other words: are we doing things right? In reference to summative evaluation, a parallel can be drawn with double-loop learning. Evaluation efforts focus on the question whether what we do is the right thing to do. This, in turn, entails gaining deeper understanding of how a system, in which an intervention takes place, functions, and which system dynamics led to the problem in the first place. Finally, developmental evaluation represents triple-loop learning, which can be defined as 'learning how to learn' (Patton, 2017, p.10). However, triple-loop learning can be also understood by addressing political dimension behind learning (Flood and Romm, 1996). From this point of view, an evaluation process represents triple-loop learning when it seeks to examine relations of power which determine what particular purposes are being privileged and valued, and how we know what the right thing to do is. Including political considerations in the evaluative endeavour resonates with the next perspective, from which reflective practice is discussed in the context of evaluation: the civil society perspective.

As Ulrich (2000) advocates for a revised definition of civil society, the issue of enabling active participation in the governance of collective affairs is brought to the forefront. Based solely on their expertise in the field, professionals, including evaluators, should not be given authority over ordinary citizens in matters that affect the latter's daily lives. Thus, practitioners of reflective practice should not shy away from dealing overtly and reflectively with conflicts of views, values, and rationality.

References

Alkin, M. (2013). *Evaluation Roots. A Wider Perspective of Theorists' Views and Influence*. Thousand Oaks, CA: SAGE Publications.

Argyris, C. & Schön, D. (1978). *Organizational learning: A theory of action perspective*. Reading, Mass: Addison Wesley.

Bateson, G. (1973). *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution and Epistemology*. London: Paladin, Granada.

Campbell, D. (1984). Can we be scientific in applied social science. *Evaluation Studies Review Annual*, *9*. Beverly Hills, CA: Sage Publications.

Chen, H.-T. (1996). A comprehensive typology for program evaluation. *Evaluation Practice*, *17*(2), 121–130. https://doi.org/10.1177/109821409601700204.

Cruise, P. (1999). Values, Program Evaluation and the New Public Management. *International Journal of Organization Theory and Behavior*, (2/3&4), 383–412.

De Laat B. & Williams, K. (2014). Evaluation Use Within the European Commission (EC): Lessons for the Evaluation Commissioner. In: M. Loud, J. Mayne (eds), *Enhancing Evaluation Use. Insights from Internal Evaluation Units* (p. 147–174). Beverly Hills, CA: Sage Publications.

Dickson, R. & Saunders, M. (2014). Developmental evaluation: Lesson for evaluative practice from the SEARCH Program. *Evaluation*, *20*(2), 176–194.

Donaldson, S. & Picciotto, R. (2016). *Evaluation for an Equitable Society*. Charlotte, NC: Information Age Publishing, Inc.
Fiol, C. & Lyles, M. (1985). Organizational Learning. *The Academy of Management Review*, (10/4), 803–813.
Flood, R. & Romm, N. (1996). *Diversity Management: Triple-loop Learning*. Chichester: John Wiley & Sons.
Gamble, J. (2008) *A Developmental Evaluation Primer*. Ottawa: The J.W. McConnell Family Foundation.
Horsch, K. (1998). Evaluation in the 21st Century. Interview with Carol H. Weiss. *The Evaluation Exchange*, (IV/2), 5–6.
Isaacs, W. (1993). Taking flight: Dialogue, collective thinking, and organizational learning. *Organizational Dynamics*, *22*(2), 24–39.
Jones, D. & Stubbe, M. (2004). Communication and the reflective practitioner: A shared perspective from sociolinguistics and organisational communication. *International Journal of Applied Linguistics*, (14/2), 185–211. https://doi.org/10.1111/j.1473-4192.2004.00059.x.
Kubera, P. (2017). Conceptual Framework for Evaluation of Economic Impacts of RDI Instruments. *Journal Association SEPIKE*, 17, 90–96.
Luo, H. (2010). The Role for an Evaluator: A Fundamental Issue for Evaluation of Education and Social Programs. *International Education Studies*, (3/2), 42–50.
Markiewicz, A. & Patrick, I. (2016). *Developing Monitoring and Evaluation Frameworks*. USA: SAGE Publications Inc.
Mc David, J., Huse I. & Hawthorn, L. (2013). *Program Evaluation and Performance Measurement. An Introduction to Practice* (2nd ed.). Thousand Oaks: SAGE Publications.
Morabito, S. (2002). Evaluator Roles and Strategies for Expanding Evaluation Process Influence. *American Journal of Evaluation*, *23*(3), 321–330.
OECD. (2010). *Glossary of Key Terms in Evaluation and Results Based Management*. Paris: OECD.
Patton, M. (1996). A world larger than formative and summative. *Evaluation Practice*, *17*(2), 131–144. https://doi.org/10.1177/109821409601700205.
Patton, M. (2007). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.
Patton, M. (2011). *Developmental evaluation. Applying Complexity Concepts to Enhance Innovation and Use.* New York: Guliford Press.
Patton, M. (2012a). Developmental Evaluation: Applying complexity concepts to enhance innovation and use. Report from an Expert Seminar with Dr Michael Quinn Patton, Centre for Development Innovation, Wageningen University & Research centre.
Patton, M. (2012b). *Essentials of Utilization-Focused Evaluation* (3rd ed.). Thousand Oaks, CA: SAGE Publications.
Patton, M. (2017). Developmental evaluation. In: J. Pokorski, Z. Popis, K. Herman-Pawłowska (eds), *Theory-based evaluation in complex environments* (pp. 7–19). Warsaw: PARP.
Patton, M.Q. (2007). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.
Preskill, H. & Torres, R. (1999). *Evaluative Inquiry for Learning in Organizations*. Thousand Oaks, CA: SAGE Publications.
Preskill, H., Zuckerman, B. & Matthews, B. (2003). An exploratory study of process use: Findings and implications for future research. *American Journal of Evaluation*, *24*(4), 423–442. https://doi.org/10.1177/109821400302400402.
Reynolds, M. & Holwell, S. (2010). Introducing Systems Approaches. In: M. Reynolds, S. Holwell (eds), Systems Approaches to Managing Change: A Practical Guide (pp. 1–24). London: Springer.
Reynolds, M. & Williams, B. (2012). Systems thinking and Equity-focused evaluations. In: M. Segone, M. Bamberger (eds), *Evaluation for equitable development results* (pp. 115–141). New York: UNICEF.

 DOI 10.7172/1644-9584.80.2

Reynolds, M. (2007). Evaluation based on critical systems heuristics. In: B. Williams, I. Imam, (eds), *Using Systems Concepts in Evaluation: An Expert Anthology* (p. 101–122). Point Reyes CA, USA: Edge Press.
Reynolds, M. & Holwell, S. (2010). Introducing Systems Approaches. In: M. Reynolds, S. Holwell (eds), *Systems Approaches to Managing Change: A Practical Guide*, Springer, 1–23.
Reynolds, M. & Williams, B. (2012). Systems thinking and Equity-focused evaluations. In: M. Segone, M. Bamberger (eds), *Evaluation for equitable development results* (p. 115–141). New York: UNICEF.
Rossi P, Freeman, H. & Lipsey, M. (2004). *Evaluation. A Systemic Approach* (7t ed.). Thousand Oaks: SAGE Publications.
Schön, D. (1983). *The reflective practitioner*. New York: Basic Books.
Schwandt, T. (2001). *Dictionary of qualitative inquiry* (2nd ed.). Thousand Oaks, CA: SAGE.
Scriven, M. (1986). New frontiers of evaluation. *Evaluation Practice*, *7*, 7–44.
Scriven, M. (1991). Beyond formative and normative evaluation. In: G. McLaughlin, D. Philips (eds), *Evaluation and Education: At quarter century* (pp.19–64). Chicago, Il: University of Chicago Press.
Shadish, W., Cook, T. & Leviton, L. (1991). *Foundations of program evaluation*. Newbury Park, CA: Sage.
Skolits, G., Morrow, J. & Burr, E. (2009). Reconceptualizing Evaluator Roles. *American Journal of Evaluation*, 30(3), 275–295.
Smith, T., (2014). *Reflective Practice, Collaboration, and Stakeholder Communication: Where Does the Field of Evaluation Stand?* PhD diss. University of Tennessee. Retrieved from: http://trace.tennessee.edu/utk_graddiss/2860/.
Stake, R. & Trumbull, D. (1982). Naturalistic generalizations. *Review Journal of Philosophy & Social Science*, 7, 1–12.
Stame, N. (2004). Theory-based Evaluation and Types of Complexity. *Evaluation*, (10/1), 58–76.
Stevahn, L., King, J., Ghere, G. & Minnema, J. (2005). Establishing essential competencies for program evaluators. *American Journal of Evaluation*, *26*(1), 43–59, http://dx.doi.org/10.1177/1098214004273180.
Stufflebeam, D., Madaus, G. & Scriven, M. (2000). Program Evaluation. A Historical Overview. In: D. Stufflebeam, G. Madaus, T. Kellaghan (eds), *Evaluation Models Viewpoints on Educational and Human Services Evaluation* (2nd ed., p. 3–18). Boston: Springer, Kluwer.
Thompson, N. & Pascal, J. (2012). Developing critically reflective practice. *Reflective Practice*, (13/2), 311–325. http://dx.doi.org/10.1080/14623943.2012.657795.
Thompson, N. (2000). *Theory and practice in the human services* (2nd ed.). Buckingham: Open University Press.
Tosey, P., Visser, M. & Sounders, M. (2012). The origins and conceptualizations of 'triple-loop' learning: A critical review. *Management Learning*, (43/3), 291–307.
Ulrich, W. (2000). Reflective Practice in Civil Society: The Contribution of Critically Systemic Thinking. *Reflective Practice*, (1/2), 247–268. Retrieved from: http://wulrich.com/downloads/ulrich_2000a.pdf.
Ulrich, W. (2005). *A brief introduction to critical systems heuristics (CSH)*. Web site of the ECOSENSUS project, Open University, Milton Keynes, UK, 14 October 2005. Retrieved from: http://www.ecosensus.info/about/index.html.
Ulrich, W. (2008). Reflections on Reflective Practice. *The Mainstream Concept of Reflective Practice*, (1/7). Retrieved from: http://www.wulrich.com/bimonthly_march2008.html.
Vince, R. (2002). Organizing Reflection. *Management Learning*, (33/1), 63–78. https://doi.org/10.1177/1350507602331003.

Weiss, C. (1988). If Program Decisions Hinged Only on Information: A Response to Patton. *Evaluation Practice*, (9/3), 15–28.

Yuthas, K., Dillard, J. & Rogers, R. (2004). Beyond agency and structure: Triple-loop learning. *Journal of Business Ethics*, *51*(2), 229–243.

 DOI 10.7172/1644-9584.80.2