

SCOPING A SYSTEMS-BASED METHOD FOR ORGANISATIONAL EVALUATION

Amanda J. Gregory
Centre for Systems Studies
Business School
University of Hull
Hull HU6 7RX
United Kingdom
a.j.gregory@hull.ac.uk

ABSTRACT

In this paper a systems approach to evaluation is proposed. It is argued that many evaluations fail to achieve a comprehensive assessment because they are essentially reductionist in nature; only focussing on limited aspects of an organisation and/or they emphasise the performance of the parts over the whole.

The argument is advanced in this paper that a systems-based approach to evaluation must put the study of the whole before that of the parts. It should stimulate an organisation's ability to learn by reflecting on the efficiency and efficacy of the interactions between its parts and the effectiveness and ethicality with which it engages with other systems in its environment. A systems approach will use a variety of different systems models to ensure that the evaluation is comprehensive, knowing that any one will be limited in terms of what it enables us to see. It is argued that at least four types of systems model are relevant. The first two can be called 'mechanistic' and 'organic' and they concentrate on building the internal capacity of an organisation. The second pair, 'stakeholder' and 'ethical' models are focussed more upon aspects of the organisation's external relationships.

Having established the four generic categories of evaluation, a summary is given of some 'emerging' forms of evaluation that may be drawn in to help us better achieve our aim of creating a truly systemic form of evaluation.

Moving closer to the ideal of being truly systemic demands more than simply adding new and emerging forms of evaluation to our tool-kit. A commitment to holism requires that we take a more critical look at how we might implement systemic evaluation in practice and critical systems thinking offers some guidelines for managing methodological pluralism in practice.

Keywords: Evaluation; Critical Systems Thinking

Systems Based Evaluation

INTRODUCTION

Evaluation “refers to the process of determining the merit, worth or value of something” (Scriven, 1991, p. 139). Unfortunately, many evaluations fail to achieve a comprehensive assessment because they do not recognise that ‘the performance of any system is not equal to the sum of the performances of its essential parts taken separately, but is a product of their interactions’ (Ackoff, 1999, p.33). What this reductionist orientation means in practice is that only limited aspects of an organisation or programme are emphasised and/or they focus on the performance of the parts over the whole.

The argument is advanced in this paper that a systems-based approach to evaluation must put the study of the whole before that of the parts. It should stimulate an organisation’s ability to learn by reflecting on the efficiency and efficacy of the interactions between its parts and the effectiveness and ethicality with which it engages with other systems in its environment. A systems approach will use a variety of different systems models to ensure that the evaluation is comprehensive, knowing that any one model will be limited in terms of what it enables us to see. It is argued that at least four types of systems model are relevant. The first two can be called ‘mechanistic’ and ‘organic’ and they concentrate on building the internal capacity of an organisation. The second pair, ‘stakeholder’ and ‘ethical’ models are focussed more upon aspects of the organisation’s external relationships.

CURRENT EVALUATION PRACTICE

Deficiencies in evaluation practice are widely reported. For example, Adams (2007) provides an insightful summary of what is wrong with current evaluation practice:

- Many reports just capture information on activities carried out and not on the outcomes of those pieces of work.
- Reports are often boring and written in a pseudo scientific style (methodology > evidence/findings > analysis > conclusions > recommendations) style which is not conducive to learning.
- Finding out about an organisation’s activities is undertaken by an external consultant and this information is not included in the evaluation report hence the consultant effectively ‘runs away’ with an organisation’s knowledge.

Three further criticisms may be added to the above:

Too Quick and Dirty

Simple methods of evaluation are too reductionist to capture the relative complexity of most organisations. If organisations are recognised to be “...vast, complex,

Systems Based Evaluation

fragmented, elusive, and multidimensional.” (Weick and Daft, 1983, p.72) then the processes created to evaluate them must have a similar level of variety. This is not to say the more complex, the better though, and it is recognised that it is impossible to capture ‘the whole’ in any evaluation. In order to make any evaluation manageable, “Investigators must make assumptions about organizations and adopt a limited perspective, however faulty to understand them.” (Weick and Daft, 1983, p.72). What is important is to make explicit the limited perspective, or boundary judgements, that is being adopted and to open it up to question. Since a whole systems understanding is an ideal, we should seek to make clear how far from that ideal our efforts will place us. Unfortunately, it is recognised that most organisational research has made low-level system assumptions (Pondy and Mitroff, 1978; Weick and Daft, 1983).

Over Elaborate and Resource Intensive

Given the foregone criticism that many evaluations are too quick and dirty to even begin to approach a whole systems view, it may seem ironic that the criticism that many evaluations are over elaborate and resource intensive is also being raised. However, emphasis on evaluation as accountability means that what can be measured is measured to the nth degree.

The criticism of over elaborate and resource intensive does not only pertain to the methods used but also to reporting procedures. Adams (2007) clearly makes the call for diversifying reporting procedures in looking for simpler and more creative methods: “While we acknowledge that reporting is important, the relationship could well be better served by using simpler procedures that enable organisations to account efficiently for inputs and outputs and look together at outcomes and impacts, for example, by using shorter reports and including structured learning events such as feedback workshops. The use of graphics is one of several more creative methods which can be useful for this. Where timelines have been used as one of the tools in an assessment it can be helpful to revisit them and to build on them, in effect developing them into graphical examples of a plan. Photographs and videos are very good examples of feedback mechanisms which enable stakeholders to engage with and ‘hear’ each other’s views.” (Adams, 2007, p.12)

Ambiguous Results and Misleading Conclusions

Evaluations can have important consequences and findings are often challenged by those whose views are contradicted (or at least unsupported) by the data (Weiss, 1970). Consequently, in designing evaluations it is important to be able to claim that the approach has validity. Constantine and Braverman (2004) make the distinction between ambiguous results, “...those that are open to multiple and conflicting conclusions” (p.237), and misleading conclusions, “...those that are inappropriately supported by either dubious results or the questionable use of results” (p.237). Hence in order to defend an evaluation against critics it is important that confident answers can be given to questions about: How can we know that organisation x has had the impact that the evaluation claims? Can we be sure that there isn’t some other reason,

Systems Based Evaluation

for instance a change in the organisation's environment, that has actually caused the change claimed?

In reflecting on these questions, it is relevant to look at two quite distinct forms of evaluation and the criteria of validity pertaining to each:

Quasi-Experimental/Experimental Evaluation

'Two-group designs are required for sound evaluation. A valid determination of impact requires comparing outcomes of a group of individuals who have participated in the program (treatment group) with an equivalent group of people who have not participated (control or comparison group). In theory, the best way to do this is by means of a randomized experiment, where individuals are assigned at random to the treatment or control group (Rossi and Freeman, 1993). Outcomes measures, chosen on the basis of program objectives, are observed at some interval after the intervention ends, with any differences between groups attributable to the program: that is, the program can be said to have caused the observed differences'. (Evaluation and Data Development, Strategic Policy, Human Resources Development Canada, 1998).

Hence, from this perspective the determination of validity requires the comparison of outcomes for a group of individuals who have participated in a program (treatment group) with an equivalent group of people who have not participated (control or comparison group).

Stakeholder-based/Ethnographic Evaluation

"...identify key 'stakeholders' (all of 'em); establish their (not the researchers') 'constructions' about a program via a prolonged period of field observation; arrange for these claims, concerns and issues to be negotiated in a 'hermeneutic dialectic circle' (act as a go-between); generate consensus with respect to as many 'joint constructions' as possible by creating a 'forum' (an evaluation arbitration and conciliation service); attempt to reconsider unresolved constructions by 'recycling' until 'enlightenment' is obtained (co-op 'til they drop)." (Pawson and Tilley, 1997, p.18).

Cronbach and Meehl (1956), in examining construct validity in psychological tests, have suggested that 'validity is dependent on the network of relationships in which a construct is embedded'. Hence it might be concluded from this that the most valid evaluations are those which most faithfully represent participants' views.

In the light of the above the notion that there is a general set of goodness criteria that might apply to any evaluation should be abandoned and it should be recognised "...goodness criteria are themselves generated from and legitimated by the self-same assumptions that undergird each inquiry paradigm, and hence are unique to each paradigm" (Guba (1988), quoted by Smith (1990), p.168).

Systems Based Evaluation

The above criticisms are important to consider in thinking about designing an approach to systemic evaluation, particularly the message that “...more is not necessarily better!” (Adams, 2007, p.3).

CHARACTERISTICS OF A SYSTEMS APPROACH

A systems approach to evaluation does not imply the use of a particular evaluation methodology but rather a way of seeing/representing the world and adopting an appropriate approach to it. Davies (2004) focuses on problems of representation and opines that “Without adequate representation it is much more difficult for an organization to propose, test and improve its ‘theory of change’, and to communicate this refined knowledge to others, enabling its wider use and impact.” (p.102).

An evaluation can only represent some aspect of reality if it has sufficient variety to capture the complexity of that reality. The question then is ‘are the evaluation approaches used sophisticated enough to capture the complexity of the organisations they are seeking to address?’ (Davies, 2004, p.102). Problems arise when evaluators seek to use an approach that is lacking in sufficient variety to capture the complexity of higher variety systems. In short, some evaluation theorists seek to use a single method of evaluation, such as the Logical Framework, in all situations even though the limitations of this approach are well documented. According to Davies (2004) ‘Arguments for the rejection of the Logical Framework typically stress how the world is not linear, mechanistic and predictable; but rather chaotic, complex and unpredictable. However, our world does include some fairly ordered environments where a relatively simple and linear view of expected changes is appropriate.’ (p.104). Unfortunately, relatively simple, stable situations are rare. More commonly, one is more likely to be operating in complex environments characterised by many interacting elements, conflict, diversity and pluralism: “As in any complex system, events at a local level are determined partly by unplanned interactions and partly by the social, organisational and political environment (patterns of power and affinity).” (Gilchrist, 2000: 269)” (cited in Barnes et al, 2003, pp.278-279).

According to Scriven (1991), valid evaluations are ones that “...take into account all relevant factors, given the whole context of the evaluation (particularly including the client’s needs) and weight them appropriately in the synthesis process.” (pp.372-373). It would be an easy mistake to assume that systemic evaluations are synonymous with valid evaluations but this is not the case. Scriven’s definition effectively designates all evaluations to the invalid category because, in referring to ‘all relevant factors’ and ‘the whole context’, what is actually defined is an inachievable ideal; a systems approach recognises this. System thinking is tempered with a realism that a whole systems view is not achievable but we should not abandon this ideal. In designing systems of inquiry such as evaluations, we should always be aware of how far from the ideal we are likely to be, recognising that, in taking an inevitably partial view, we will be privileging some perspectives whilst down-playing others. In recognising the partiality of evaluation

Systems Based Evaluation

design, the question might be raised ‘why bother with a systems approach if it is partial like any other?’. It is perhaps worthwhile to make clear the potential contribution that a systems approach might make and why it might be regarded as a more developed form of evaluation:

Fit for purpose

Variety theory (Ashby, 1956) tells us that simple organisations require simple methods of evaluation whereas complex organisations require complex methods. In both instances, though, the method should be designed with the purpose of capturing a whole systems view whilst being designed according to lean system principles, involving those actions, and only those actions, necessary for conducting the evaluation.

Dynamic

Most evaluation approaches are relatively static; if they are dynamic in any way it is through the comparison of a snapshot of an organisation at one point in time with it at another point in time. This approach in no way is able to explain why there may be any difference between the two. A form of evaluation that includes a technique such as System Dynamics should provide an understanding of the dynamics of the organisation and its broader context.

Capture the emergent

The performance of any system is a product of synergistic relations between the parts, not the parts operating in isolation. Hence, performance is said to be an emergent property. A system approach would be open to capturing unanticipated features, insights or variables and counter-intuitive results.

Environmentally aware

Where the organisation-environment boundary is drawn is essentially a value call in any evaluation. Too often the drawing of this boundary serves to define what is relevant (internal) to the evaluation and that which is regarded as irrelevant (external). This effectively disregards the many factors (political, economic, legal, social) that constrain or facilitate the organisation’s capacity for action and development.

Understands the embedded nature of systems

Systems exist within the context of wider systems. To make a change at one system level impacts on other levels. Consequently, it is important to consider the implications of change at the sub-system, system, and meta-system levels. For example, some arrangements might seem illogical at the system or organisational level, e.g. duplication of activities or functions, and it is only when reference is made to the wider or meta-system that things make sense.

The systematic literature review undertaken as part of this project has enabled the deficiencies in current evaluation practice to be articulated and, in the light of these, the potential contribution of a critically systemic approach to evaluation made clear.

Systems Based Evaluation

SCOPING A SYSTEMS APPROACH TO EVALUATION

There are many different strands of systems thinking and Critical Systems Thinking (CST) (Flood and Jackson, 1991; Jackson, 2003; Midgley, 1996) seeks to make sense of this diversity and bring them together in a coherent way. In summary, CST involves (Midgley et al, 1998):

- Critical awareness: Examining taken-for-granted assumptions and the conditions that give rise to them.
- Improvement: Defined temporarily and locally, but in a widely informed manner, taking issues of power into account.
- Methodological pluralism: Using a variety of methods in a theoretically coherent manner to address a corresponding variety of issues, and in awareness of their strengths and weaknesses.

In seeking to apply CST to evaluation, Gregory and Jackson (1992a,b) recognized that evaluation is essentially an abstract concept; to have meaning, it needs to be applied to something. Consequently, evaluation theory cannot be said to be an independent homogenous body of knowledge; each strand of evaluation theory has developed out of a parent discipline that determines the focus of attention. Hence organisational evaluation has changed in response to developments in organisational theory (for example, accountability and learning oriented evaluations reflect, in turn, mechanistic and organic models of the organisation).

On the basis of a Leverhulme Trust funded project with the National Association of Councils for Voluntary Service, Gregory and Jackson outlined a systemic approach to evaluation based on a classification of evaluation approaches:

Goal: This approach is based on the mechanistic model and is concerned with 'efficiency', i.e. with using the minimum of resources to get the output required. Organisations are viewed as goal seeking. Once goals have been set, achievement of them is evaluated using quantitative measures of inputs, processes and outputs. Decision makers are viewed as rational beings: able to define both their goals and the appropriate actions to achieve them, and willing to adjust these goals and actions in accordance with feedback from an evaluation. Logical framework, theories of change, Realistic Evaluation (Pawson and Tilley, 1997) approaches are typical of examples of this kind of positivist, goal-based evaluation.

System-Resource: This approach is based on the organic model and is concerned with 'efficacy', i.e. with the system's overall capacity to deliver outputs. Evaluation is undertaken with viability in mind, and this requires an organisation to be structured in such a way as to enable it to adapt to changing internal and external conditions. Gregory and Jackson (1992a), referring to the work of Yuchtman and Seashore (1967), argue that evaluators assume 'the organisation is in a perpetual

Systems Based Evaluation

bargaining relationship with the environment...The better the organisation's position in this relationship, the better able it is to pursue its higher level goals' (p 23). Therefore, evaluation focuses on the quality of the organisation's processes that ensure it is in the healthiest 'bargaining position' possible. Existing organisational processes are compared with a model of the ideal adaptive organisation (usually a model derived from the management or systems literature such as the Viable System Model (VSM)) to generate ideas for improvement. Other approaches associated with this form of evaluation focus on understanding the pattern of deep structures that drive behaviour both in the organisation and environment.

Multi-actor: This approach is based on the stakeholder model and is concerned with whether an organisation is 'effective' in doing the right thing in terms of what its stakeholders require. It is based on the idea that organisations are primarily constituted by the interactions of social actors with their own unique subjective perspectives about what the organisation is, or should be, doing. Therefore, surfacing the perspectives of a variety of stakeholders should be the essence of evaluation. According to Fitzpatrick (2004), House (2003) calls for stakeholder involvement to be refined according to breadth and depth. (p.545). This form of evaluation assumes that the desires of 'external' stakeholders who are seeking to influence the organisation (e.g., customers and lobby groups) are as legitimate as those of 'internal' stakeholders (e.g., owners and employees). The role of the evaluator is to assess people's levels of satisfaction and help to minimize constraints on the organisation's ability to satisfy its stakeholders' objectives. Guba and Lincoln's Fourth Generation Evaluation (1989) is a good example of this type of evaluation.

This work was further developed by Gregory in 1996 when she acknowledged that her initial approach, while being an advance on the work of 'isolationists' (people who claim that only one kind of approach is valid or useful), was overly 'mechanical'. She went on to advocate using different methods in parallel, while holding a 'reflective conversation' between them (Gregory, 1996).

Most recently, the development of a systemic approach to evaluation has been picked up by Boyd et al (2007). Using Gregory's work as a foundation, they make several changes:

- The models were relabelled to, they claim, 'enhance accessibility' ('goal-based' evaluation remains the same, 'systems-resource' becomes 'organisational' evaluation, 'multi-actor' becomes 'stakeholder' evaluation).
- Three of the models of evaluation were connected in a simple but intuitive ordering (see Figure 3). As Boyd et al state: "A stakeholder evaluation can lead to the setting of community-sensitive goals, the achievement of which can be measured through goal-based evaluation, and pursuit of the goals will be enhanced by organizational evaluation." (p.1312).

Systems Based Evaluation

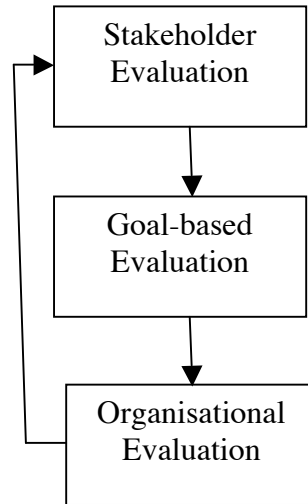


Figure 3 The Relationship Between the Three Approaches to Evaluation
Source: Boyd et al, 2007

The connection of the three models of evaluation in this way is significant as it emphasises that each form of evaluation should not be seen in isolation, for to do so might have serious consequences. For example, to use only a goal-based form of evaluation without a consideration for stakeholder participation might result in an organisation with rather dubious objectives being rated highly. However, the model is perhaps too serial and linear to capture the complexity of the organisational capacity development effort. This point will be taken up again later.

Importantly, Boyd et al place a heavy focus on identifying minority interests through boundary critique throughout the process of evaluation but they do not particularly enable evaluators to deal with this by making specific a form of evaluation based on the ethical model. Such an approach, which would be focussed upon how the organisation treats its various stakeholders, might be drawn from the systems literature and question whether the organisation is it acting to promote diversity, to advance democracy and social justice, with a view to ecological sustainability. An appropriate approach to consider here is Critical System Heuristics (CSH), created by Ulrich (1983).

CSH recognises that the design of any program or organisation involves making judgements about what lies within and without the 'system' boundary. This important stage normally goes unrecognized and, because of this, questions about who should be involved and how they should be involved are not addressed. CSH enables stakeholders to identify value and boundary judgements that are being made about what is regarded as relevant to a situation, and to propose alternative judgements. Such judgements should not be regarded as the sole province of evaluation experts or other powerful stakeholders: ordinary citizens should be viewed as equally capable of rational inquiry into values and boundaries. Hence CSH provides a structured way of questioning (12 critical questions

Systems Based Evaluation

posed in both 'is' and 'ought' modes) to challenge what is to count as relevant knowledge in an evaluation and thus expose the politics of evidence that would otherwise not be revealed. Gregory (1997) sees that CSH may be supplemented with Responsive Evaluation (Stake, 1980 and 2004) as the two approaches may be mutually supportive: responsive evaluation provides an explicit process that reveals and reports minority value-positions, and also helps prevent stakeholders from abdicating control to evaluation experts, while CSH provides questions for challenging the value and boundary judgements made by the different stakeholders.

EMERGING FORMS OF EVALUATION

Having established the four generic categories of evaluation, it is relatively easy to associate a specific methodology with its underlying form thus making clearer underpinning assumptions. It is worth recognising that there is much development work going on by proponents of the different forms of evaluation and, in this section, a summary will be given of some 'emerging' forms of evaluation that may be drawn in to help us better achieve our aim of creating a truly systemic form of evaluation.

Data Envelopment Analysis

This approach is based on the mechanistic model. Data Envelopment Analysis (DEA) is a technique that has been widely applied to measuring efficiency and productivity in the three decades since it was first proposed in the seminal article by Charnes et al (1978). Basically, this approach compares inputs and outputs across a number of 'Decision Making Units' (DMUs), comparing each DMU with others in the sample to identify which are the efficient operators. The DMUs can be firms or any other "...entity responsible for converting inputs into outputs and whose performances are to be evaluated" (Cooper et al, p.22). DEA has been widely applied to non-profit-making bodies such as libraries and universities.

Quasi-Experimental Evaluation and Synthetic Indices

Crawford et al (2004) identify two practical challenges in evaluation:

- *measurement*: determining changes in the circumstances encountered by intended beneficiaries
- *attribution*: identifying causality between the intervention and observed changes.

According to Crawford et al "...the measurement and attribution of non-tangible changes in complex environments are problematic. Synthetic indices are frequently used by statisticians in other contexts to capture complex changes and may prove useful to describe changes effected by aid projects." (p.175).

Crawford et al go on to look at how a 'Background Change Index' could be constructed, perhaps based on the 'STEER' (social, technological, economic, environmental and political) framework, to categorize the drivers/inhibitors of change. They state that "An index based on this framework could prove useful in tracking variables such as climate (e.g. temperature and rainfall), social conditions (e.g. employment, civil disturbances,

Systems Based Evaluation

personal security) and economic conditions (e.g. world commodity prices, local market commodity prices).” (p.187). Such an index might help to explain programme effects, for example, was a programme a failure because there was a rise in unemployment or was it a success because there was a massive rise in unemployment in other regions but only a small rise in the project locality?

System Dynamics

This approach is based on the organic model. System Dynamics is an approach created more than 35 years ago (Forrester, 1971) that involves building a model that captures the dynamic nature of systems and aims to identify the underlying structure of social systems that produce counterintuitive behaviour. The approach is increasingly being recognised as a useful tool for evaluation purposes and Fredericks et al (2008) declare that it adds value in two important ways, “First, it provides a way to diagram endogenous relationships in a way that illustrates where programs have potential unintended consequences. Second, the process is iterative, with some similarity to grounded theory. Therefore, the qualitative diagramming process is a flexible tool in the evaluation that would provide a deeper understanding of potential problems in the system. This would be a tool that complements traditional evaluation techniques by showing the “structure” that might explain the behaviour of important outcome variables over time.” (Fredericks, p.266). The development of a System Dynamics model draws on three sources of information: numerical data, program documentation, and the expert knowledge of key stakeholders in the system (Richardson & Pugh, 1981, cited in Fredericks, 2008). The first step in building a System Dynamics model is qualitative in nature, involving the articulation of a conceptual diagram of the relationships within the system, identifying causal loops, and exploring their dynamic nature. The second step is more quantitative and is based on simulation modelling, which explores the relationships in the system and provides insights into the implications of different approaches.

The development of a System Dynamics model may be useful at the planning and evaluation stages (Fredericks et al, 2008). At the planning and implementation stages, potential unintended consequences can be surfaced, and a shared understanding of the structure and process of the system achieved. At the evaluation stage, a System Dynamics model can be used to provide context, explain competing goals and to identify important feedback processes and the potential for unintended consequences.

Chaos and Complexity Theory

This approach is based on the organic model. Complexity theory is rooted in Chaos theory. Chaos does not mean an absence of order but rather a domain between deterministic order and randomness which is complex. Haynes (2008) recognises that “Complexity accepts the inevitability that individuals are often subservient to social structures, but realizes that feedback from individuals, however, small in power, can contribute in unpredictable ways to the future organization and representation of structures.” (p.402).

Systems Based Evaluation

Haynes (2008) proposed a model of evaluation based on complexity theory: it “...begins with a storyline account of the events in a given period. This is a qualitative attempt to speculate about what have been the key events and policy instruments...Second, the dynamic nature of complex systems requires that careful attention is given to how they perform over time...by way of the examination of quantitative timelines. It is argued that the examination of key trends over a given time period is not enough and that the rate of change is an additional key factor that must be considered...The third stage is to consider qualitatively the combined interaction of these dynamics. The aim in this final stage is to explore further the concept of an attractor. In complexity methodology an attractor is a point of order within a system of change. In mathematical terms it defines the boundaries, or limits of fluctuating change over time. In qualitative terms, in a policy system, an attractor can be argued to be a set of values or logics that give a policy system a general characteristic of relative stability in a given time period.” (p.405).

Soft Systems Methodology

This approach is based on the stakeholder model. Soft Systems Methodology (SSM) created by Checkland (1981) has proven one of the most influential systems methodologies with a wide field of implementation and, in Jackson's view (2000), generated a “...paradigm revolution in systems thinking which liberated the discipline from the intellectual straightjacket in which it had been locked and, at the same time, made it much more obviously relevant to managers”. It focuses on building shared understanding in order to bring about change in value systems, cultures and philosophies that exist in organisations. SSM is a methodology, which can assist us to examine an organisation's problems, consider the implications of possible solutions (using systems models) and arrange participative debate about feasible and desirable change in order to generate commitment to action.

Social Network Analysis

This approach is based on the stakeholder model. Penuel et al (2006) consider “Social network analysis (SNA), which focuses on understanding the nature and consequences of ties between individuals or groups (Scott, 2000; Wasserman & Faust, 1994), has become an increasingly popular method within the social sciences for exploring human and social dynamics.” (p.437). Penuel et al point out that there are many use of SNA from a professional/ethical perspective (e.g. identifying people that have not agreed to be involved in the evaluation process) and this may well limit its potential as a mainstream evaluation tool (it may be useful though to assess projects in which the level of connectivity is an appropriate measure of effectiveness).

A COMMENT ON THE COMPLEMENTARITY OF EVALUATION APPROACHES

The literature review undertaken has revealed a range of new approaches to evaluation that might be considered important in realizing our aim of achieving a more systemic

Systems Based Evaluation

approach to evaluation. To be clear, we are striving to achieve a ‘whole systems’ evaluation. Seeking not to adopt a partial view that privileges, either deliberately or accidentally, some aspects of organisation at the cost of others but rather, like Seashore (1983), sees the variety in models of evaluation as “...nicely complementary referring to different but interdependent facets of organizational behavior” (p.61).

A sound basis is now in place for the further elaboration of this form of evaluation to enable it to fulfil its potential of moving closer to the ideal of being truly systemic but simply adding new and emerging forms of evaluation to our tool-kit is not sufficient. Our commitment to holism requires that we take a more critical look at how we might implement systemic evaluation in practice. Although Boyd et al’s ordering of the different evaluation models is appealing in its simplicity, it is perhaps too serial and linear to capture the complexity of most organisations (Boyd et al, 2007).

Any attempt to equip an organisation to bring about beneficial social change must make use of the four systems models previously identified in carrying out evaluation. Because the models are not just complementary but also, to a degree, contradictory in what they tell us about organisations and how to improve them, such multi-methodological work requires us to live with and manage paradigm incompatibility.

Critical systems practice accepts that different systems models, and their related methodologies, are based on opposing paradigms and set forth incompatible philosophical assumptions, and that they cannot, therefore, be integrated without something being lost. It seeks to manage paradigm diversity by encouraging them to confront one another on the basis of ‘reflective conversation’. (Jackson, 2003; Gregory, 1996). No paradigm or methodology is allowed to escape unquestioned because it is continually confronted by the alternative rationales offered by others. The form of evaluation suggested is difficult to operationalise but it is the only kind appropriate given the heterogeneity, complexity and turbulence of the world we currently inhabit.

Critical systems thinking offers some guidelines for managing methodological pluralism in practice. It insists that it is desirable, to ensure paradigm diversity, that methodological pluralism be applied at all stages of an evaluation. It is tempting to allocate different methodologies to the various stages because they seem most suited to those stages; for example front-ending an evaluation with a more subjective approach to ‘deal with’ multiple perceptions before moving on to a ‘harder’ methodology for measurement. There is no theoretical justification for such a procedure. To theorists of an objectivist persuasion, the analysis stages of an evaluation carried out according to a subjectivist logic are not ‘richer’; they are simply misguided. To subjectivists, the softer issues of culture, politics and power cannot simply be made to disappear. Critical systems practice’s preference (Jackson, 1999) is to observe methodological pluralism at each and every stage of an evaluation. Pollack (2009), in a recent article, calls this a ‘parallel’ approach to multi-methodology as opposed to a ‘serial’ approach.

Of course, adopting a parallel approach to multi-methodology makes life more complicated. To cope with this, critical systems practice has developed the device of

Systems Based Evaluation

working with ‘dominant’ and ‘dependent’ methodologies in creative combination. An explicit choice of ‘dominant’ methodology is made to run an evaluation with ‘dependent’ methodologies, reflecting alternative paradigms, in the background. The relationship between dominant and dependent methodologies changes as the evaluation proceeds in order to maintain flexibility. Making explicit the rationality underpinning the methodology with which we are operating, and being ready to switch rationality and methodology, makes the initial choice of ‘dominant’ approach less committing.

The manoeuvre of switching paradigms and methodologies places significant demands on evaluators. Brocklesby (1997) has explored the ‘cognitive difficulties’ involved for individuals in becoming multi-methodology literate; concluding that it is difficult but by no means impossible. Pollack (2009), however, describes a project in which “...frequent swaps were made between the hard and soft paradigms by a single practitioner with little apparent difficulty” (p.163). Becoming multi-methodology literate depends on a detailed understanding of the different philosophies underpinning the various systems approaches and evaluation methodologies. Soft systems methodology, to take an example, is employed in a radically different fashion by someone who grasps its subjectivist assumptions than by an analyst who tries to reconcile it with an unchallenged objectivist mindset.

<i>Type of Evaluation</i>	<i>Methodology</i>	<i>Underlying Model</i>	<i>Measure of Success</i>
Goal	<ul style="list-style-type: none"> • Logical Framework • Theory of Change • Data Envelopment Analysis • Quasi-Experimental Evaluation and Synthetic Indices 	Mechanical	Efficiency
System-Resource	<ul style="list-style-type: none"> • Viable System Model • System Dynamics • Complexity Theory 	Organic	Efficacy
Multi-Actor	<ul style="list-style-type: none"> • Fourth Generation Evaluation • Soft Systems Methodology • Social Network Analysis 	Stakeholder	Effectiveness
Ethical	<ul style="list-style-type: none"> • Critical System Heuristics • Responsive Evaluation 	Ethical	Ethicality

Table 1 A Classification of Evaluation Types

In summary, a critically systemic approach must pay attention to matters of efficiency, efficacy, effectiveness and ethicality. This can be achieved by employing a variety of systems approaches and evaluation methodologies, based on alternative models and paradigms, in parallel (see Table 1). Critical systems practice suggests how this can be

Systems Based Evaluation

operationalised by evaluators becoming multi-methodology literate. Whilst this demands much of the evaluator, it is surely worth the effort.

CONCLUSION

In this paper a systems approach to evaluation was proposed. It was argued that many evaluations fail to achieve a comprehensive assessment because they are essentially reductionist in nature; only focussing on limited aspects of an organisation and/or they emphasise the performance of the parts over the whole.

The argument was advanced that a systems-based approach to evaluation must put the study of the whole before that of the parts. It should stimulate an organisation's ability to learn by reflecting on the efficiency and efficacy of the interactions between its parts and the effectiveness and ethicality with which it engages with other systems in its environment. A systems approach uses a variety of different systems models to ensure that the evaluation is comprehensive, knowing that any one will be limited in terms of what it enables us to see. It was argued that at least four types of systems model are relevant. The first two can be called 'mechanistic' and 'organic' and they concentrate on building the internal capacity of an organisation. The second pair, 'stakeholder' and 'ethical' models are focussed more upon aspects of the organisation's external relationships.

Having established the four generic categories of evaluation, a summary was given of some 'emerging' forms of evaluation that may be drawn in to help us better achieve our aim of creating a truly systemic form of evaluation.

Moving closer to the ideal of being truly systemic demands more than simply adding new and emerging forms of evaluation to our tool-kit. A commitment to holism requires that we take a more critical look at how we might implement systemic evaluation in practice and critical systems thinking offers some guidelines for managing methodological pluralism in practice.

REFERENCES

- Ackoff, RL (1999) *Re-creating the Corporation*, OUP : Oxford.
- Adams, J (2007) *Learning and Accountability: A Monitoring & Evaluation Consultant's Perspective*, PraxisNote No. 32, INTRAC available at:
<http://www.intrac.org/docs.php/2634/PraxisNote%2032%20%20Learning%20and%20Accountability.pdf>
- Ashby, WR (1956) *An Introduction to Cybernetics*, Methuen, London.
- Barnes, M, Matka, E, and Sullivan, H (2003) Evidence, understanding and complexity: Evaluation in non-linear systems, *Evaluation*, 9(3).

Systems Based Evaluation

- Boyd, A, Geerling, T, Gregory, WJ, Kagan, C, Midgley, G, Murray, P, and Walsh, MP (2007) Systemic evaluation: A participative, multi-method approach, *Journal of the Operational Research Society*, 58.
- Brocklesby, J (1997) Becoming multi-methodology literate : an assessment of the cognitive difficulties of working across paradigms, in : J Mingers and A Gill (eds) *Multimethodology*, Wiley, Chichester, pp.189-216.
- Charnes, A, Cooper, WW and Rhodes, EL (1978), Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research*, 2(6).
- Checkland, PB (1981) *Systems Thinking, Systems Practice*, Wiley, Chichester.
- Cooper, WW , Seiford, LM and Tone, K (1999) *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Kluwer, Boston.
- Constantine, NA, and Braverman, MT (2004) Appraising evidence on program effectiveness, in: Marc T Braverman, Norman A Constantine, and Jana Kay Slater (eds), *Foundations and Evaluation: Contexts and Practices for Effective Philanthropy*, Jossey-Bass, San Francisco.
- Crawford, P, Perryman, J, and Petocz, P (2004) Synthetic indices: A method for evaluating aid project effectiveness, *Evaluation*, 10(2).
- Cronbach, LJ, and Meehl, PE (1956) Construct validity in psychological tests, *Minnesota Studies in Philosophy and Science*, 1.
- Davies, R (2004) Scale, complexity and the representation of theories of change, *Evaluation*, 10(1).
- Evaluation and Data Development (1998) Strategic Policy, Human Resources Development Canada.
- Fitzpatrick, JL (2004) Exemplars as case studies: Reflections on the links between theory, practice and context, *American Journal of Evaluation*, 25(4).
- Flood, RL, and Jackson, MC (1991) *Creative Problem Solving: Total Systems Intervention*, Wiley, Chichester.
- Forrester, J (1971) Counterintuitive behaviour of social systems, *Technology Review*, 73(3).
- Fredericks, KA, Deegan, M, and Carman, JG (2008) Using system dynamics as an evaluation tool: Experience from a demonstration project, *American Journal of Evaluation*, 29(3).
- Gregory, AJ, and Jackson, MC (1992a) Evaluation methodologies: A system for use, *Journal of the Operational Research Society*, 43.
- Gregory, AJ, and Jackson, MC (1992b) Evaluating organizations: A systems and contingency approach, *Systems Practice*, 5.
- Gregory, AJ (1996) The road to integration: Reflections on the development of organizational evaluation theory and practice, *Omega*, 24(3).
- Gregory, AJ (1997) Evaluation practice and the tricky issue of coercive contexts, *Systems Practice*, 10.
- Guba, EG, and Lincoln, YS (1989) *Fourth Generation Evaluation*, Sage, Newbury Park.
- Haynes, P (2008) Complexity theory and evaluation in public management: a qualitative systems approach, *Public Management Review*, 10(3).
- Jackson, MC (1999) Towards coherent pluralism in management science, *Journal of the Operational Research Society*, 50.

Systems Based Evaluation

- Jackson, MC (2000) *Systems Approaches to Management*, Kluwer/Plenum, New York.
- Jackson, MC (2003) *Systems Thinking: Creative Holism for Managers*, Wiley, Chichester.
- Midgley, G (1996) What is this thing called CST?, in: RL Flood and NRA Romm (eds) *Critical Systems Thinking: Current Research and Practice*, Plenum, New York.
- Midgley, G, Munlo, I, and Brown, M (1998) The theory and practice of boundary critique: developing housing services for older people, *Journal of the Operational Research Society*, 49.
- Pawson, R and Tilley, N (1997) *Realistic Evaluation*, Sage, London.
- Penuel, WR, Sussex, W, Korbak, C, and Hoadley, C (2006) Investigating the potential of using social network analysis in educational evaluation, *American Journal of Evaluation*, 22(4).
- Pollack J (2009) Multimethodology in series and parallel : strategic planning using hard and soft OR, *Journal of the Operational Research Society*, 60.
- [Pondy](#), L and Mitroff, I (1978) Beyond Open Systems Models of Organizations. In: LL Cummings and B Staw (eds) *Research in Organizational Behavior*, JAI Press, Greenwich, CT.
- Scriven, M (1991) *Evaluation Thesaurus*, 4th ed, Sage, London.
- Seashore, SE (1983) A framework for an integrated model of organizational effectiveness, in: KS Cameron and DA Whetten (eds) *Organizational Effectiveness. A Comparison of Multiple Models*, Academic Press, London.
- Smith, JK (1990) Alternative research paradigms and the problem of criteria, in: EG Guba (ed) *The Paradigm Dialogue*, Sage, London.
- Stake, RE (1980) Program evaluation, particularly responsive evaluation. In: WB Dockrell and D Hamilton (eds) *Rethinking Educational Research*, Hodder and Stoughton, London, pp 72–87.
- Stake, RE (2004) *Standards-Based and Responsive Evaluation*, Sage, Thousand Oaks, CA.
- Ulrich, W (1983) *Critical Heuristics of Social Planning: A New Approach to Practical Philosophy*, Haupt, Bern.
- Weick, KE and Daft, RL (1983) The effectiveness of interpretation systems, in: KS Cameron and DA Whetten (eds), *Organizational Effectiveness. A Comparison of Multiple Models*, Academic Press, London.
- Weiss, CH (1970) The politicization of evaluation research, *Journal of Social Issues*, 26(4).
- Yuchtman, E and Seashore, SE (1967) A system resource approach to organizational effectiveness, *American Sociological Review*, 32.