MERIT RESEARCH JOURNALS
www.meritresearchjournals.org

Merit Research Journal of Education and Review (ISSN: 2350-2282) Vol. 11(9) pp. 145-159, November, 2023
Available online http://meritresearchjournals.org/er/index.htm

DOI: 10.5281/zenodo.10323516

Original Research Article

A Critical Review of Theories and Approaches to Interpreter Training Programme Evaluation

Azambou Ndongmo Laurent Roger

Advanced School of Translators and Interpreters (ASTI), University of Buea, Cameroon

Email: azamboulaurent@yahoo.fr

Abstract

The number of programmes offering interpreter training courses has tremendously increased over the years and in many parts of the world. Caminade and Pym (1995) list more than 250 university programmes in more than 60 countries, but since the 1990s, many more programmes have beenset up, in particular in China where the government has recently decided to set up MTIs, Masters in Translation and Interpreting (Gile, 2009:17). This has led to programme evaluation becoming increasingly relevant as a research sub-area, both for students and researchers interested in interpreter training as a research area. Given this developing interest, the current review sets out to present the various theories, models, methods, and approaches that have been developed from the 1960s till the 2000s to guide programme evaluation, especially training programme evaluation. They are presented and discussed in detail to enable researchers in the sub-area to make informed choices depending on the purpose of the evaluation and its expected outcomes. Until more tools emerge, this contribution may stand as a handbook for researchers in need of theoretical tools to evaluate complete training programmes or specific aspects related to that, especially in the area of interpretation. This attempt is all the more useful as conference interpretation is a relatively young academic discipline. It is the author's hope that this contribution provides the new pool of researchers who are emerging and specialising in interpretation research in Cameroon and beyond with a useful and ready-to-use tool for interpreter training programme evaluation.

Keywords: Programme evaluation, Interpreter training programme, Interpreter education, Informed choices, Evaluation theories

INTRODUCTION

Todate, three areas of focus have drawn the attention of researchers in interpretation studies: the cognitive aspect of interpretation, training in interpretation, and professional topics. Pöchhacker (2004:3, quoted in Motale, 2018:2) would add a fourth area, which is the history of interpreting. However, none of the above areas have witnessed important developments, probably because very few interpreters believe and develop an interest in research. Dodds and Katan (1997) state as follows:

At a round table on the interaction between interpreter training and research held at a conference in Turku, Finland in 1994, Dodds (Dodds and Katan 1997) reached the conclusion that the participants were not very interested in the application of research to training.

Sandrelli (2005:2) asserts that: Indeed, conference interpreting is a relatively young academic discipline and although the literature on interpreting pedagogy and interpreter training exercises is now fairly abundant, there are very few published contributions on the overall

structure and contents of the interpreter training curriculum.

However, other fields of studies have developed evaluation theories, approaches, and models that can be adopted and adapted for evaluating training programmes in general, interpreter training programmes, and a particular interpreter training programme. This is of no surprise given the developingnature of the discipline and the preeminence of the training rather than the education approach at the early stages. Only the time factor, aided by the reforms that affected the whole tertiary education sector in the 1980s and early 1990s, would overcome resistance to any form of research. Not only did the reforms make tertiary education institutions more professional, but they also made vocational institutions more academic because most of those which were purely vocational were linked to academic departments which also conducted research or underwent transformations to take into account the research component. Pöchhacker (2003:106) is of the opinion that the absence of studyon interpreter training is because interpreting research has focused too intently on the cognitive mechanics of the process, especially in simultaneous interpreting, with no great breakthroughs being made. Thus, among the three main areas of research focus, one seems to have monopolised most researchers with rather poor results. But there are still hopes, especially with the publication, since 2007, of *The Interpreter and Translator Trainer*, an international journal dedicated to research on training (Kelly and Martin, 2009:299).On another note, in this era of Information and Communications Technology, its pervasive evolution spares neither interpretation practice nor interpreter training.

According to al-Zahran (2007), three factors led to the development of Interpretation Studies (IS) in general and conference interpretation in particular: i. Degree and research training programmes that led to the production of a great number of theses at the MA level and a certain number of dissertations at the PhD level, with the latter playing a crucial role in the development of IS. ii. Conferences thatenabled researchers from all over the world to meet, discuss and exchange ideas. iii. The publication of interpreting journals on a regular basis, which creates a demand for research papers. Another important factor that might have played a role is globalisation, a phenomenon which started in the 1990s.

This paperattempts to review and discuss a number of interpreter training programme evaluation theories, approaches, and models.

Defining Some Key Concepts

This section attempts to define the following concepts: Programme, Curriculum, Syllabus, Evaluation, and Programme Evaluation.

Programme

The term programme is one of those terms so commonly used that no one actually stops and reflects on its real content. The impression is, therefore, that everyone knows what it means. However, in a paper having the term as one of its main keywords, it is important to clarify the twists and turns it takes. Although the dictionary is not the best source of definitions of technical terms, it may be a good place to start explaining their meaning, given their centrality in this study. This applies for 'Programme' and 'Evaluation'. The 2000 Merriam-Webster electronic version proposes for 'Programme'[4 b (1)]: "a plan determining the offerings of an educational institution: CURRICULUM".

This definition already suggests a relationship between 'Programme'and 'Curriculum'. Besides, by referring to 'educationalinstitution', it takes us closer to the understanding of 'programme'in this study, i.e. 'educational programme'. Indeed, literature abounds on 'programme evaluation' but little technical literature is found which defines 'programme'as understood in the context of this study in unambiguous terms.It is when a qualifier (educational) is appended to the term that its real meaning as per this research starts emerging. It will thus be defined as 'educational programme' rather than simply 'programme'. Indeed, the latter pertains first and foremost to the business and management area. It is borrowed from that area into educational institutions perceived as businesses which are to be managed according to admitted and well-grounded management principles. Therefore, what is an 'educational programme' (or 'training programme')?

According to Gile (1995: 12, cited by Ulrych, 2005:11),

Most professional interpreter and translator training programs worldwide consist essentially of practical interpretation and translation exercises: a source-language text or speech is selected, students are invited to interpret or translate it, and the result is commented on and corrected by or under the guidance of the instructor. Judging by the literature, there does not seem to be any disagreement between teachers, practitioners, or students as to the principle that training should consist essentially of such exercises, although there are differing opinions as to implementation with respect to duration, progression, types of materials used, admission standards, graduation standards, etc.

From the above, it is understood that an educational programme is characterised by its duration, its progression, the types of materials used, its admission standards, graduation standards, etc. A more straight-forward definition is borrowed from Fiola (2003) about professional translation, and can also apply to professional interpretation: "Le terme « programme d'études » désigneici un ensemble intégré de mesuresvisant au développement des savoirs, des savoir-faire et des savoir-êtrenécessaires pour qu'un-

apprenantsoit compétent entraduction professionnelle [The term 'educational programme' refers here to an integrated set of measures aimed at developing the knowledge, know-how, and interpersonal skills required for a learner to be competent in professional translation]" (My translation).

From another perspective, Gile (2009:7) defines an interpreter training programme through the functions it performs at individual, social, and professional levels. At the individual level, it can help individuals who wish to become professional interpreters enhance their performance to the full realisation of their potential; it can also help them develop their interpretation skills more rapidly than through field experience and self-instruction. At the social and professional level, it can help raise general professional standards in the marketplace by selecting the best candidates at admission and the truly skilled at graduation. This may, in turn, positively impact the social status of interpreters, especially if standards are set at postgraduate academic level. Through the professional circles they are connected to, training programmes can also help beginning interpreters and translators start their professional careers by introducing them to professional organisations and clients. This is a particularly important function in conference interpretation, as interpretation schools maintain close links with major international organisations and other institutional clients and invite their representatives to take part in graduation examinations. Though important, these social functions are context-dependent and vary considerably from one country to the next and from one market to the next. In contrast, the didactic function of a formal training programme is essentially invariant, although the application of common principles is also context-dependent, as operational aims vary according to the nature, duration, resources and baseline status of the relevant training programmes. This latter Gile (2009)'s perception of a programme through the terms 'nature, duration, resources and baseline status of the relevant training prograrmmes'meets up with the perception that was already manifest in the same author's (1995) indented quotation above through the terms 'duration, progression, types of materials used, admission standards, graduation standards'. This convergence portrays the broad characteristics of what should be understood as an interpreter educational programme in the present paper.

Curriculum

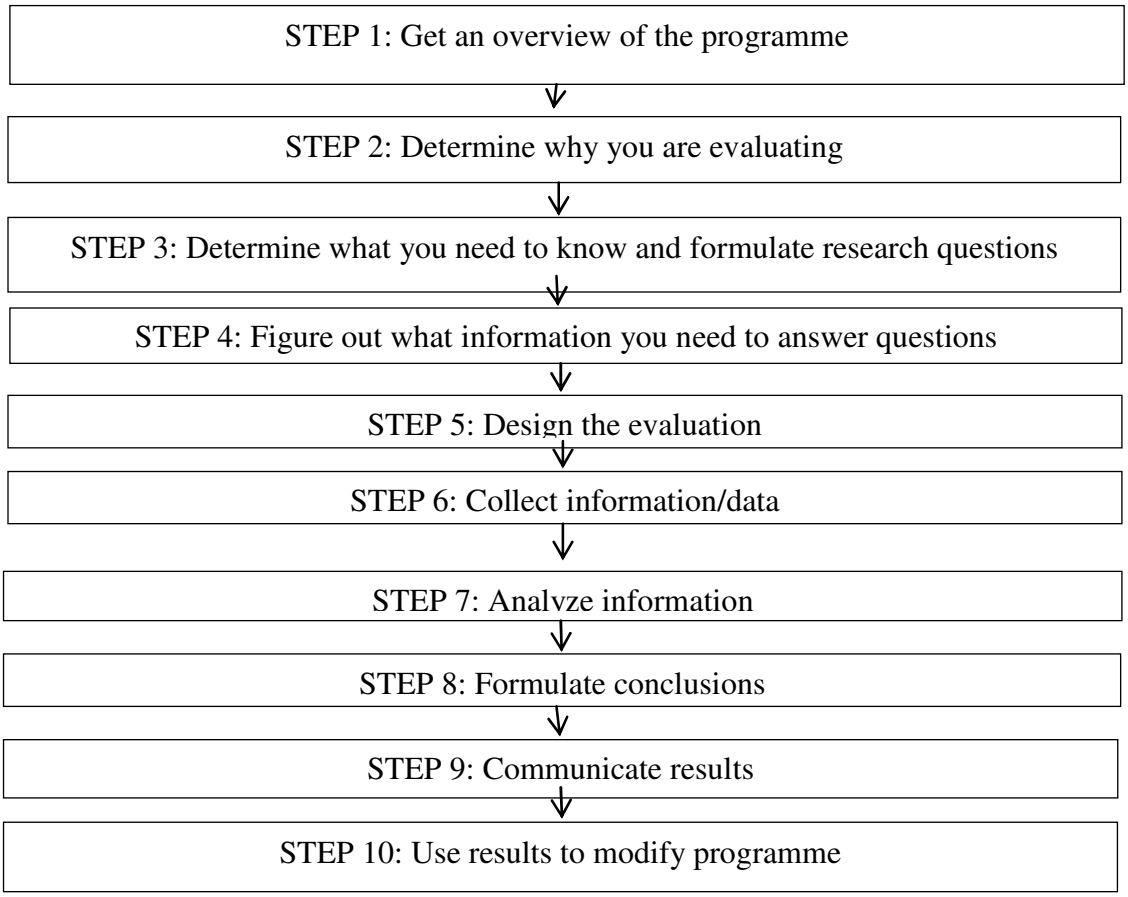
The words curriculum and programme sometimes confuse the minds of some people because they are often used interchangeably. However, the two words are closely related but different from each other. A curriculum is the content of a programme; it is set or determined by an external body that has the authority to administer the programme to students. The rule of demand and supply may influence change in programmes' curricula. So, within the same programme, a student may choose a range of courses to specialise in some aspects of the programme. A curriculum also includes the way the content is administered, including teaching and performance assessment methods. Olivia (2011), in a post online, summarises the difference as follows: "Different degree or diploma courses available in various streams of study are labelled as programmes, whereas the content that is used to make up these study programmes and the way it is administered is called curriculum". Another word that sometimes adds to the confusion is 'syllabus'.

Syllabus

Just like programme and curriculum, curriculum and syllabus are two important words in the field of education, which are often confused as if they mean the same thing. Curriculum is a word that refers to the subjects that are studied or prescribed for study in a programme, and it covers the whole course experience. It is a much broader concept, whereas syllabus is much narrower, and covers only a part of the course experience. Syllabus is, therefore, a subset of curriculum. Syllabi make a curriculum. A particular syllabus is covered each year for a specific subject throughout the duration of a programme. A syllabus is descriptive because it is formed to create an understanding between the teacher and the student. A curriculum that contains all the objectives of the whole programme, is prescriptive; it is a guide the institution follows for the programme as long as it lasts.

Evaluation

As was the case for 'Programme'(see section 2.1 Programme above), the dictionary meaning of 'Evaluation' will first be considered. The 2000 version 2.5 of the Merriam-Webster states that 'Evaluation' is "the act or result of evaluating: JUDGMENT, APPRAISAL, RATING, INTERPRETATION". While running a programme, important decisions must be made, especially if theprogramme's environment, as can be expected, undergoes significant changes. When, for example, competition is mounting in the immediate environment of the programme (within the country), evaluation results may be used to revise the programme operation, obtain additional funding to expand the programme, or simply terminate the programme (Holden and Zimmerman, 2008). When systematically conducted, evaluation can also provide the information needed for continuous improvement in a globalised environment where stagnation is equivalent to death.



Source: (www.evaluationwiki.org/wiki/index.php/Evaluation, 17 January 2017).
Figure 1. The ten distinct steps each evaluation usually follows

Trochim (2006, cited by Zinovieff and Rotem, 2008:5) first defines evaluation as "the systematic assessment of the worth or merit of some object." Secondly, he states that evaluation is "the systematic acquisition and assessment of information to provide useful feedback about some object."

The second technical definition is less judgemental and more flexible than the former. Some evaluations, like in descriptive studies or formative evaluations, may process information for feedback purposes. Thus, evaluation is not carried out for the sake of evaluating, but its main goal is to provide some type of useful, meaningful feedback to stakeholders. The information is not only acquired, but it is assessed, systematically, like in an organised manner or using a method. Various types of evaluation will be discussed in subsequent sections, as well as the major models and methodologies commonly applied in evaluation processes.

Muraskin (cited by Zinovieff and Rotem, 2008:6) sees evaluation as the systematic collection and analysis of data needed to make decisions. As can be seen, feedback will lead to decision-making; this takes this definition closer to that provided by Trochim above. Whether or not an assessment of the worth or merit of the programme results in a change is not much the concern of the evaluator, but of the people in charge of making decisions about the programme. The ten distinct steps each evaluation usually follows are: Figure 1

The dimensions, goals, approaches, purposes, and types of evaluation will be discussed in subsequent sections.

Programme Evaluation

The concepts of 'Programme'and of 'Evaluation'have been discussed separately; the following statement from Zinovieff and Rotem (2008:13) now summarises what should be understood by 'Programme Evaluation':

Programme evaluation is carefully collecting information about a programme or some aspect of a programme to make necessary decisions about it. Programme evaluation can include any or a variety of at least 35 different types of evaluation (according to Patton, others have identified even more). The type and model of evaluation one undertakes to improveone's programmes

depends on what one wants to learn about the programme. One should worry less about what type/model of evaluation one needs and worry more about what one needs to know to make the programme decisions one needs to make and also worry about how one can accurately collect and understand that information.

Theoretical Perspectives on (Programme) Evaluation and Other Relevant Theories

Many specialists have presented their views on evaluation in general and programme evaluation in particular.

Evaluation: A Multifaceted Task

Zinovieff and Rotem (2008) proposed an interesting review of some facets of programme evaluation. Their approach includes two moves: first, they extensively discuss evaluation in general through its definition (see section 2.4. above), dimensions, goals, approaches, purposes, and types; secondly, they specifically target programme evaluation and lengthily describe various models/methods of programme evaluation, with two sub moves of interest to this study: objective/goal-based programme evaluation, and systems-based programme evaluation. Following is the review of their discussion:

Dimensions

In evaluation literature, dimensions include the fundamental concepts of process, outcome and impact.
i. **Process evaluation** assesses programme activities. It describes the interventions to which any findings about outcomes may be attributed.
ii. **Outcome evaluation** is concerned withthe immediate or direct effects of the programme on participants. Beyond knowledge and attitudes, this can extend to behavioural effects.
iii. Finally, **impact evaluation** considers the long-term as well as the unintended effects of the programme.
All the three dimensions discussed above by these authors are relevant to this review. However, the review goes further and investigates some dimensions, like context and initial input, which are not covered here.

Goals

Evaluation generally aims at providing useful feedback to a variety of actors. It is expected that the empirical data provided by the evaluation would influence decision-making or policy formulation.

Approaches

These approaches are a general way of looking at or conceptualising evaluation. Duignan (2001) describes seven main evaluation approaches:
a- Utilisation-focused evaluation – determines methods on the basis of what is going to be most useful to different audiences;
b- Empowerment evaluation – emphasises that the evaluation process and methods should be empowering to those who are being evaluated;
c- Stakeholder evaluation – looks at the differing perspectives of all of a programme's stakeholders (those who have an interest in it);
d- Goal-free evaluation – in which the evaluator's task is to examine all of the outcomes of a programme, not just its formal outcomes as identified in its objectives;
e- Naturalistic or 4th generation evaluation – emphasises the qualitative uniqueness of programmes and is a reaction against the limitation of quantitative evaluation approaches;
f- Theory-based evaluation – emphasisesdetailing the assumptions on which a programme is based (intervention logic) and follows those steps to see if they occur;
g- Strategic evaluation – emphasises that evaluation design decisions should be driven by the strategic value of the information they will provide for solving social problems.

Among the above approaches, three can be of particular interest because of their practical value: utilisation-focused evaluation which talks about the methods,goal-free evaluation which suggests going beyond the formal outcomes as identified in the programme's objectives,and strategic evaluation which means that design decisions should be driven by the strategic value of the information they will provide for solving social problems. The others are less practical.

Purposes

The evaluation purpose identifies what evaluation activity is going to be used for. A particular evaluation activity can, however, have more than one purpose. It may also relate to the intent of the evaluation (Zinovieff and Rotem, 2008):
a- Gain insight – provide the necessary insight to clarify how programme activities should be designed to bring about expected changes;
b- Change practice – improve the quality, effectiveness, or efficiency of programme activities;
c- Assess effects – examine the relationship between programme activities and observed consequences;
d- Affect participants – use the processes of evaluation to affect those who participate in the inquiry. The systematic reflection required of stakeholders who

participate in an evaluation can be a catalyst for self-directed change. Evaluation procedures themselves will generate a positive influence.

Types

Literature traditionally differentiates between the formative and the summative types of evaluation (see, for example, Trochim, 2008). However, two other concepts have emerged in the recent years and received much attention from researchers (see, for example, Dessinger and Morley, 2003): confirmative and meta-evaluation.

a- **Formative evaluation** examines the delivery of a programme by considering the quality of its implementation, assessing the organisational context, personnel, procedures, inputs, etc. (Zinovieff and Rotem, 2008); it includes several sub-types:

i. **Needs assessment**: who needs the programme, how great the need is, and how to meet the need.

ii. **Evaluability assessment**: is an evaluation feasible? Do stakeholders consider it useful?

iii. **Structured conceptualisation**: this sub-type helps stakeholders define the programme, its target population and possible outcomes.

iv. **Implementation evaluation**: takes care of the fidelity of the programme delivery.

v. **Process evaluation**: is concerned with the process of delivering the programme and considers alternative delivery methods.

b- **Summative evaluation** examines the effects or outcomes of a programme by describing what happens subsequent to the delivery of the programme and investigates the causal relationship between the programme and the outcome (Zinovieff and Rotem, 2008). It also includes several sub-types:

i. **Outcome evaluation**: did the programme cause demonstrable effects on specifically defined target outcomes?

ii. **Impact evaluation**: it is a broader assessment of the overall or net effects (intended or unintended) of the programme.

iii. **Cost-effectiveness and cost-benefit analysis**: this aspect of evaluation addresses questions of efficiency by standardising outcomes in terms of their costs and values.

iv. **Secondary analysis**: this is a re-examination of existing data to find answers to new questions or use different methods.

v. **Meta-analysis**: this sub-type uses the outcome estimates from multiple studies to arrive at an overall or summary judgement on an evaluation question.

Apart from these two traditional evaluation types (formative and summative), two other types have emerged and are gradually gaining ground. With the formative and summative types of evaluation, the confirmative and meta types combine to form what is referred to as 'full-scope evaluation' (Dessinger and Morley, 2003).

c- Confirmative evaluation goes beyond formative and summative evaluations in that:

The evaluation and training practitioner collects, analyses, and interprets data related to behaviour, accomplishment, and results in order to determine the continuing competence of learners or the continuing effectiveness of instructional materials and to verify the continuous quality improvement of education and training programmes (Zinovieff and Rotem, 2008:11).

Indeed, the evaluation task will not be complete if feedback is not obtained from the relevant stakeholders (behaviour, accomplishment). This feedback will serve the purpose of at least ensuring the 'continuous effectiveness of instructional materials'.

Apart from assessing level one (focus on the needs of learners and developers at draft form level) and level two (monitoring of the programme during its full implementation, focusing on the needs of learners and the programme objectives) of the evaluation process, what formative and summative evaluations do, confirmative evaluation goes a step beyond (level three) and assesses the transfer of learning to the real world.

d- Meta evaluation is all about evaluating the evaluation and includes formative, summative and confirmative evaluations. The evaluator investigates how the evaluation was conducted to validate its inputs, process, outcomes and outputs. Zinovieff and Rotem (2008:11) report two types of meta evaluations.

i. Type one meta evaluation is conducted con-currently with the evaluation process. This is why it is literally referred to as the formative evaluation of evaluation.

ii. Type two meta evaluation is the more common approach. It is conducted after formative, summative, and after at least one cycle of confirmative evaluation.

What actually happens subsequent to the delivery of the programme (see discussion of confirmative evaluation above) is as important as what happens before and during the programme processes.

The various types of evaluation discussed above could also be tackled more thematically rather than conceptually (Zinovieff and Rotem, 2008:12); in this frame of mind, they would be organised in terms of goal-based evaluation, process-based evaluation, and outcomes-based evaluation.

e- A goal-based evaluation evaluates the extent to which a programme is meeting predetermined goals or objectives. The following questions are to be asked while designing this type of evaluation:

i. How were the programme goals (and objectives, if applicable) established?

ii. Was the process effective?

iii. What is the status of the programmeprogress toward achieving the goals?

iv. Will the goals be achieved according to the timelines specified in the programme implementation or operations plan? If not, then why?
v. Do personnel have adequate resources (money, equipment, facilities, training, etc.) to achieve the goals? (Zinovieff and Rotem, 2008).

f- A process-based evaluation is geared toward fully understanding how a programme works – how it producestheresults that it does. This type of evaluation is useful for long-standing programmes, especially if they have changed over the years and if, for example, there are reports from stakeholders of inefficiencies. It is also useful in portraying to outside parties how a programme operates. There are numerous questions that might be addressed in a process evaluation, examples of which include:
i. On what basis is it decided that products or services are needed?
ii. What is required in order to deliver the product or services?
iii. How is the personnel trained on how to deliver the product or services?
iv. How do customers or clients come into the programme?
v. What is required of customers or clients? (Zinovieff and Rotem, 2008).
These questions can be selected carefully dep-ending on what is important to know about the programme.

g-An outcomes-based evaluation tries to ascertain if the right programme activities are carried out to bring about the expected outcomes. Outcomes are in terms of enhanced learning (knowledge, perceptions/attitudes or skills) or conditions, and are often confused with programme outputs or units of services, e.g. the number of clients who went through a programme (see www.unitedway.org/outcomes/).

The themes of goal, process, and outcome may all be concerns of a particular study, just like the concepts of formative, summative, confirmative, and meta types of evaluation. In that case, such a study will use both the thematic and conceptual types of evaluation to establish its findings.

The above review of Zinovieff and Rotem (2008)'s approach covers several facets to consider when evaluating a programme. However, it was designed more specifically to assess fellowship training programmes, especially in their long-term impact(s). Therefore, as such, its findings are not necessarily applicable to more formal educational training programmes. Indeed, important aspects like the context or the initial input that need special consideration while evaluating educational training programmes, are not given due attention by the two authors. Though highly inspiring, their method cannot always be privileged.

Other Relevant Theories

This section presents other relevant training programme theories.

Programme Theory

The evaluation field has seen programme theory gradually gaining interest. Multiple terms are used to refer to it including programme theory, theory-based (Weiss, 1995, 1998, cited by Rogers, 2000:30), theory driven (Chen, 1990, cited by Rogers, 2000:30), and programme theory evaluation (Rogers, 2000a, cited by Sharpe, 2011:72). It refers to a variety of ways of developing a causal model linking programme inputs and activities to a chain of intended or observed outcomes, and then using this Model to guide the evaluation (Rogers et al., 2000).

The function of a programme theory is to ascertain the theoretical sensibility of the programme. A programme theory consists of a set of statements that describe a particular programme, explain why, how, and under what conditions the programme effects occur, predict the outcomes of the programme, and specify the requirements necessary to bring about the desired programme effects (Sedani and Sechrest, 1999, cited by Sharpe, 2011).

The conceptual foundation constitutes the primary stage to any programme development. The programme theory can then be used to develop the outcomes of the programme. It should be developed prior to the commencement of the programme (Bickman, 1987; Prosovac and Carey, 1997; Rogers et al., 2000), though this is not always the case (Bickman, 1987; Reynolds, 1998: Rogers et al, 2000; Stufflebeam, 2000). In the event the programme is underway, it is important for a programme theory to still be developed. Therefore, programme theories can be developed during the operation of the programme (Rogers et al, 2000) or prior to evaluating a programme (Bickman, 1987). The development of a programme theory is necessary when hoping to determine why a programme is succeeding or failing and if and where programme improvement should be focused (Sharpe, 2011). Earlier, Sharpe (2001:72) had stated the following:

Program theory modelling uses three components to describe the program: the program activities or inputs, the intended outcomes or outputs, and the mechanisms through which the intended outcomes are achieved (Reynolds, 1998; Rogers, 2000; Rogers et al, 2000; Sedani & Sechrest, 1999). A description of the critical inputs defines the components of the program, describe how these components are delivered, define the strength or amount of treatment required to induce the outcome (Sedani and Sechrest, 1999), and outline the required

Table 1. Simple, Complicated and Complex Problems (Glouberman and Zimmerman, 2002)

Simple: Following a recipe	Complicated (lots of parts): Sending a rocket to the moon	Complex (uncertain and emergent): Raising a child
The recipe is essential.	Formulae are critical and necessary.	Formulae have a limited application.
Recipes are tested to assure easy replication.	Sending one rocket to the moon increases assurance that the next will be OK.	Raising one child provides experience but no assurance of success with the next.
No particular expertise is required but cooking expertise increases success rate.	High levels of experience in a variety of fields are necessary for success	Expertise can contribute but is neither nor sufficient to assure success
Recipes produce standardised products	Rockets are similar in critical ways	Every child is unique and must be understood as an individual
The best recipes give good results every time	There is a high degree of certainty of outcome	Uncertainty of outcome remains
Optimistic approach to problem-solving	Optimistic approach to problem-solving	Optimistic approach to problem-solving

aspects vital in producing the expected outcomes (Lipsey, 1993). The processes that the outcome is contingent upon (Lipsey, 1993) and that follow the inputs should be described. […]. Implementation issues or resources necessary for carrying out the program's services (Bickman, 1987; Lipsey, 1993; Sedani and Sechrest, 1999) should also be detailed. For example, resources and implementation issues may include supplies, materials, and skills (Sedani & Sechrest, 1999).

Thus, a programme theory can clarify the programme's perspective; the quality of the programme can then be based on this defined perspective.

In a theory-based evaluation, the impact of clients onthe programme, utilisation of services, exogenous factors, relationships and interrelationships among programme components, the complexity of the programme, etc., must be considered (Sharpe, 2011).

Complexity Theory

Rogers (2008) brings together the programme theory and the complexity theory and explains their relationship and how both may be used for evaluating interventions with complicated or even complex aspects. "Complexity theory provides a promising framework for organising and conducting social work research and evaluation" (Wolf-Branigin 2013). This author goes further to write:

Although complexity theory evolved within the natural sciences, it serves as an approach for understanding the interactions of networks of services, and the evolution of policies and services within our discipline. Following social work's long tradition of framing social behavior in the social environment using systems theory, recent alternative schemes based on postmodernism and quantum mechanics arose. These schemes support the notion that outcomes from social work interventions were unpredictable (Pozatek 1994, quoted in Wolf-Branigin 2013: 3).

The following comparative table summarises the three-part distinction between what is complicated (lots of parts) and what is complex (uncertain and emergent): Table 1

The evaluation of simple interventions is intended to either develop or test the 'recipe'that others can then follow. Complicated interventions that have many components pose challenges to evaluations, given the limited number of variables that can be identified and empirically investigated. However,it is complex interventions that present the greatest challenge for evaluation and for the utilisation of evaluation because the path to success is so variable and cannot be articulated in advance (Rogers, 2008:31).

However, systems cannotalways be classified in a Manichean way as either simple, complicated, or complex. Some complicated or complex systems may well have some simple aspects in them. Such simple aspects may be subjected to known causal processes. Even though simple logic models may not always be appropriate for human service interventions such as education, it is often most beneficial within the framework of an evaluation endeavour to adapt a simple logic model for a better understanding of the hypothesised relationships between the various conceptual compartments of the programme at its differentdiachronic stages. However, external factors, such as environmental factors, prejudices, assumptions, etc., are to be considered. Figure 2 below gives an idea of a simple logic model.

The Model suggests that a work is planned through resources/inputs and a number of activities from which outputs and outcomes are expected to make impacts.

Systems Thinking Theory

In an era of increasing complexity (changes in the job market, technological innovation, globalisation, etc.), there is a need for organisational leaders to shift away from traditional (almost linear) leading models for innovative models that take uncertainties into account.

Leading in an increasingly uncertain and complex environment has prompted scholars to advocate for new

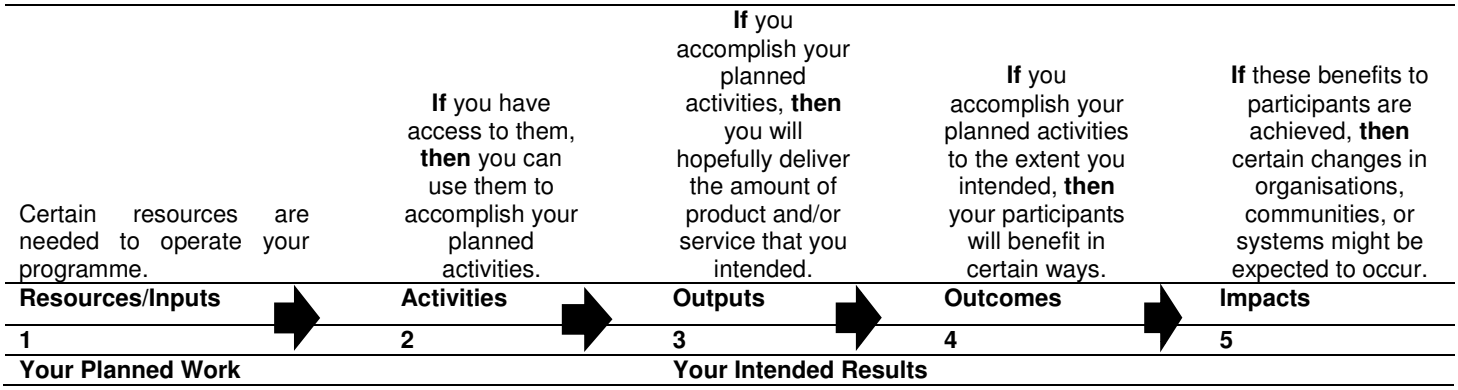


Figure 2. A Simple Logic Model (W. K. Kellogg Foundation, 2004, cited by Rogers, 2008:33)

models of leadership informed by complexity theory and systems thinking toward the goal of moving leadership from a traditional bureaucratic model to a more adaptive model better suited for today's complex organizations (Marion & Uhl Bien, 2001; Senge, 1990; Uhl-Bien & Marion, 2009; Uhl-Bien et al., 2007; cited by Davis, 2014:2).

The systems thinking paradigm refers to systems thinking as a 'world view', that is, seeing things holistically and interconnectedly (Maani and Maharaj, 2001). "Systems thinking refers to an approach that views systems as wholes rather than compilations of individual components and allows one to see the inter-connectedness and interdependencies of agents within systems" (Davis, 2014:2). According to Richmond (1997a, cited by Maani and Maharaj, 2001:3), the following thinking skills are the 'operational' guide to thinking systematically:

i. **Dynamic thinking** which is a mental application of the behaviour, is over time graph. "The trajectory should thus have a historical segment, a current state and one or more future paths" (Richmond, 1997b:6).

ii. **System-as-Cause thinking** which enables the determination of plausible explanations for the behaviour patterns identified with dynamic thinking. "Relationships which are not under the control of decision-makers within a system should be eliminated from consideration" (Richmond, 1997c:6).

iii. **Forest thinking** which is seeing the 'big picture'. The picture shows "the system of relationships that link the component parts"(Richmond, 1997d:6).

iv. **Operational thinking** which attempts to determine causality,looks at the structure of 'physics' of relationships, how one variable affects another, not just that they affect each other (Richmond, 1998a, cited by Maani and Maharaj, 2001:3).

v. **Closed-loop** thinking which helps to identify the principle of closed-loop structure, and enables a person to see that causality does not run in just one direction, but rather that an 'effect' usually feeds back to influence one or more of the 'causes', and that the 'causes' themselves affect each other (Richmond, 1997a, cited by Maani and Maharaj, 2001:3).

Some evaluation studies aim at having a holistic perspective on the programme being evaluated and its functioning. To that end, a systemic approach is crucial if as many aspects of the programme as possible are to be captured. This makes Systems Thinking one of the theories of interest for such studies.

The Major Programme Evaluation Models/Methods

The thematic categorisation of evaluation types corresponds in the literature to the various evaluation "models" (some prefer the term "methods"). This is of prime importance in the effort to identify an evaluation framework for any evaluation study. Discussions concerning the definition, dimensions, goals, approaches, purposes, and types of evaluation (see section 3.1. above) may be helpful. Still, the crux of an evaluation "framework"[…] lies in determining, as far as it is feasible, the respective relevance of these models to evaluating a particular programme (see Zinovieff and Rotem, 2008:11). There exists a myriad of such models.Bramley (1991), Worthen & Sanders (1987), cited by Eseryel (2002), identify six of such general approaches in a recent article on"approaches to evaluation":

i) Goal-based evaluation which begins with goals in mind and seeks to determine if those goals were achieved;

ii) Goal-free evaluation which does not seek to confirm or deny a pre-determined outcome or goal, but which rather aims to discover any benefits that result from the intervention;

iii) Responsive evaluation is an approach based on client requirements.Though a common approach, it can present unique challenges for the evaluator;

iv) Systems evaluation, which focuses on whether the intervention was efficient and effective;

Table 2. Summary of the Kirkpatrick's 4 levels and the questions the evaluation needs to address

Level		Measurement focus	Questions addressed
1-	Reaction	Trainees' perceptions	What did (do) trainees think of the training?
2-	Learning	Knowledge/Skills gained	Was there an increase in knowledge or skill level?
3-	Behaviour	Worksite implementation	Is new knowledge/skill being used on the job?
4-	Results	Impact on organization	What effect did the training have on the organisation?

Zinovieff and Rotem (2008:14) underline the following:

v) Professional review, which uses external expert appraisal to evaluate instead of other commonly used and accepted methods;
vi) Quasi-legal, which is infrequently practisedbut uses an actual court-of-inquiry format to present evidence, take testimonials, and evaluate an intervention or product.
Despite their existence in the literature, these models are not very frequently used in practice, and the literature seems to put more emphasis on goal or objective-based vs systems-based models, especially for the evaluation of training programmes (Phillips, 1991).

Goal-based Models of Programme Evaluation

Several goal-based models of programme evaluation have been discussed in the literature.

Donald Kirkpatrick's 4 levels

This Model is considered the "bible" of evaluation models. It is one of the most applied evaluation models developed by Donald J. Kirkpatrick (1959) in his "Evaluating Training Programs". He described 4 levels of training evaluation:
i) *Reaction* – a measure of satisfaction (what the trainees/ fellows thought and felt about the training); evaluation here focuses on the reaction of individuals to the training or other improvement intervention;
ii) *Learning* – a measure of learning (the resulting increase in knowledge or capability); evaluation here assesses what has been learned as measured with the end of course tests;
iii) *Behaviour* – a measure of behaviour change (extent of behaviour and capability improvement and implementation/application); evaluation here measures the transfer of what has been learned back to the workplace;
iv) *Results* – a measure of results (the effects on the institutional environment resulting from the fellows' performance); evaluation here tries to measure the impact of the training on overall organisational results (in the private sector on business results).

Comparing these levels of evaluation to the types of evaluation described in 3.1.5., one can observe that levels 1 (reaction) and 2 (learning) can be seen as part of formative evaluation. In contrast, levels 3 (behaviour) and 4 (results) are associated with summative evaluation. Attempts have been made to establish a level 5, measuring the impact at a societal level. In this case, levels 4 and 5 would correspond to normative and meta levels; an ideal 'full-scale'evaluation would have then been achieved. It should be noticed that the higher the level of evaluation, the more difficult it is to implement. Therefore, most organisations would limit themselves to lower levels [1 (reaction) and 2 (knowledge and skills), for example] evaluations. Still, these may give a deceptive sense of everything being on the right track. Forinstance, participants may express positive feelings about the programme without the latter or the individuals themselves improving in any aspect of their performance.

Level 1 produces what has been dubbed the "smile sheet", which measures how well the trainees like the training. Level 2 techniques are most used in the public sector/academic settings; they are most reliable when pre- and post-evaluations are utilised. The difficulty with level 3 is its possible subjectivity as it relates to human behaviour; this is why some believe level 4 evaluations may actually be easier to accomplish than level 3, since information at level 4 can be measurable. Some trainers believe that a positive level 3 evaluation implies success at level 4. Level 4 (the hardest to conduct) is the leastimplemented; it helps determine whether the programme has value. Depending on the information required, it may not be desirable, practical, or necessary to do all levels of evaluation. The questions the evaluation needs to address should be clearly identified so as to know which level(s) the investigation can limit itself to. Kirkpatrick (1959, cited by Zinovieff and Rotem, 2008:13) expressed this idea in Table 2:

Today, many evaluators point out that while the Kirkpatrick's Model is useful to evaluate a) whether learners liked their instruction, b) whether they learned something from it, and c) whether it had some positive effect onthe organisation, its weakness is that it cannot be used to determine the cost-benefit ratio of training (ROI). These modern evaluators have consequently recommended adding the so-called fifth level to Kirkpatrick's Model, at least for some programmes.

Table 3. Fictitious Kirkpatrick's 5th level

Level		Measurement focus	Questions addressed
1-	Impact	Impact at a societal level	What effect did the training have at the societal level?

A fictitious Kirkpatrick's 5th level would be tabulated as follows (Table 3):

Another limitation of the Kirkpatrick's Model is that all its four (or even five) levels of evaluation are not necessarily systematically conducted. Besides, a favourable level 1 evaluation (positive reaction from participants) may not have any direct relationship to learning, just like there is no such link between level 2 (learning) and level 3 (behaviour change). Different evaluation models/methods must, therefore, be considered with a view to opening larger evaluation angles.

Jack Phillips's Return on Investment (ROI)

Away from the previous perspective of training, which seeks to satisfy trainees, Jack Phillips's (2003, cited by Zinovieff and Rotem, 2008:14)' ROI perspective focuses on improving organisational performance. It is the positive impact of training on an organisation that is measured. However, this requires converting the level 4 data (results) into monetary values and comparing those results with the cost of the training programme, a constraint that the previous model does not carry.

Hamblin's 5 Levels

Hamblin (1974, cited by Zinovieff and Rotem, 2008:15) was one of the first to modify Kirkpatrick's Model. Both authors agree about the first three levels, but Hamblin splits the fourth level into two: organisation and ultimate value. The five-level model is therefore:

• *Level 1*: Reactions;
• *Level 2*: Learning;
• *Level 3*: Job behaviour;
• *Level 4*: Organisation – the effects on the organisation, from participant's job to performance changes;
•*Level 5*: Ultimate value – the financial effects, both on the organisation and the economy.

Guskey's critical levels

Thomas Guskey (2002) has also elaborated Kirkpatrick's 4 levels into 5; while developing his levels, he had students and educational environments in mind:

• Level 1: Participant reaction
- Purpose: to gauge the participants' reactions about information and basic human needs;
- Technique: usually a questionnaire;
- Key questions: was your time well spent? Was the presenter knowledgeable? Did participants like it? Did the material make sense? Will it be useful? Did the physical conditions of the activity support learning?

• Level 2: Participant learning
- Purpose: examine participants' level of attained learning;
- Technique: test, simulation, personal reflection, full-scale demonstration;
- Key question: did participants learn what was intended?

• Level 3: Organisational support and change
- Purpose: analyse organisational support for skills gained in staff development;
- Technique: minutes of district meetings, questionnaires, structured interviews or unobtrusive observations;
- Key questions: were problems addressed quickly and efficiently? Were sufficient resources made available, including time for reflection? What was the impact on the organisation? Did it affect organisational climate or procedures? Was implementation advocated, facilitated, and supported? Wasthe support public and overt? Were successes recognised and shared?

• Level 4: Participant use of new knowledge and skills
- Purpose: determine whether participants are using what they learned and using it well;
- Technique: questionnaires, structured interviews, oral or written personal reflections, examination of journals or portfolio, or direct observation;
- Key question: are participants implementing their skills, and to what degree?

• Level 5: Student learning outcomes
- Purpose: analyse the correlating student learning objectives;
- Technique: classroom grades, tests, direct observation;
- Key questions: what was the impact on students? Did it affect student performance or achievement? Did it influence students' physical or emotional well-being? Are students more confident as learners? Is student attendance improving? Are dropouts decreasing?

Indiana University (1996) Taxonomy

Indiana University's evaluation taxonomy includes six strata whichare not classified according to their importance. The first and last strata provide additions to Kirkpatrick's framework:

• *Stratum 1*: activity accounting – which examines training volume and level per participant
• *Stratum 2*: participant reactions

• *Stratum 3*: participant learning
• *Stratum 4*: transfer of training
• *Stratum 5*: business impact
• *Stratum 6*: social impact
The sixth stratum examines the impact of changed performance on society, and as such, is similar to Hamblin's ultimate (Level 5) value.

Industrial Society Stages

Founded in 1918, The Industrial Society (Work Foundation since 2002) is a British not-for-profit organisation and independent authority that provides advice, consultancy and research on the future of work, improving the quality of working life, leadership, economic and organisational effectiveness. It developed a six-stage circular model, which starts with a planning phase. The stages are:
• *Stage 1*: identify the business need;
• *Stage 2*: define the development objectives;
• *Stage 3*: design the learning process;
• *Stage 4*: experience the learning process;
• *Stage 5*: use and reinforce the learning;
• *Stage 6*: judge the benefits to the organisation (quality, customer satisfaction, and financial benefits).

Kearns and Miller KPMT Model

Many similarities exist between Kearns and Miller's KPMT (1997) Model and Phillips's (1991) work. They argue that clear objectives are an essential component of a training evaluation model. "Where they differ is in their aim to provide a sort of toolkit to help evaluators work through the process of identifying bottom-line objectives by means of questioning techniques, evaluating existing training, and using process mapping to identify the added value to organisations"(Zinovieff and Rotem, 2008:17). Just like several other models reviewed so far, the four-stage KPMT Model is much business inclined. The objectives are clarified from a business perspective rather than from the trainees. The evaluation levels, however, look very similar to Kirkpatrick's:
• Reaction to training and development;
• Learning;
• Transfer to the workplace/behaviour;
• Bottom line added value, measured in relation to the base level measures taken.
Kearns and Miller differ from some of the other models in that they believe that return on investment can only be looked at in hard terms, which confirms their over-business inclination.

Nine Outcomes Model

The "Nine Outcomes" Model aimsto measure whether training has been successful. In identifying the 9 outcomes, Donovan and Townsend (2004, cited by Zinovieff and Rotem, 2008:17) pose 9 questions with the training participants in mind:
• *Reaction* to training – did they like it?
• *Satisfaction* with the organisation of a training event (facilities, logistics, meals, etc.)
• *Knowledge* acquisition – did they learn anything?
• *Skills* improvement – can they do something new or better?
• *Attitude* shift – have they changed their opinions about something?
• *Behaviour* change – have they changed their way of doing things following the training?
• *Results* – how did the training impact on the organisation's key success factors?
• *Return on investment* – to what extent did the training give back more than it cost?
• *Psychological capital* – how did the training affect corporate image?
All 4 of Kirkpatrick's levels, incidentally, are included among the 9 outcomes. The Model therefore goes beyond Kirkpatrick's and includes ROI (Jack Phillips) and other aspects.

Organisational Elements Model

Kaufman and Keller (1994) argue that Kirkpatrick's Model was intended for evaluating training; they then suggest some framework modification because other types of development events need evaluation in organisations. Just like many other goal-based models went above the s fourth level to include a fifth and even a sixth level, they expanded Kirkpatrick's Model to include societal contribution as an evaluation criterion (for example societal consequences of manufacturing like pollution and safety) as well as some additions at the other levels, such as a review of the availability and quality of resources, etc. According to them, evaluation at all levels should be planned and designed prior to the implementation of any intervention. The six levels of the Model are the following:
• *Level 1: Input* – similar to Kirkpatrick's reaction level, but has been expanded to include the role, usefulness, appropriateness and contributions of the methods and resources used;
• *Level 2: Process* – this level also has similarities to the reaction level, but is expanded to include an analysis of whether the intervention was implemented properly in terms of achieving its objectives;
• *Level 3: Micro (acquisition)* – this is similar to the learning level and examines individual as well as small-group mastery and competence;
• *Level 4: Micro (performance)* – links closely to the behaviour level and examines the utilisation of skills and knowledge. The focus is on application rather than

transfer of skills and knowledge;
• *Level 5: Macro* – relates to the results level and examines organisational contributions and payoffs;
• *Level 6: Mega* – an additional level which looks at societal outcomes.

Systems-based Models of Programme Evaluation

Just like with goal-based models of evaluation, there exist several systems-based models, though the distinction between "goal-based"and "systems-based" models of evaluation is sometimes ambiguous.

Management-Oriented Evaluation [Context, Input, Process and Product (CIPP)] Model

Daniel Stufflebeam developed the CIPP Modelfrom 1971 onwards. Its original purpose was to "delineate, obtain, and provide useful information for judging decision alternatives" (Stufflebeam, 1971:267). The Model distinguishes four types of evaluation:
• *Context evaluation* – which helps in planning and developing objectives;
• *Input evaluation* – which helps to determine the design by examining capability, resources and different strategies;
• *Process evaluation* – which helps to control the operations by providing on-going feedback;
• *Product evaluation* – which helpsto judge and react to the programme attainments in terms of outputs and outcomes.
• Product evaluation,
The core parts of this Model are context, input, process, and product (CIPP) evaluation. The questions asked by each part are the following:
- What needs to be done?
-How should it be done?
- Is it being done?
- Did it succeed?
In this checklist, the "Did it succeed?" or product evaluation part is divided into four subparts, notably impact, effectiveness, sustainability, and transportability evaluations. The questions asked by each subpart are respectively the following:
- Were the right beneficiaries reached?
- Were their needs met?
- Were the gains for the beneficiaries sustained?
- Did the processes that produced the gains prove transportable and adaptable for effective use in other settings?
This Model allows focus on processes and not simply on outcomes.

IPO (Input, Process, Output)

This Model was developed by Bushnell in 1990. It focuses more on the inputs to training and helps to monitor progress by setting performance indicators at each stage. Its stages are:
• *Input* – such as the instructor experience, trainee qualifications, and resources;
• *Process* – the plan, design, development and delivery of the training;
• *Outputs* – the trainees' reactions, knowledge and skills gained and improved job performance;
• *Outcomes* – profits, customer satisfaction and productivity.

TVS (Training Valuation System)

The Training Valuation System (TVS) was developed by Fitz-enz in 1994. It is a four-step process similar to Kirkpatrick's framework at steps 3 and 4:
• *Step 1*: Situation analysis – this is similar to an in-depth training analysis. Like Kearns and Miller (1997), he suggests that the manager's answers are continuously probed until some visible, tangible outcome is revealed; besides, the questions should initially focus on the work process and not on the training;
• *Step 2*: Intervention – this involves diagnosing the problem and designing the training;
• *Step 3*: Impact – this examines the variables that impact on performance;
• *Step 4*: Value – this step places a monetary worth on the changed performance.

Pulley's Responsive Evaluation Model

This other system-based evaluation model focuses on the purpose of evaluation; it is known as the "responsive evaluation" model and was developed by Pulley in 1994. Responsive evaluation is a tool for communicating evaluation results more effectively by tailoring it to the needs of the decision-makers. Pulley argues that the objective of the evaluation should be to provide evidence for key decision-makers to determine what they want to know about the programme. The stages involved are:
• Identify the decision-makers so as to ascertain who will be using the information and what their stake in it is;
• Identify the information needs of the decision-makers – what do they need to know, and how will it influence their decisions?
• Systematically collect both quantitative and qualitative data;
• Translate the data into meaningful information;

• Involve and inform decision-makers on an on-going basis.
Clearly, using this evaluation model requires a commission from decision-makers which specifies their needs.

E-Learning Models

E-Learning models (2001) of evaluation are more recent. They were elaborated for evaluating "new technology delivery", as part of a series of training options available to human resources managers. They seem to have been designed specifically to evaluate technical and scientific achievements (de Moraes and Machado, 2017).

Duignan's Framework for Outcomes Systems

Paul Duignan (2004, cited by Zinovieff and Rotem, 2008:21) designed a very complex system which develops intervention or programme logics using the Outcome Hierarchies diagramming approach. Intervention logics set out the connections between the outcomes an individual, organisation, or group of organisations are trying to achieve and the steps, stages or intermediate outcomes which are needed to accomplishthis. It culminates in the identification of 7 high-level outcome attribution evaluation designs, as follows:

- True experiment design;
- Regression discontinuity design;
- Time series analysis design;
- Constructed comparison group design;
- Exhaustive causal identification and elimination design;
- Expert judgement design;
- Key informant judgement design.

United Way of America Outcomes Evaluation

Outcomes evaluation (which is also an evaluation type) can be less complex than in the Duignan's system. The UWA (1966) describes a step-by-step plan for this kind of evaluation. In this description, differences are established between the following:
• *Outputs* – which indicate little about changes in trainees and are usually just numbers;
• *Outcomes* – which indicate real changes in trainees;
• *Outcome targets* – which specify how much of one's outcome one hopes to achieve;
• *Outcome indicators* – which suggest progress towards the outcome targets.

Charities Evaluation Service Outcomes Triangle (2003)

The CES (2003) defines outcomes as the effects of activities, the changes, benefits or learning that occur as a result of work carried out. Outcomes are considered different from outputs or user satisfaction. Outcomes are the detailed activities, services and products of organisations, whereas outputs or user satisfaction usually involve asking clients/trainees what they think about different aspects of the services provided. The CES applies a triangular "Outcomes Learning Cycle" to help in the process of clarifying and measuring outputs, in analysing what can be learned from the outcomes achieved, and planning/implementing changes as a result of such learning.

Additional Evaluation Methods

The models/methods noted above do by no means exhaust the field. Several other methods or approaches exist which are not considered and discussed in this review of the literature, mainly because they are already included in many aspects, in the ones discussed. A few of these are: Dixon's (1996) six steps, Brinkerhoff's (1987) six stages, Bramley's (1996) goal-based, Wade's (1996) high impact, Shapiro's(2001) matrix, Pershing's (2000) perspectives, Warr, Bird and Rackham's (1970) CIRO (Context, Input, Reaction, Outcome), Preskill and Torres' (1999) Evaluative Inquiry, Kraiger's (1993) Learning Outcomes, Kaplan and Norton's (1990) Balanced Scorecard, etc.

CONCLUSION

To undertake an evaluation, one must choose an appropriate "type", as identified in section 3.1.5. above, and then determine if anyone "model" or "method", or combination of such, best fits one's situation. It should also be reminded that goal-based evaluations evaluate the extent to which programmes are meeting predetermined goals or objectives (Zinovieff and Rotem, 2008:12), while systems-based models seem to be more useful in terms of thinking about the overall context and situation (Ibid., 2008:15); and that "the distinction between the 'goal-based' and 'systems-based' is sometimes ambiguous" (Ibid., 2008:22). This study set out to critically review the literature available on programme evaluation. It compiled and discussed the works of programme evaluation specialists. These works are the theories, approaches, and models/methods which are relevant enough to serve as a theoretical handbook to researchers in the emerging and fast-developing area of training programme evaluation.

REFERENCES

Al-Zahran, A. (2007). The consecutive conference interpreter as intercultural mediator: a cognitive-pragmatic approach to the interpreter's role. PhD Dissertation. European Studies Research Institute (ESRI) School of Languages, University of Salford, Salford, UK.

Bickman, L. (1987). The functions of program theory. New Directions for Evaluation, 33, 5-18.

Davis, A.P. (2014). A Conceptual Model of Systems Thinking Leadership in Community Colleges. [Online]. (URL http:// journals.isss.org/index.php/proceedings58th/article/viewFile/2 256/758). (Accessed 9 March 2017).

De Moraes, R.M., & Machado, L.S. (2017). Continuous Evaluation of Training Systems Based on Virtual Reality. [Online]. (URL https://www.researchgate.net/publication/228957347_Continuous_ Evaluation_in_Training_Systems_Based_on_Virtual_Reality_Usi ng_ Fuzzy_Rule_Based_Expert_Systems). (Accessed 25 March 2018).

Dessinger, J., & Morley, J. (2003). Full-Scope Evaluation: Raising the Bar. [Online]. (URL http://catalogimages.wiley.com/images/db/ pdf/0787965006.excerpt. pdf). (Accessed 19 March 2018).

Dongho, J.R. (2015). Needs Analysis for Translator Education in Cameroon: A Case Study of ASTI, University of Buea. Ph.D. Thesis. ASTI, University of Buea.

Duignan, P. (2001). Introduction to Strategic Evaluation: Section on Approaches, Purposes, Methods and Designs. [Online]. (URL www.strategicevaluation.info/se/documents/104f.html). (Accessed 17 January 2017).

Eseryel, D. (2002). Approaches to Evaluation of Training: Theory and Practice. [Online]. (URL http://www.ifets.info/journals/5_2/esery el.html). (Accessed 2 February 2017).

Fiola, M. A. (2003). Prolégomènes à une didactique de la traduction professionnelle. [Online]. (URL http://id.erudit.org/iderudit/00 7594ar). (Accessed 19 January 2017).

Fitz-enz, J. (1994). Yes, you can weigh training value. Dallas (Texas): McGraw Hill Higher Education.

Gile, D. (1995a). Basic Concepts and Models for Interpreter and Translator Training. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Gile, D. (2009). Basic concepts and models for interpreter and translator training. Rev. ed. The Netherlands: John Benjamins Publishing Co.

Holden, D.J., Zimmerman, M.A. (2008). A Practical Guide to Program Evaluation Planning: Theory and Case Examples. London: Sage Publications, Inc.

Kearns, P., & Miller, T. (1997). Measuring the impact of training and development of the bottom line. London: Pittman Publishing.

Kelly, D., & Martin, A. (2009). Training and Education. In Baker, M. & Saldanha, G. (2009), eds. Routledge Encyclopedia of Translation Studies. 2nd ed. London and New York: Routledge Taylor and Francis Group, pp. 294-299.

Kirkpatrick, D. (1996). Evaluation. In: Craig, R., ed. The ASTD Training and Development Handbook: A Guide to Human Resource Development. 4th Edition. New York: McGraw-Hill. Ch.14.

Maani, K.E., & Maharaj, V. (2001). Systemic Thinking and Complex Problem Solving: A Theory Building Empirical Study. [Online]. (URL http://www.systemdynamics.org/conferences/2001/papers/Maani_ 1. pdf). (Accessed 9 March 2017).

Merriam-Webster (2000). Merriam-Webster Unabridged Dictionary. Electronic Version 2.5.

Motale, B.T.R. (2018). Vers une autonomisation du processus d'apprentissage de l'interprétation au Cameroun. M.A. Thesis. ASTI, University of Buea.

Olivia (2011). Difference between curriculum and program. [Online].(URL http://www.differencebetween.com/difference-between-curriculumand-vs-program/). (Accessed 28 February 2016).

Phillips, J. J. (1991). Handbook of training evaluation and measurement methods. (2nd ed.). Houston, TX: Gulf.

Pulley, M.L. (1994). Navigating the Evaluation Rapids. [Online]. (URLhttps://www.questia.com/magazine/1G1-16309518/navigating- the evaluation-rapids). (Accessed 1 November 2018).

Reynolds, A. J. (1998). Confirmatory program evaluation: A method for strengthening causal inference. American Journal of Evaluation, 19(2), 203-221.

Richmond, B. (1997b). Dynamic Thinking: A Behavioral Context. The Systems Thinker, 8(6).

Richmond, B. (1997c). System-as-Cause Thinking. The Systems Thinker, 8(8), 6-7.

Richmond, B. (1997d). Forest Thinking. The Systems Thinker, 9(2), 6-7.

Rogers, P. J., Petrosino, A., Huebner, T. A., & Hacsi, T. A. (2000). Program theory evaluation: Practice, promise, and problems. New Directions for Evaluation, 87, 5-13.

Rogers, P.J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. [Online]. (URL http://journals.sagepub.com/doi/abs/10.1177/1356389007084674. (Accessed 9 March 2017).

Sandrelli, A. (2005). Designing CAIT (Computer-Assisted Interpreter Training) Tools: Black Box. [Online]. (URL http://eurocon-ferences.info/proceedings/2005_Proceedings/2005_San drelli_Annalisa.pdf). (Accessed 1 January 2017).

Sharpe, G. (2011). A Review of Program Theory and Theory-BasedEvaluations. [Online]. (URLhttp://www.aijcrnet.com/journals/ Vol_1_No_3_November_2011/10. pdf). (Accessed 9 March 2017).

Ulrych, M. (2005). Training translators: Programmes, curricula, practices. Chapter 1. In: Tennent, M., (Ed.) Training for the New Millennium:Pedagogies for Translation and Interpreting. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Wolf-Branigin, M. (2013). Using Complexity Theory for Research and Program Evaluation (Pocket Guide to Social Work Research Methods): New York: Oxford University Press.

Zinovieff, M. A., & Rotem, A. (2008). Review and analysis of training impact evaluation methods, and proposed measures to support a UN system fellowships evaluation framework. Geneva: WHO's Department of Human Resources for Health. [Online]. (URL http://docplayer.net/4653283-Review-and-analysis-of-training impact-evaluation-methods-and-proposed- measures-to-support-a united-nations-system-fellowships-evaluation-f ramework.html). (Accessed 5 February 2016).