DEVELOPING A SYSTEMIC PROGRAM EVALUATION METHODOLOGY: A CRITICAL SYSTEMS PERSPECTIVE

Maria Alejandra Torres Cuello

ABSTRACT

In recent years there has been an increased interest within the program evaluation field for introducing systems thinking concepts in it. However, most of these attempts have been primarily directed towards supporting the practice of evaluation and not towards making theoretical advancements. This article is focused on introducing systems thinking concepts at a theoretical level, particularly those related to boundary critique in the program evaluation field by reframing the Fourth Generation evaluation methodology. I will introduce the general ideas for carrying out such reframing as well as describing the major changes produced in the methodology and how the introduction of these concepts may be beneficial for conducting an evaluative process.

Keywords: Systems thinking, program evaluation, boundary critique, Fourth Generation evaluation.

INTRODUCTION

In the American Evaluation Association Conference held in 2017, Michael Quinn Patton along with other panelists, addressed the need for leaders and evaluators of developing a critical evaluation thinking in order to achieve the United Nations Sustainable Development Goals (SDG). However, this need was not only recognized to be applicable to the SDG´s but to the practice of evaluation in other fields as well. By these means, the panelists recognized the need of thinking outside the box, joining leadership and evaluation, and the demand for capacity building in trading for evaluation building across the globe by posing the following question: How do you use evaluation to improve the life of people? In responding such question, they realized the importance of interconnectedness. In this discussion, Patton identified a series of leadership qualities that should be reinforced: think beyond projects, think beyond nation-state borders, think beyond siloes when referring to the SDG´s, to move beyond top-down vs bottom-up to dynamic local-global interconnections, and to move from the independence to interdependences. This represents the need of a transformational change in leadership and evaluation in which systems thinking and particularly critical systems thinking can play a fundamental role for developing the current evaluation agenda.

When exploring what has been done in introducing systems thinking concepts in evaluation, and particularly in program evaluation, advancements can still be made. Although interest in systems thinking has been raised in recent years, most of it has been directed towards supporting the practice of evaluation and not towards producing theoretical advancements. The practical side of the introduction of systems thinking within the program evaluation field

is exemplified by the work of Burke, 2006; Dick, 2000; Sharma & Gutierrez, 2004 and Williams & Hummelbrunner, 2011. On the other hand, the theoretical development of the field is exemplified by the work of Cabrera *et al.* (2008) and Imam *et al.* (2006) in identifying a series of concepts related to interrelationships, perspectives, boundaries, and the system on its own (a concept introduced only by Cabrera *et al.* (2008)) as key concepts of the systems thinking field that could be useful for evaluators. However, aside of this recognition no further work has been conducted in exemplifying ways by which such elements could be introduced within the field a critique has been directed towards attempting to unify the systems thinking field by means of these concepts (Nowell, 2008; Midgley, 2008).

Furthermore, although several evaluators claim to be using or to have used systems thinking ideas and methods in evaluation, the extent and ways in which evaluators are drawing on these ideas is not well understood or operationalized (Gates, 2016; Reynolds et al, 2012; Walton, 2016). To date, there has been no systematic examination of how these ideas and methods contribute to the [program] evaluation field or a framework for understanding when and why to use them (Gates, 2016). Moreover, the way in which evaluation theory and practice operate is a concern as well. Recent research on evaluation illustrates that, despite many noteworthy efforts to integrate evaluation's theory into its practice and its practice into its theory, more often than not, the gap between theory and practice remains a chasm (Christie, 2003; Christie, Quinones, & Fierro, 2014; Miller & Campbell, 2006).

This suggests the presence of a gap regarding the use of systems thinking in the evaluation field, not because attempts of producing such inclusion have not been done but because they are perceived as incomplete or not completely clear. There is a need then of not only incorporating systems thinking at a theoretical level in evaluation but also on detailing how such amalgamation works, why and how is it useful and how they will both work in practice.

THEORETICAL REVIEW

The theoretical review associated to the development of the methodology described in this paper, involved the review of the systems thinking field, the program evaluation field, and the work that has been carried out in relating both of them. Each one of them will be briefly described in what follows.

Systems thinking

According to Jackson (1991) and Midgley (2000), the development of this field could be described as taking place in three distinctive waves for which the system concept defined by Angyal (1941) as well as the work of Von Berthalanffy (1950, 1956) and Boulding (1956) on the General Systems Theory constitute a foundation. The first wave is characterized by having an objective character in which the achievement of goals is the main aim, the subject-object duality is the founding concept, and the system is understood in terms of component

parts and not wholes for which it can be shaped by means of a mathematical model. The second wave emerged as a result of the critique of authors such as Ackoff (1979), Checkland (1985), and Churchman (1970) of the underlying concepts of the first wave. This critique was particularly directed towards the objective structure of the wave, the way in which the system concept was conceived, and how being oriented towards goal achievement restricted the type of situations that could be tackled by means of the approaches developed in the first wave. Taking this into account, the second wave of systems thinking is characterized by having a subjective character grounded on interpretivism for which it is directed towards ill structured situations that are modelled by the different perspectives of those who are viewing them. In this conception the subject-object duality that was once the first wave foundation, is avoided and is replaced by the idea of the existence of multiple perspectives by which systems are structured and for which multiple realities exist. Just like the second wave emerged, the emergence of the third wave took place by sharing the critique of the second wave towards the first one, but also by critiquing the second wave for not being radical enough in its attempts to generate change, instead it was seen as reinforcing the status quo of the system (Mingers, 1980; Jackson, 1985, 1987, 2000, 2003).

The third wave development, took place in a different way when compared with the first two waves as it was developed in two parallel strands oriented towards power (Ulrich, 1983) and pluralism (Flood & Jackson, 1991; Schechter, 1991; Midgley 1996, 1998) and which initially had a competitive character but were later reconciled in the work of Midgley (1998, 2000) in developing systemic intervention. The pluralism oriented strand has at its core three themes, critical awareness, improvement, and methodological pluralism. According to Midgley (1998), critical awareness is directed towards reflecting over the taken for granted assumptions as well as over the conditions that gave rise to them. Improvement, is defined temporarily and locally, but in a widely informed manner, taking issues of power (which may affect the definition) into account. Finally, methodological pluralism refers to the use of a variety of research methods in a theoretically coherent manner for addressing a variety of issues. On the other hand, the power oriented strand is grounded on two concepts, boundary judgments and boundary critique. A boundary judgment defines the lenses by which one sees the world. In terms of Churchman (1971), boundary judgments define what is to be included in analyses, and whose views have credibility. Taking into account the definition of a boundary judgment, boundary critique is directed towards carrying out a systematic effort of handling boundary judgments critically, a process in which one is self-critical and critical towards external things as well (Ulrich & Reynolds, 2010). Midgley (2000) extends the work on boundary critique by introducing an additional concept of marginalization. According to Midgley (2000), when setting boundaries one must reflect not only over what is included in it but also over what has been left out, for which choosing a particular boundary will marginalize a series of elements that are left out of it. Then, one must reflect over the implications of such marginalization and over making a particular selection or another.

Developing a Systemic Program Evaluation

Program Evaluation

In order to talk about program evaluation, one must first need to define two key elements that are traditionally understood as being alike, but which I understand differently, these are evaluation and program evaluation concepts. The American Evaluation Association (AEA) defines evaluation as a systematic process to determine merit, worth, value or significance (2014). On the other hand, the definition of program evaluation considered in this research which was proposed by Rossi *et al.* (2004), refers to program evaluation as the use of social research methods to systematically investigate the effectiveness of social intervention programs in ways that are adapted to their political and organizational environments and are designed to inform social actions in ways that improve social conditions. The main difference between both concepts is that the former can be directed towards evaluating personnel, processes, or programs on their own. However, the latter is solely directed towards assessing programs in ways in which some sort of social improvement can be produced. The broader definition of evaluation does not involve the consideration of improvement or the generation of a social impact.

There are many other elements within the program evaluation field such as, the role of the evaluator, the purpose with which the evaluation is conducted, the methods used and the degree in which the stakeholders are involved within the evaluative process that are important when carrying out an evaluation. For this reason, all these traits are summed up in a series of approaches or models of evaluation that dictate a general standard by which evaluations can be carried out following different variations for each one of these traits. Utilization Focused Evaluation, Responsive Evaluation, Cost-Effectiveness Analysis, and the Fourth Generation Evaluation are examples of evaluation models carried out with different purposes, involving stakeholders in different degrees among other features that make them appealing to evaluators. Fourth Generation evaluation will be the evaluation approach of interest for developing the methodology described in what follows.

Fourth Generation Evaluation is an evaluation model developed by Guba & Lincoln (1989) grounded on the constructivists paradigm and the Responsive evaluation characterized by having a participative character. This participation is grounded on the dialectic-hermeneutic circle for stakeholder in which their interaction is divided by groups for construction formation which are later improved and critiqued by other groups leading to a reframing process until consensus is achieved. Figure 1 presents the general methodology of the original Fourth Generation evaluation model.

Developing a Systemic Program Evaluation
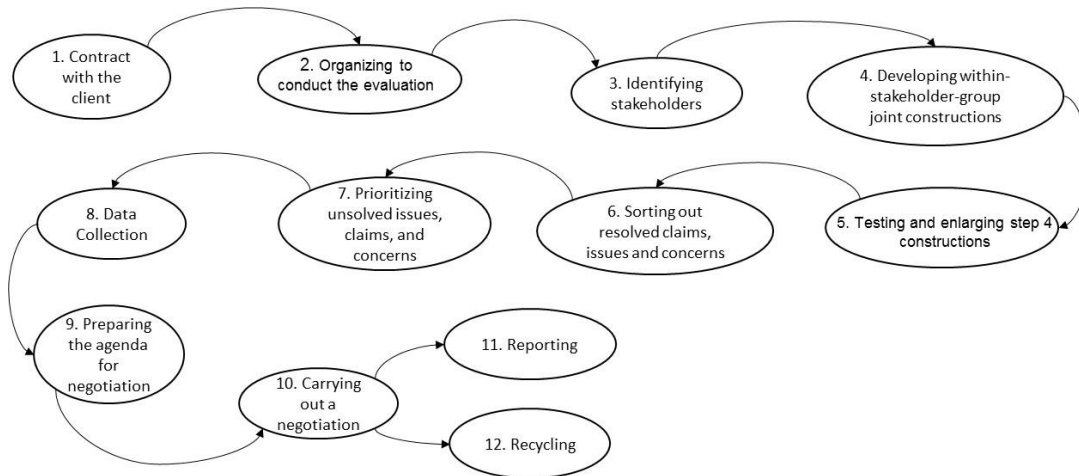
Figure 1. Original Fourth Generation Methodology.

The Fourth Generation methodology consists of 12 steps in which the evaluator can go back and forth when needed. The main characteristic of this form of evaluation is the use of dialogical processes and negotiations in several of its stages, as well as the development of constructions in order to include issues, claims, and concerns of the different stakeholders.

Program Evaluation and Systems Thinking

Throughout the evaluation field, problem situations have been recognized as simple, complicated, or complex (Glouberman & Zimmerman, 2002; Snyder, 2013; Rogers, 2008; Patton, 2011; Gregory & Jackson, 1992). The latter are the ones of particular interest for us. If the world is complex, the used evaluation theories and methods should mirror that complexity (Forss, Marra, & Schwartz, 2011). For this reason systems thinking and complexity science have been regarded as promising options to tackle complex situations in the evaluation field. Although there have been attempts at both theory (Imam *et al.*, 2006; and Cabrera *et al.*, 2008) and practice (see Burke, 2006; Williams & Hummelbrunner, 2011; Newman *et al.*, 2003; Frederick *et al.*, 2008; Martin & MacDonald, 2012; Derwisch & Lowe, 2015; Fitch, 2006; Hart & Paucar-Caceres, 2017; Reynolds, 2006; Ulrich & Reynolds, 2010; and Gregory, 1997) of exploring how systems thinking concepts could serve program evaluation, to date, there has not been a systematic examination of how these ideas and methods contribute to the evaluation field or a framework for understanding when and why to use them (Gates, 2016). Even though the connection of systems thinking and evaluation in practice has not been deeply explored, benefits of using systems thinking and/or complexity science in evaluation have been recognized. However, this paper will focus only on exploring the introduction of systems thinking concepts in evaluation.

According to Imam *et al.* (2006), Reynolds (2007), Hummelbrunner (2011), Reynolds *et al.* (2012), and Williams & van´t Hof (2014) there are three concepts of systems thinking evaluators need to know. They correspond to perspectives, boundaries and interrelationships. Cabrera *et al.* (2008) extends them with a fourth concept represented by the system itself.

Each one of these concepts relate to the other in a particular kind of way so that they cannot be taken separately. According to Cabrera *et al.* (2008), distinctions refer to the existence of concepts that could be differentiated, meaning that in order to make a distinction one must establish an identity and exclude another. Distinctions relate to what Imam *et al.* (2006) denote as boundaries given that making distinctions involves a boundary that differentiates between what/who is inside from that which is outside the boundary that is being set. Although a distinction is made between two or more elements, there still exists a relationship between them, bringing to the fore the consideration of a second pattern: relationships. In Cabrera *et al.* (2008) terms, if there is any distinction between two concepts or agents, there must be some concept of relationship between them. Relationships have as direct implication the emergence of affecting and effecting elements. Affecting refers to the action taken by an agent and effect refers to the result of that action on or to another entity. The sole existence of relationships between different concepts, increases the interconnectedness among the elements of a system so that its internal behavior gets to be reinforced. In that sense, once again a key pattern emerges, in the form of system, as the simple form of system is one considered as a whole made up of at least two related parts. By having distinctions and the emergent relationships between them, a system emerges as well. Finally, the perspective pattern emerges, this pattern is implicit in the consideration of a concept as it refers to a particular way of framing it. Perspectives relate in a particular way to the other three patterns considered by Cabrera *et al.* (2008) as they have the potential to instantly transform whole systems, rearrange distinctions, and cause relationships to appear or disappear.

THE PROPOSED REFRAMED METHODOLOGY

Taking into account the methodology proposed by Guba & Lincoln (1989) and the systems thinking concepts expressed in previous sections, I will now proceed to describing how the Fourth Generation evaluation methodology is reframed and concepts like boundary critique are introduced in doing so.

General characteristics

One can reframe Fourth Generation evaluation by identifying the presence of the concepts proposed by Cabrera *et al.* (2008) and Imam *et al.* (2006) in different ways; multiple perspectives and boundaries represented by the conceptions of the stakeholders involved within the evaluation stages and the way in which they frame the situation, the interrelationships not only between the stakeholders but also between the issues, claims, and concerns that interest them, and the system as a whole by understanding and taking into account the contextual characteristics surrounding the evaluand and the evaluation, and which are at the same time affecting them and vice versa. These elements will be explained in more detail in what follows.

Developing a Systemic Program Evaluation

Given that Fourth Generation evaluation is based on responsive evaluation as well as constructivism, as Guba & Lincoln (1989) express, one of its key components is the need to conduct program evaluation by taking into account the perspectives of those who have a stake in it. This means that it does not prioritizes the way the client of the evaluation or the evaluation team seek to orient the evaluation. Instead, it places all the stakeholders in a similar position regarding the ability of designing the evaluation, based on the issues, claims, and concerns that all the involved consider as being relevant to be evaluated. An important element of this multiplicity of perspectives is that it not only eases the way in which the stakeholders can express their view points, by means of making constructions, but also by giving the same relevance to every construction. In that sense, prioritization of issues, concerns, and claims is not based on who makes them but on the direct implications of choosing one or another and on the results that can be achieved by doing so. By choosing particular viewpoints to be addressed, a boundary process takes place, as the evaluation starts to be bounded by the considerations of those involved in it. However, the key element that should be explicit is the fact that the consideration of multiple perspectives, does not only implies different ways of framing situations, but also on considering that the background from which these perspectives come have different elements that influence and shape them such as values, ideas, and beliefs. At this point is where perspectives and boundaries are particularly intertwined.

On the other hand, another important element is the role of the evaluator as a facilitator throughout different stages of the evaluative process that involve the participation of different individuals and stakeholder groups. When playing the role of a facilitator, Guba & Lincoln (1989) recognize the evaluator as having no special license, elite status, or superior power, nor is he warranted in exerting any special control, that is as being neutral. According to Rifkin *et al.* (1991), neutrality could be identified in two senses: first, impartiality, that is, ensuring against bias; and second, "equidistance", which is to say ensuring that the parties participate "equally" within the process. However, taking into account such definition, the role of the evaluator could not be neutral given that even if the evaluator was in a different position within the evaluative process, it will always be biased in some sense, in other words it will always have a partial character. Given that partiality is understood as being grounded on biases, the stakeholders´ understanding, including the evaluator, will be partial as they will be biases by their own interests, goals, values, beliefs, perspectives, and distinctions they can make. Such biases will also lead to the formation of different boundaries. Biases will be explored in further chapters of this research. On the other hand, the evaluator cannot be neutral in terms of ensuring an exactly equal participation of the parties involved in the evaluative process. Although one of the key elements of both, the original and the reframed form of Fourth Generation evaluation is the stakeholders´ inclusion, this does not guarantee equality in their participation. Setting the adequate conditions for their participation does not guarantee it either, given that once again the issues which condition the human perception

and behaviour come to the fore in the way in which they participate and understand situations as well.

Guba & Lincoln (1989), shift from a realist ontology to a relativist one. In their conception, there is no reality except that created by people as they attempt to make sense of their surrounds. At the same time, they also conceive truth as a matter of consensus among informed and sophisticated constructors, not of correspondence with an objective reality. In order to develop my proposal, I draw upon Guba & Lincoln at a methodological level given that my work has an epistemological character. This does not imply the adoption of a relativist reality, the rejection of the possibility of an ultimate reality, or even embracing solipsism. In order to take advantage of Guba & Lincoln ideas on Fourth Generation evaluation, one does not need to adopt a particular ontological position.

Reframed Methodology in Detail
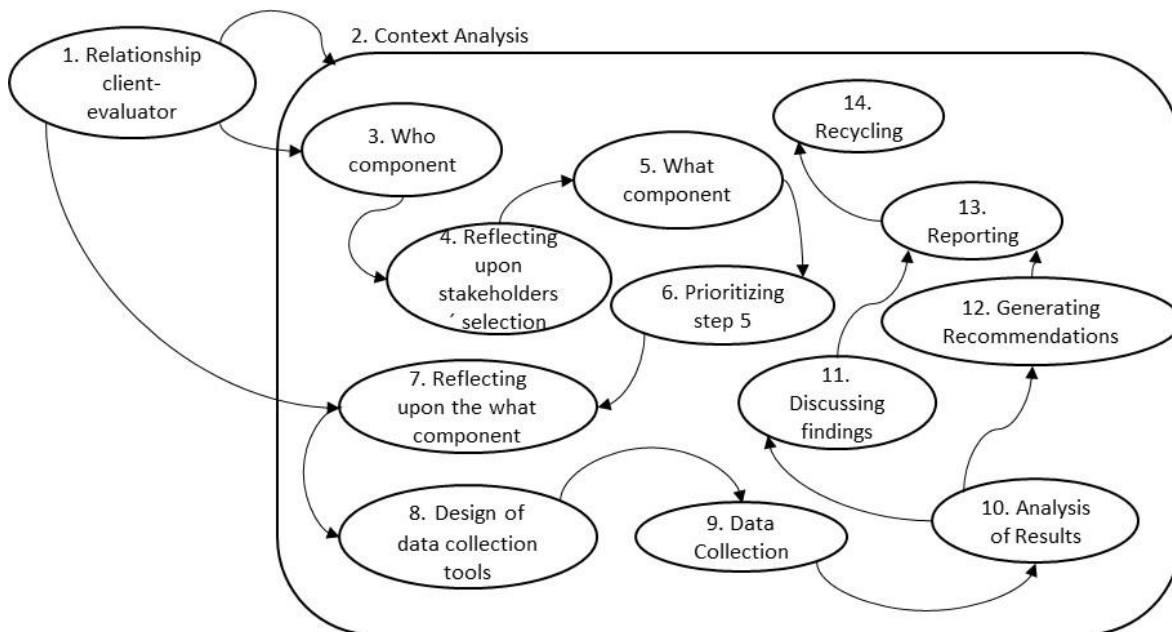
Figure 2. Proposed Reframed Methodology.

Figure 1 and Figure 2 show the original methodology and the reframed one. With regards to the underlying conceptions of the original methodology, boundary critique becomes a fundamental element in identifying the "what" and the "who" components of the evaluative process and in framing the evaluation as a whole. The rationality followed by boundary critique is grounded on the fact that throughout the evaluative process the establishment of boundaries takes place several times, for which it becomes very important not only to recognize such boundaries but to be able to reflect over the implications of choosing and setting one or another and also on recognizing what is left outside of them. Steps 4 and 7 of the evaluative process shown in Figure 2 are explicit means by which reflexivity needed for boundary critique takes place. Being able to introduce boundary critique in the evaluation

has as a direct implication raising awareness about the decision making processes of the evaluation design.

Guba & Lincoln (1989) recognize in their original development of Fourth Generation evaluation the importance of considering the stakeholders´ interests, values, ideas, among other characteristics, as they are the ones shaping the issues, claims, and concerns of interest of those involved within the evaluative process. However, the way in which the authors propose to conduct the methodological development of Fourth Generation evaluation, does not explore how one individual´s ideas, interests, values, and other characteristics may influence other individual´s perceptions and how they may shift them. Instead, they focus on the individual contributions in the formation of constructions, given that the participation of multiple stakeholders does not necessarily imply an interaction between them. The proposed reframed methodology, shifts from this conception towards one in which construction formation takes place simultaneously not only by involving different individuals but different stakeholder groups as well. Hence, the reframed proposed methodology implies a constant reflective process of setting boundaries around the different elements of its design and implementation, shifting between them and reshaping them until a particular boundary is reached. Such boundary will be the one that ultimately will shape the way in which the evaluation will take place in practice. It also becomes important not only to reflect over the implications of setting a particular boundary but also over which are the assumptions, conceptions, and values that lead to the formation of particular boundaries. This is a process which at the moment does not take place explicitly in the Fourth Generation Evaluation approach.

An example of such elements are the "what" and the "who" that should be considered in conducting the evaluation. This reflective process implies the inclusion of new information which has the possibility of modifying the existing boundary, as well as putting the initial boundaries under the scope so that they can be placed in a position where they are critiqued. However, when using the term critique, this does not mean that they are targeted in a negative way so that their flaws might be put forward. Instead critique in a systemic sense implies that none of the constructions are taken for granted or assumed as an unalterable truth, so that they can be enriched and modified by putting them under a different scope from which they were originally conceived. On the other hand, according to Ulrich (1983) being critical means to become self-reflective in respect to the presuppositions flowing into one´s judgments, both in the search for true knowledge and rational action. This means that being critical in the reframed form of Fourth Generation evaluation is not only oriented outwards but inwards as well. This means that it is equally important and valuable to question not only other people assumptions but also one´s own assumptions. This would lead not only to the possibility of modifying other stakeholder boundary judgments but for reframing the own as well, as well as to trying to understand how and by what means they have been formed.

Developing a Systemic Program Evaluation

Ulrich (1988) highlights the importance of considering boundary judgments in evaluation. According to the author, evaluation shows an inability to act effectively with regards to value judgments particularly because of their uncritical acceptance of the prevailing objectivism of science upon which most evaluation approaches rest upon. Ulrich (1988) proposes the critical use of systems ideas in order to tackle this issue. When using system concepts in any real world issue or situation such as public policy and evaluation, a priori assumptions about what belongs to the systems and its environment, are made. These assumptions are defined by Ulrich (1988) as boundary judgments. As I have stated earlier, boundary judgments as well as Fourth Generation evaluation, have a subjective character. This means that there is no right or wrong way of bounding but different perspectives in which this process can take place. A particular way of framing determines further ways in which problems and improvements are conceived, different boundaries bring to the fore different ways of appreciating improvement and defining problem situations. Boundaries are not only considered in framing the system but in considering the problem context as well and the consequences of bounding it in one way or another. This is another element that is being introduced in reframing Fourth Generation Evaluation under the scope of system concepts and particularly those related to boundary judgments and boundary critique. Guba & Lincoln propose four different purposes with which an evaluation could be carried out. However, two of those purposes do not take into account the contextual elements surrounding the evaluand and the evaluation. As such, those evaluation purposes should not be considered when conducting a reframed Fourth Generation methodology as this would dismiss the systemic character of the proposed methodology given that they do not take into account the evaluand and the evaluation context for which they do not explore the relationship between wider systems [contextual framing] and the sub systems [evaluand and evaluation]. The interaction between the wider systems and the sub-system will be determined by the way in which the former is framed; given that it could be framed in multiple ways, one can have multiple wider systems to explore.

Another important issue regarding boundaries, is the concept of marginalization. It is well known that when a boundary is chosen, elements outside of it are marginalized, however the way in which this takes place when carrying out an evaluative process acquires a particular connotation. When issues, claims, and concerns are marginalized within the reframed evaluative process, this does not mean that they are completely left outside the evaluative process. Instead, this means that they are left in a position in which they can be re-addressed for future stages of the same evaluative process or in future evaluation. This can happen particularly because even though they may not be useful at the present moment, they may become useful in the future as the evaluation evolves. However, it is important to understand that the shifting character of an issue, claim or concerns needs to be taken into account. This means that although a particular issue will be taken into account at different points in time, its behaviour or characteristics may be different.

Another important element within the systemic character of this evaluative process is the relevance given to the contextual elements of the evaluand and the evaluation. The contextual

elements are an important factor for both the evaluand and the evaluation as they shape the performance of the former and the structure of the latter. For this reason, as can be seen in Figure 2 the contextual factors do not only need to be identified throughout the evaluative process, but they also need to be considered for comprehending and analysing the operation of the evaluand, and for customizing the evaluative process. Within the original design of Fourth Generation methodology, the identification of contextual elements takes place within step 2 of the evaluation. However, it is taken as having a static character with regards to how it takes place and the usefulness of carrying out such task.

Generating recommendations was another step that was not defined in the original methodology. However, it is a key element of the reframed Fourth Generation methodology as it is a stage I attempt to produce by following a participatory approach by involving the most relevant stakeholders. By doing so, it would be more likely that the produced recommendations will be put into practice given that they affect each one of the stakeholders in a different way. In order to produce such recommendations a similar structure to the one in selecting the "who" and the "what" components of the evaluative process is followed.

CONCLUSIONS

As can be seen systems thinking elements are very useful for framing and conducting a program evaluation approaches. The introduction of systems thinking concepts in evaluation is useful for addressing situations more richly than it would be done without them given that by doing so there is an awareness about what is being included in the evaluation and why this is being done. Such awareness, gives those involved in the evaluation planning the possibility of reflecting over the implications of addressing the evaluative process in one way or another before conducting it in practice. By addressing such implications before the implementation of the evaluation, they have the possibility of modifying the parameters over which the evaluation is being grounded, the way in which the evaluation is being conducted, as well as the boundaries used to frame its different components. As can be seen, different systems thinking concepts are helpful not only for theoretically developing program evaluation but also to put such developments into practice, given that their design and propositions are related to aspects that may influence the evaluation in real life. Probably the most important benefit of introducing systems thinking concepts in reframing Fourth Generation Evaluation is that it produces a shift in the way the evaluand and the evaluative process are conceived as the interconnection between the elements within and outside the program are realized and attempts to understand them as well as their implications are conducted, and also by realizing that such consideration has an impact on the way in which the evaluation is planned and conducted as well.

The main benefit of using boundaries in an explicit manner is that at plain sign the stages of the evaluation seem to be easily conducted and consensus presents itself as easily achievable.

However, there are several processes that need to be carried out and taken into account in order to reach this point. It is of particular importance realizing that although one might count with few or several stakeholders, the main difficulty of conducting participatory evaluations that give a more active role to the stakeholders is that it is difficult not only to make them aware of what matters to them, but also on how what matters to one or another can be seen as equally useful to carry out an evaluation. Another important element, is that by using boundary critique the evaluator can be aware of the decisions made throughout the evaluation design and implementation, as they not only shape the way in which the evaluation was conducted but also how results could be interpreted and recommendations could be produced. In that sense, since the conception, the entire evaluation process is interrelated and affected by the decisions made throughout the evaluation.

REFERENCES

Ackoff, R. (1979). Resurrecting the future of operational research. *Journal of the Operational Research Society Vol 30*, 189-199.

Angyal, A. (1941). A Logic of Systems. In F. E. Emery, *Systems Thinking* (pp. 17-29). Suffolk: Penguin Books.

American Evaluation Association. (2014). *American Evaluation Association. What is Evaluation?* Retrieved from http://www.eval.org/p/bl/et/blogid=2&blogaid=4

Boulding, K. E. (1956). General System Theory - The skeleton of science. *Management Science*, 197-208.

Burke, D. D. (2006). System Dynamics-based Computer Simulations and Evaluation. In B. Williams, & I. Imam, *Systems Concepts in Evaluation* (p. 226). Inverness: EdgePress.

Cabrera, D., Colosi, L., & Lobdell, C. (2008). Systems thinking. *Evaluation and Program Planning*, 299-310.

Checkland, P. (1985). From Optimizing to Learning: A Development of Systems Thinking for the 1990s. *Journal of the Operational Research Society,* 757-767.

Christie, C. A. (2003). What guides evaluation? A study of how evaluation practice maps onto evaluation theory. *New Directions for Evaluation, Vol 97*, 7-36.

Christie, C. A., Quiñones, P., & Fierro, L. (2014). Informing the Discussion on Evaluator Training A Look at Evaluators' Course Taking and Professional Practice. *American Journal of Evaluation, Vol 35 (2)*, 274-290.

Churchman, C. W. (1970). Operations Research as a Profession. *Management Science*.

Derwisch, S., & Lowe, P. (2015). Systems Dynamics Modelling in Industrial Development Evaluation. *Institute of Development Studies Bulletin,* 44-57.

Developing a Systemic Program Evaluation

Dick, B. (2000). *Soft systems methodology*. Obtenido de Session 13 of Areol – action research and evaluation on line.: Available at http://www.scu.edu.au/schools/gcm/ar/areol/areol-session13.html

Fitch, D. (2006). A Cybernetic Evaluation of Organizational Information Systems. En B. Williams, & I. Imam, *Systems Concepts in Evaluation* (pág. 226). Inverness: EdgePress.

Flood, R., & Jackson, M. (1991). *Creative Problem Solving Total Systems Intervention.* Wiley & Sons.

Forss, K., Marra, M., & Schwartz, R. (2011). *Evaluating the Complex: Attribution, Contribution and Beyond.* Transaction Publishers.

Fredericks, K., Deegan, M., & Carman, J. (2008). Using Systems Dynamics as an Evaluation Tool. Experience from a Demonstration Program. *American Journal of Evaluation,* 251-267.

Gates, E. (2016). Making sense of the emerging conversation in evaluation about systems thinking and complexity science. *Evaluation and Program Planning*, 62-73.

Glouberman, S., & Zimmerman, B. (2001). *Complicated and Complex Systems. What would succesfull reform of Medicare look like?* Commission on the future of health care in Canada.

Gregory, A. (1997). Evaluation Practice and the Tricky Issue of Coercive Contexts. *Systems Practice*, 589-609.

Gregory, A., & Jackson, M. (1992). Evaluation Methodologies: A Systems for Use. *The Journal of the Operational Research Society*, 19-28.

Guba, E., & Lincoln, Y. (1989). *Fourth Generation Evaluation.* Newbury Park: Sage Publications.

Hart, D., & Paucar-Caceres, A. (2017). A utilisation focussed and viable systems approach for evaluating technology supported learning. *European Journal of Operational Research,* 626-641.

Hummelbrunner, R. (2011). Systems Thinking and Evaluation. *Evaluation, Vol 17(4)*, 395-403.

Imam, I., LaGoy, A., & Williams, B. (2006). Introduction. In B. Williams, & I. Imam, *Systems Concepts in Evaluation* (pp. 3-10). Inverness: EdgePress.

Jackson, M. (1987). New Directions in Management Science. In M. Jackson, & P. Keys, *New Directions in Management Science* (pp. 133-159). Aldershot : Gower.

Jackson, M. (2000). *Systems Approaches to Management.* New York : Plenum.

Jackson, M. (2003). *Systems Thinking: Creative Holism for Managers.* Chichester: John Wiley & Sons.

Developing a Systemic Program Evaluation

Jackson, M. (1985). Social Systems Theory and Practice: The Need for a Critical Approach. *International Journal of General Systems*.

Jackson, M. (1991). *Systems Methodology for the Management Sciences.* New York: Plenum Press.

Joint Committee on Standards for Educational Evaluation. (1994). *Program Evaluation Standards.* Sage.

Martin, E., & MacDonald, R. (2012). *A System Dynamics-Based Evaluation of the New York State HIV Testing Law.* New York: University of Albany.

467-478.

Midgley, G. (1996). What is this thing called CST? In R. Flood, & N. Romm, *Critical Systems Thinking* (pp. 11-24).

Midgley, G. (1998). The Theory and Practice of Boundary Critique: Developing Housing Services for Older People. *Journal of the Operational Research Society,* 467-478.

Midgley, G. (2000). *Systemic Intervention: Philosophy, Methodology and Practice.* London: Kluwer Academic/Plenum Publishers.

Midgley, G. (2008). Response to paper "Systems Thinking" by D. Cabrera et al.: The unification of systems thinking: Is there any gold at the end of the rainbow? *Evaluation and Program Planning,* 317-320.

Miller, R. L., & Campbell, R. (2006). Taking stock of Empowerment Evaluation an empirical review. *American Journal of Evaluation, Vol 27*, 296-319.

Mingers, J. C. (1980). Towards an appropriate social theory for applied systems thinking: critical theory and soft systems methodology. *Journal of Applied Systems Analysis*, 41-50.

Newman, J., Martin, L., Velasco, M., & Fantini, A. (2003). A Systems Dynamics Approach to Monitoring and Evaluation at the country level: An Application to the Evaluation of Malaria control Programs in Bolivia. *Fifth Biennial World Bank Conference on Evaluation and Development* (págs. 1-16). Washington : OED Conference.

Nowell, B. (2008). Response to paper "Systems Thinking" by D. Cabrera et al.: Conceptualizing systems thinking in evaluation. *Evaluation and Program Planning,* 329-331.

Patton, M. (2017). *Change Agents: Advancing the Global SDGs by Ensuring the Right to Evaluation by All.* Washington DC: American Evaluation Association Conference - Presidential Strand.

Patton, M. Q. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use.* Thousand Oaks: Sage.

Reynolds, M. (2006). Evaluation Based on Critical Systems Heuristics. In B. Williams, & I.

Imam, *Systems Concepts In Evaluation* (pp. 101-121). Inverness: EdgePress.

Reynolds, M., Forss, K., Hummelbrunner, R., Marra, M., & Perrin, B. (2012). Complexity, systems thinking and evaluation - an emerging relationship? *Evaluation Connections: The European Evaluation Newsletter*, 7-9.

Rifkin, J., Millen, J., & Cobb, S. (1991). Toward a New Discourse for Mediation: A Critique of Neutrality. *Mediation Quarterly, Vol 9*, 151-164.

Rogers, P. J. (2008). Response to paper "Systems thinking" by D. Cabrera et al.: Is it systems thinking or just good practice in evaluation? *Evaluation and Program Planning*, 325-326.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A Systematic Approach.* Sage Publications.

Sharma, S., & Gutierrez, I. A. (2010). An evaluation framework for viable business models for m-commerce in the information technology sector. *Systems Practice and Action Research,* 33-52.

Schechter, D. (1991). Critical Systems Thinking in the 1980s: A connective summary. In R. Flood, & Michael Jackson, *Critical Systems Thinking: Directed Readings* (pp. 213-227). Willey & Sons.

Snyder, S. (2013). The Simple, the Complicated and the Complex: Educational Reforms Through the Lens of Complexity Theory. *Organisation for Economic Co-operation and Development*, 1-35.

Ulrich, W. (1983). *Critical Heuristics of Social Planning: A New Approach to Practical Philosophy.* Berne: Haupt.

Ulrich, W. (1988). Churchman´s "process of unfolding"- its significance for policy analysis and evaluation. *Systems Practice,* 415-428.

Ulrich, W., & Reynolds, M. (2010). Critical systems heuristics. En M. Reynolds, & S. Holwell, *Systems approaches to managing change: A practical guide.* London: Springer.

Von Bertalanffy, L. (1950). The theory of open systems in physics and biology. *Science,* 23-29.

Von Bertalanffy, L. (1956). General Systems Theory. *General Systems*, Vol 1, 1-10.

Walton, M. (2016). Expert views on applying complexity theory in evaluation: opportunities and barriers. *Evaluation, Vol 22 (4)*, 410-423.

Williams, B., & Hummelbrunner, R. (2011). *Systems Concepts in Action.* Stanford, California: Stanford University Press.

Williams, B., & van´t Hof, S. (2014). *Wicked Solutions. A Systems Approach to Complex Problems.* Gumround.