Theory, Method and Tools for Evaluation Using a Systems-based Approach

Päivi Jokela[1], Peter Karlsudd[2] and Martin Östlund[1]
[1]School of Communication and Design, University of Kalmar, Sweden
[2]School of Human Sciences, University of Kalmar, Sweden
paivi.jokela@hik.se
peter.karlsudd@hik.se
martin.ostlund@hik.se

Abstract: This work introduces an evaluation model for examining organisations that in one form or another depend on modern information technology for their core activities. The evaluation model, named SUV, is based on a systems science approach and has been developed at the eHealth Institute (*eHälsoinstitutet*) at the University of Kalmar.

The central mechanism of the model, the SUV diagram, partitions the scope of evaluation into seven categories and three levels. The seven categories have been selected on a systems science basis with inspiration from systems thinking. The rationale for making this the starting point for the evaluation model is that it may be justifiably stated that the framework provided by systems thinking truly encompasses the breadth of human activity. Making use of this general framework, the SUV diagram provides a general roadmap to guide the evaluation effort. The three levels (organisation, technological and individual) were selected to add detail to the analysis complement the categories and enrich the analysis of each category as well as to the dynamic interplay among them.

The SUV methodology is based on a continuous evaluation process whose driving-force is the wish of the interested parties to develop their own activities. It is a methodology within whose framework any and all methods for data gathering deemed appropriate for the evaluation at hand can be used. The use of multiple methods is explicitly encouraged for the purpose of gaining a multi-perspective view of the organisation/activity under scrutiny.

Based on the accumulated findings from pilot studies, the model was operationalised in the form of an IT application for electronic surveys. The application is expected to contribute in making the evaluation process more efficient and add structure to this process.

Keywords: systems science, systems thinking, evaluation methodology, multi-method, electronic surveys

1. Introduction

Taking appropriate strategic or tactical action based on evaluation results is essential for any organisation to be able to adapt and survive in a changeable environment. However, the process of evaluation is fraught with problems and its findings are often regarded as incomplete, erroneous or biased (Barroso et al., 2006, Hartford et al., 2007). One reason is that systems involving human interaction are inevitably messy and complex. One way to handle this is to divide the system into smaller parts and examine each one at a time. This is a common tactic to manage complexity, but the downside is that the relationships between the parts and the ones between the part and the whole are neglected. Systems science provides a framework within which complex systems can be understood and it offers structure and method to describe and analyze parts within their context and systems as wholes (von Bertalanffy, 1968, Boulding, 1985, Checkland, 1981). The usefulness of the systems science view and the holistic perspective in examining organisations has been both theoretically justified (von Bertalanffy, 1968, Checkland, 1981, Beer, 1979) and is supported by empirical findings from diverse fields (Sandelin and Sarvimäki, 2000, Marine, 2002, Cohen et al., 2007).

The systems science approach can be used to examine a system at any level. Within the boundaries of a system there exist an infinite number of subsystems (von Bertalanffy, 1968, Boulding, 1985, Checkland, 1981, Edwards, 2005). The process of deconstruction of the system should be guided by the stated purpose for the study and be regarded as part of defining its scope. This is similar to what Mulej (2007) lables requisite holism/realism. The inherent complexity of any human system, with its infinite number of subsystems and complex contexts, implies that any description or understanding of, a system must be based on incomplete evidence (Backlund, 2002, Jacucci and Hanseth, 2006, Simon, 1969). Furthermore, human activity is also fleeting, constantly changing and morphing, both at surface level and at a deeper level (Cohen et al., 2007). Still, all but the most temporary and volatile systems tend to have some kind of stable core, which contributes to order in the system, although the relationship between structure and order is not fully understood (Bohm and Peat, 2000).

Reference this paper as:
Jokela, P., Karlsudd, P. and Östlund, M. “Theory, Method and Tools for Evaluation Using a Systems-based Approach.” *The Electronic Journal Information Systems Evaluation* Volume 11 Issue 3 2008, pp. 197 - 212, available online at www.ejise.com

By letting the purpose guide the process of analysis, a useful, albeit incomplete, representation can be constructed of the system under scrutiny. In practice, this means that some parts of the system will be analyzed in greater detail while others will be analyzed in less detail. This kind of goal-directed, focused analysis is necessary to be able to manage the process and also the product of analysis. At the same time one must not lose track of the whole picture. The marriage between details and the whole (or the wholes) is the main challenge to be overcome in order to fulfil the potential of a systems science based approach (Boulding, 1985, Checkland, 1981, Koestler, 1967, von Bertalanffy, 1968). Systems science aids us in this process by providing a language to describe the systems and their parts; and to provide structure and guidance in the traversal of systems levels when shifting between different levels of granularity (Cottam et al., 2005, Edwards, 2005).

The model described here, systematic evaluation – also called the SUV model, takes a systems science based approach to evaluation and uses systems thinking based methods to perform evaluative work and analyse findings. SUV can be characterised as theory-based evaluation with a holistic approach. SUV-evaluation is mainly formative, in a sense that it is aimed to provide evaluation participants with feedback and to develop new knowledge, which is needed in order to continuously improve the evaluand and its performance. Involvement of the participants and interactivity between the evaluator and participants are essential throughout the evaluation process; from the beginning when the diverse evaluation issues and goals are established to the end when action must be taken to make changes in the organisation, based on the evaluation outcomes.

SUV has been developed by the authors while working at the e-Health Institute (eHälsoinstitutet), University of Kalmar. Previous work on the model is described in (Brandt et al., 2003). The SUV tool is especially suitable for organisations that *use* (as opposed to having it as their main business) modern information and communication technology; for example: e-learning (Harun, 2002) and health services (Zhao et al., 2002). This, of course, applies to most organisations today (in the western world), and in the near future any organisation that does not, will have a hard time surviving. In the methodology based on the SUV model, the use of multiple methods is encouraged. One method that is given special attention here is a web-based questionnaire tool has been constructed to provide technical support and practical guidance in formulating evaluation questionnaires.

The basic principles of the SUV model may be summarised as follows:

- Analysing holistically by systematically considering the parts and their mutual relations.
- Analysing activities starting from three levels: ***Individual***, ***Technology*** and ***Organisation*****,** and seven categories: ***Goal Seeking, Hierarchy and Relations, Differentiation and Entropy, Inputs, Transformation Process, Outputs*** and ***Regulation***.
- Enabling previous evaluations to be analysed by the same principles
- Within the holistic framework the use of complimentary methods of data capture is encouraged.

The following section will give a brief orientation into the theory on which the model is based. Then follows a description of the method and the program supporting the model and creating the conditions for evaluation construction, implementation, analysis and innovation.

2. Evaluating on a systems-based approach

Systems theory is a theoretical approach for formulating general laws for systems, regardless whether they are physical, biological, social etc. Within *organisation theory* systems thinking involves that an organisation, like an organism, is affected by its environment to which it has to adapt for its survival (Boulding, 1985, Churchman et al., 1969, Litterer, 1973, von Bertalanffy, 1968).

Systems have been defined in the literature in various ways. A frequently quoted specification by Hall and Hagen (1969) reads “A system is a set of components (objects) containing relations between the components (objects) and between their properties.”

A system to be evaluated can be defined as a whole organisation but also as a department or a specific computer system. It is important to notice that the definition of system boundaries as well as the division into separate system levels does not follow any strict rules or principles but is a result of the goals and attitudes of the evaluator (Katz and Kahn, 1980). In an open system information and material are continuously interchanged with the environment. This continuous interaction with the environment makes it complicated to

distinguish an open system from its surroundings (Checkland, 1981). The aim of the evaluation model is to facilitate the choice of limitations by using an area division derived from systems theory.

2.1 Main principles for a system

The properties of a system may be explained according to the following basic principles:

1. Holism: A change in part of a system affects the whole system (Boulding, 1985, Litterer, 1973, von Bertalanffy, 1968).
2. Non-summativity: The whole is more than the sum of the parts (ibid).
3. Equifinality: All open systems can attain the same state or result via different starting-points and approaches (ibid).
4. Multifinality: A common starting-point and/or the use of different roads may lead to different results (ibid).
5. Circular and manifold causality, non-linearity. A causal connection is not linear but a more refined explanation model is needed, in that many parts of the system may be affected on different occasions (de Shazer, 1994).

2.1.1 Holism

Evaluating an activity on a system-theoretical basis entails seeing the whole by systematically studying the parts and then feeding these results back into the activity from a holistic perspective. Most systems may be regarded as hierarchically organised, consisting of a long series of subordinate partial systems which all interact within the superordinate system (Boulding, 1985, Edwards, 2005, Koestler, 1967, Litterer, 1973, von Bertalanffy, 1968).

2.1.2 Non-summativity

Every partial system is of importance to the whole and should thus be examined separately but also in relation to the other parts. A hackneyed but relevant expression is that “no system is stronger than its weakest link”. Therefore it is essential that all parts of importance to the activity become visible and are included in the evaluation while at the same time the holistic perspective must be the guiding principle (Boulding, 1985, Litterer, 1973).

2.1.3 Equifinality and multifinality

In contrast to closed systems, all open systems lack an unambiguously defined point of equilibrium. Activities and organisations are never static but change more or less continuously. This is the reason why an evaluation must be considered from a dynamic perspective and why the result must always be regarded as more or less preliminary. One important concept in systems theory is *equifinality*, which means that all open systems can reach the same condition or result from different starting-points and by different approaches. In other words, an activity may attain the same goal by different methods. Similarly, the principle of *multifinality* may be explained as a common starting-point where similar conditions may lead to different results. Even though an organisation, for instance, gives the same directions and provides the same conditions for its departments the result may turn out to be completely different (Boulding, 1985, Litterer, 1973, von Bertalanffy, 1968).

2.1.4 Causality

A *cause-effect* relation or a linear causality, as it may also be formulated, is common in Western thinking. An explanation model like this briefly means that all events derive from *one* cause. This event may in turn give rise to other events. Linear causality thus means that *one* cause leads to *one* event (Petit and Olsen, 1994). Linear causality is not suitable for systems theory, as it provides a simplified image of reality. In systems thinking various causality theories have been brought forward, one of the most frequently mentioned being *circular causality*. The gist of it is “that an event is considered in relation to other events taking place simultaneously, interfering with each other and affecting each other reciprocally. An individual’s actions are affected by the environment, while these actions simultaneously affect the environment.” (Schöjdt and Egeland, 1994), p. 83). The consequence of this is that all events may be regarded as both cause and effect. All parts or events are dependent on one another (Petit and Olsen, 1994). Circular causality as an explanation model is not sufficient, however, for studying more complex systems. A more refined explanation model is *manifold causality*, according to which it is meaningless to try to find *one* cause of an event, but that there may be many causes behind what has happened. If circular and manifold causality are combined, an

even more thoroughgoing and refined causal explanation is attained, entailing that all parts of a system affect each other directly and indirectly (ibid.).

2.2 Human activity systems

Organisations that can be the target of a SUV-evaluation are human activity systems, in a sense that they can be modeled as a range of human activities that are related to each other within an organisational framework and these activities are supported or mediated by information technology. The perception of such systems is associated with the individual observers' image of the world, or according to Checkland's terminology their *Weltanschauung* (Checkland, 1981). Although these perceptions may sometimes converge into an agreement of what constitutes a meaningful system, there is no objective method available to create a joint description of the system. The special nature of the human activity system hinges on the concept that the system is aware of itself, in the same way that a person is aware of him- or herself. If a member of the human activity system was to observe the system he/she is part of, this would alter the system. Consequently, any member of the human activity system is barred from participating in a group observing this system with the purpose of building absolute, objective public knowledge about the system. However, a pre-requisite for building public knowledge is that anyone and everyone can participate in observing something. The exclusion of the members of the human activity systems from observing group makes it impossible for knowledge about human activity systems to ever reach the level of objective public knowledge. This implicates that the knowledge that is needed to evaluate human activity systems must be gained by using a different methodology.

In this work, we postulate that the process of evaluative inquiry can be organised as a learning system, where multiple issues originating from different *Weltanschauung* are identified, presented and debated within a systems framework. This discourse is the basis of shared knowledge as well as new knowledge created jointly by the evaluation participants, and the deeper understanding may eventually lead to improvements of the studied organisation. (Jackson and Kassam, 1998, Checkland, 1981)

2.3 The main principles of the evaluation model

In the SUV evaluation model structures and processes of importance are divided into seven categories and three levels. The categories may be illustrated as the spokes of a wheel, which give stability and direction to the activity; the levels show the recurring layered structure of an organisation (Figure 1).

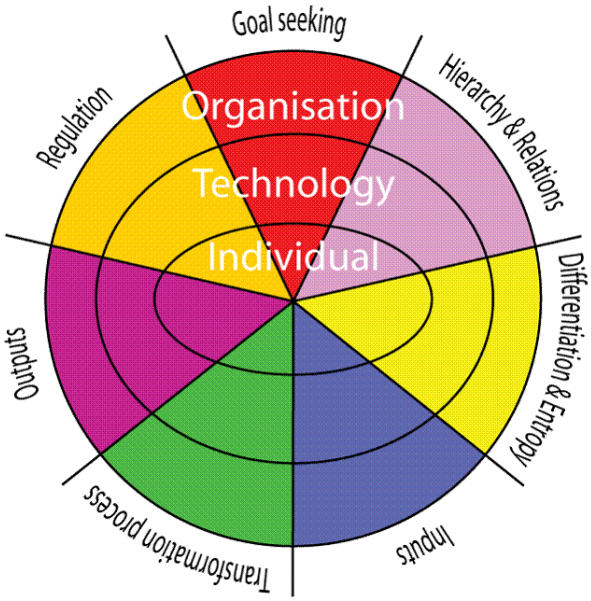

Figure 1: The seven categories and three levels in a SUV evaluation.

The three levels of the model can be seen as a nested holarchy, a concept that was introduced by Koestler (1967) and further developed by Edwards (2005). Each level not only includes but also transcends the previous levels. The complexity increases successively when one moves towards higher levels (Edwards, 2005). In the SUV model, the fundamental level comprises individuals and interactions between them. At the next level, information technology is added to the system so that the technology level includes individuals,

computer systems and networks as well as the human-computer interactions. The third level emphasises how the individuals and IT-systems are embedded in the organisational structure.

The seven categories are gleaned from work in the field of systems science on the epistemological view of the world as a system. The categories can, with due justification, be said to encompass the breadth of human and natural activity, albeit at a high abstraction level (Beer, 1979, Checkland, 1981, Skyttner, 2001, von Bertalanffy, 1968). For each category an example is given to illustrate the type of issues that are covered by the category and how they can be operationalised in the form of questions on the personal, technical and organisational levels, respectively. The examples are from projects at the eHealth Institute.

2.3.1 Goal-seeking

As organisations consist of a number of different parts and hierarchical levels there is often an amalgamation of goals that have to be defined. The structure of some goals are more overarching and static and will be relevant to the activity for a long time, whereas others are more mobile and changeable. It is important to consider organisational goals from different hierarchical levels, as well as making both the formal and informal organisation goals visible. (Boulding, 1985, Litterer, 1973, von Bertalanffy, 1968).

Project: Evaluation of a web-based customer service application for 3D-mediated advice-giving on pharmaceutical products	Status: Planned
Personal	What are the personal goals of the pharmaceutical counsellors when using the service?
Technical	What technical equipment and software will be required for users of the service?
Organisational	Can the introduction of this service broaden the range of questions that can be handled over distance?

2.3.2 Hierarchy and relations

In most organisations it is possible to identify different hierarchical levels and specialised functions. Elucidating these structures and the distribution of responsibility among them creates important conditions for the evaluation. Since the holistic perspective is a matter of cooperation and mutual relations between the parts, the whole cannot build on a collection of autonomous elements. In this part of the evaluation the external and internal boundaries of the organisation are made clear. (ibid)

Project: Evaluation of a distance learning course for medical professionals	Status: Completed
Personal	What is the work climate like in the student groups?
Technical	What is the structural relationship between the different technical solutions that are used in examination with regard to authentication and security issues?
Organisational	How are responsibilities and task distributed among the persons involved in running and administering the course?

2.3.3 Differentiation and entropy

When the parts of an organisation begin to specialise to be able to perform certain tasks, the redundancy (surplus) as well as the entropy of the system decreases. In the long run the result of this differentiation is that the various parts become more dependent on each other while simultaneously, to some extent, the total efficiency of the system increases. A balanced degree of work distribution creates stability and safety but too much differentiation reduces the flexibility of the organisation and may make reality vulnerable. (Boulding, 1985, Litterer, 1973, Simon, 1969, von Bertalanffy, 1968).

Project: Evaluation of a web-based monitoring system for home care patients	Status: Underway
Personal	What skills do the various professionals who are in contact with the patients have?
Technical	What alternative technical systems are available should the primary system fail?
Organisational	What routines are in place for the training of the different categories of professionals involved in running the system?

2.3.4 Inputs

The input category makes visible the resources and conditions available to implement a certain process (Boulding, 1985, Litterer, 1973, von Bertalanffy, 1968). This comprises human resources like the knowledge, competencies and experiences of the individuals involved in the process. It is also important to study the attitudes and motivations of the people involved, as well as their demands, expectations and fears with regard to the process. The physical circumstances include the ergonomic design of the work place and the access to work tools. The IT environment comprises hardware, software and network resources as well as IT security.

Project: Evaluation of a web-based customer service application for 3D-mediated advice-giving on pharmaceutical products	Status: Planned
Personal	How many real life birthing sessions has the user participated in?
Technical	What technical capabilities are there in the hardware to sense and react to the physical actions performed by the midwifery student?
Organisational	Is the simulator training co-ordinated with the clinical training?

2.3.5 Transformation process

The transformation process is the actual interaction where the resource inputs are transformed into products and services (ibid). This kind of interaction may take place between, for example, management – employee, employee – customer or human – computer.

Project: Evaluation of a health information web site	Status: Completed
Personal	How does the service support the user's knowledge learning process?
Technical	In what way do the technical features of the service support the user's learning?
Organisational	Is the organisation well-adapted to support the purpose of increasing its users' knowledge on health issues?

2.3.6 Outputs

The outputs are a result of the processes taking place in the organisation and are usually synonymous with the services or products it delivers (ibid). They thus occupy a central position in following-up systems. In an evaluation it is important to remember that there may be several different goals for what may be considered a good result. One important question that has to be asked is whether the results created in the evaluation of the outputs will be effectively fed back into the activity.

Project: Evaluation of a pharmaceutical telephone-based call centre service	Status: Completed
Personal	How satisfied are the clients with the information they have received?
Technical	What level of audio quality is delivered?
Organisational	How can the average call time be optimised?

2.3.7 Regulation

Regulating the parts and process included is necessary for directing the transformation process towards the goals formulated. Regulation takes place synchronously and continuously on three levels. The first “primary” regulation is chiefly based on policy, norms and fixed rules in the organisation and its activities. The second or “secondary” regulation may use parts of the feedback occurring in a continuous evaluation in order to better adapt the contents to external and internal demands. The third “adaptation” level is attained when the organisation has been made part of the environment by opening up new paths and goals for its activities. All the three levels are essential to an open system, because static goals would mean that the system dies of itself. (Boulding, 1985, Churchman et al., 1969, Litterer, 1973, von Bertalanffy, 1968). The result and the efforts made around the evaluation may be regarded as the regulation of systems and partial systems, especially on the second and third levels, which makes the evaluation very important for maintaining the vitality of the system.

Project: Evaluation of an information exchange system for the co-ordination of activities for special needs kids	Status: Planned
Personal	In what way can the system users contribute to the way the service is developed?
Technical	In what way can the features of the system be adapted to suit the individual preferences of users?
Organisational	How are proposals for improvements and changes managed?

2.4 Evaluation research

Evaluation is an important control instrument in all activities, especially in goal-oriented and decentralised ones, where a clear feedback both to decision-makers and to individuals in the operative activity is required. There are a number of different models and strategies in the evaluation field, some of which models evaluate products and others processes (Langerth Zetterman and Lindblad, 2001). Depending on which type of evaluation model is used, the possibilities of different interested parties to have their needs and interests elucidated and noticed are affected (Guba and Lincoln, 1989, House, 1980, Madaus and Kellaghan, 2000, Patton, 1986). It is not uncommon that evaluations focus more on models and methods than on a theoretically founded research methodology, and frequently criteria set up in advance are applied. In a comparison with what has been mentioned above the SUV model may be positioned as a theoretically founded model, which gives the evaluator the freedom to choose and create new criteria in both product- and a process-oriented evaluations.

There are many points of contact between evaluation and research (Karlsson, 1999). An essential overarching similarity is that researchers and evaluators use the same methods for data gathering and analysis. The term *evaluation research* is relatively frequently used in describing the application of scientific research methods in evaluation contexts. However, there are differences between evaluation and research. One such difference is that evaluation often takes place in a task framework that is narrower than the theory frame used by researchers. This means that the evaluator is under stricter control than the researcher. A further difference is that the evaluation task often involves assessment, which is not expected of research. SUV enables both of these directions depending on which approach is employed.

An evaluation effort based on systems science theory broadens the issue and opens up to a more analytical, critical and reflective approach. Theory-based evaluation has been recommended and developed by Chen and his colleagues in reaction to the more product-oriented evaluation tradition. If the theoretical foundation is too narrow, there is a risk that the approach of the evaluation is confirmatory rather than descriptive and explanatory (Smith, 1994). Closely related criticism has been vented by Patton (1989), who argues that theory-based evaluations have gone too far in reducing the focus on and generalisation of causal connections, criticism that he somewhat tones down in a later work (Patton, 1996). It is important to emphasise that evaluation according to the SUV model does not focus on causal connections but rather on pointing to a number of possible influential factors.

The evaluation field contains basically three directions. The first is ***goal-result evaluation.*** This means evaluating towards goals in terms of measurable results. The process is viewed as an implementation of decisions which need not be motivated. The results are often presented as quantitatively measurable results. The second direction is ***process evaluation***, which is conducted against criteria for what characterises a good process. The interest is geared towards the implementation, for instance in order to assess work

methods and contents. The result is assumed to be positive if it can be established that the implementation corresponds to the requirements agreed on. The third direction is ***interactive evaluation***, which is characterised by the participation of the interested parties. Questions on the implementation are formulated on the basis of what different interested parties want to know. The interest focuses on what different interested groups consider important to evaluate from their various perspectives. Interactive evaluation forms a pattern developed in the 1980s and 1990s, by which cooperation between different interested parties increases the relevance of the evaluation questions and results and strengthens their influence (Cronbach et al., 1981). The goal of the SUV model is to thoroughly involve all the interested parties' experiences and ideas.

Earlier on, the idea prevailed that evaluation meant being able to stay neutral vis-à-vis the various interests by making objective measurements of goals which everyone was assumed to share. Today most evaluation researchers agree that this is not feasible – possibly except for narrowly delimited situations or relatively uncontroversial issues. It is probably more usual to be faced with different views on what should be examined and against which criteria this should be done. In a situation like this it is important to find forms which may increase the chances of various groups to make themselves heard. In the SUV model goal-seeking is done at an early stage and continues throughout the evaluation process, one important principle being to view the goals from different system levels. All the individuals involved should be offered the opportunity to participate in goal-seeking. The object is to be able to make formal and informal as well as short- and long-term goals visible.

According to Franke-Wikberg (1992), the evaluation aim is formulated primarily from one ***vertical*** and one ***horizontal*** perspective. The first vertical evaluation, "Type 1", is top-down, while vertical evaluation, "Type 2", is bottom-up. Horizontal evaluation takes place on the same organisation level. An example of this is a personnel group's internal evaluation, where the personnel "own" the results, deciding among themselves what they want to convey to other levels within the organisation. To these three levels may also be added a level outside the organisation consisting of various external interested parties. There is reason to point out that evaluation according to these different descriptions can function in parallel. However, confusion often arises or attempts to hold up one type as the only right one (Karlsson, 1999). The SUV model enables formulating the aim from all of these starting-points, which is often an advantage, since it sheds light on the issue to be evaluated from different perspectives.

3. Methodology

The SUV methodology is based on a continuous evaluation process whose driving-force is the wish of the interested parties to develop their own activities. What are in several evaluation methods termed process and product are studied from several points of view on the basis of a holistic approach. By continuously viewing the strengths, deficiencies and connections of the evaluated systems major and minor changes may be carried out continuously.

The five basic principles of the SUV methodology may be summarised as follows:

- A holistic approach
- All interested parties take an active part in the evaluation
- A flexible model enables the use of several different theories and methods
- The result of the evaluation is used for improving and developing activities
- A continuous evaluation process

The evaluation model described here is valid for use in both quantitative and qualitative approaches and can be incorporated in inductive and deductive studies. An example of deductive use is to construct a questionnaire based on the SUV categories and levels. SUV also encourages to building up explanations via an inductive approach in a continuously explorative process, which to some extent reminds of the theory-generated approach called Grounded Theory as formulated by Glaser & Strauss (1967). An example of inductive use is the post factum categorisations of previously performed evaluations not using the model, see 3.2.

The SUV model is not restricted to any *one* method for gathering data. It is a methodology within whose framework any and all methods deemed appropriate for the evaluation at hand can be used. Moreover, the use of multiple methods, used simultaneously, in parallel or sequentially; is explicitly encouraged for the purpose of gaining a multi-perspective view of the organisation/activity under scrutiny. A survey tool has been developed for questionnaire construction, web based delivery of surveys and compilation and analysis

of results. The survey tool is used exclusively for quantitative studies and while it serves as a good example of how the SUV principles can be put to practical use, it represents only one out of many methods that can be used with SUV. The choice of methods depends on the purpose of the evaluation and should be given careful consideration at the outset of the evaluation project. Figure 2 presents an overview of the conceptual foundations for the SUV methodology and its inclusive meta-structure within which any suitable method of data inquiry and any suitable approach for data analysis can be used.

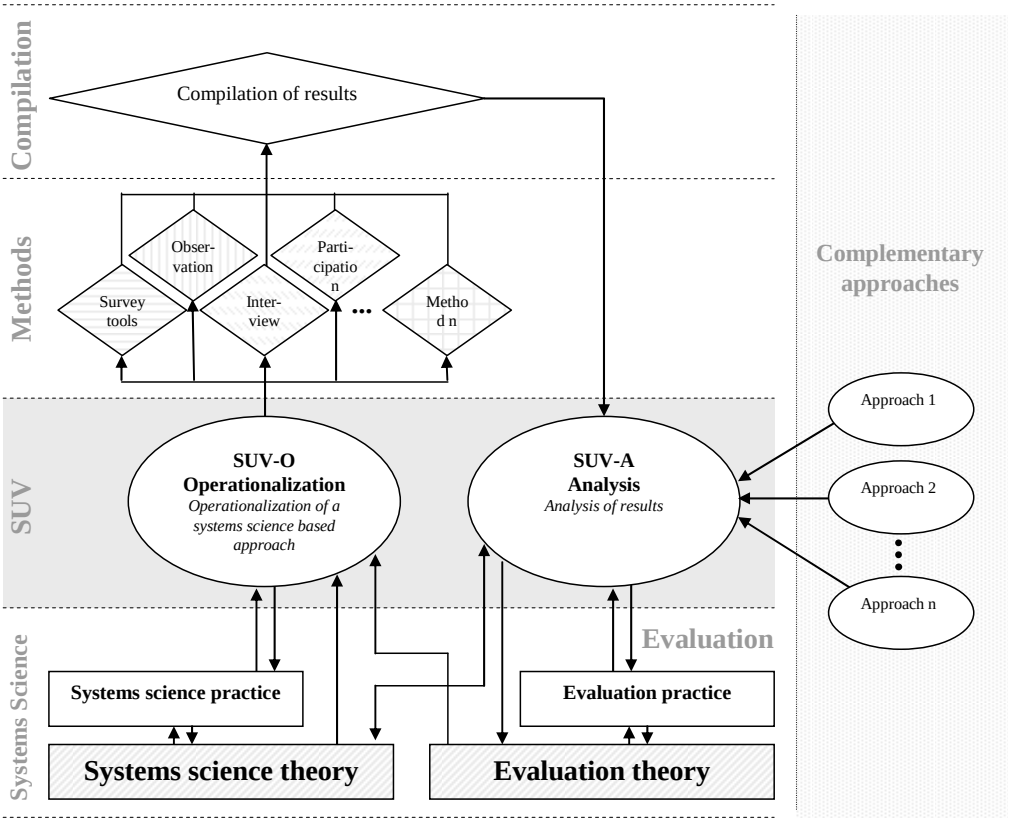

Figure 2: SUV is based on Systems Science Theory and Practice. The figure also illustrates the independence of method.

Figure 3 describes the process of an evaluation study according to the SUV methodology. The iterative quality in the process of evaluation is emphasized with the resulting measures and innovations serving as input in the next cycle of evaluation. In a large scale evaluation project, several parallel processes of the kind illustrated in the figure can be active at the same time starting and ending at different points in time. Going backwards and forwards between sub-steps in an iterative fashion will (and should) also occur. All the processes of evaluation should, however, follow the same general progression starting with introductory delimitation and problematisation and ending in proposed measures and innovations.

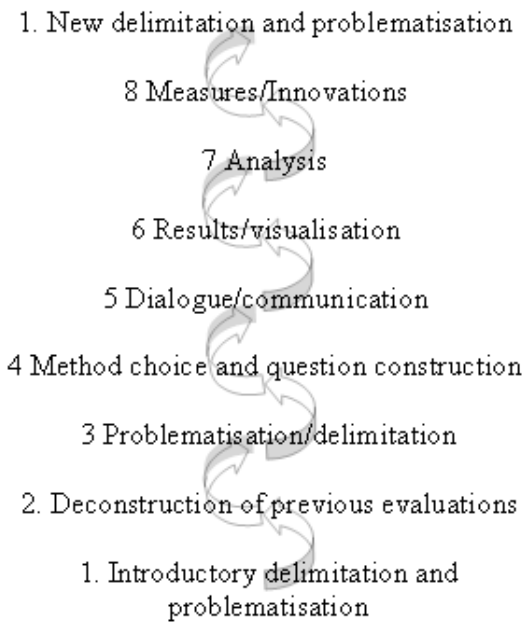

Figure 3: SUV, a continuous evaluation process

3.1 Step 1 – Introductory delimitation and problematisation

This step includes an introductory discussion about what is to be evaluated, what is the background of the task and for what context is the effort planned. Preliminary system delimitation is made in the form of a sketch of what and who should be included in the evaluation. The evaluator should also formulate the overarching goal as clearly as possible and make a list of what is considered to be most relevant to the evaluation. It may also be useful to reflect on evaluator's pre-understanding and how it may affect the planning of the study.

The following questions may serve as guidelines during the introductory phase:

- Why should there be an evaluation?
- What knowledge is desired?
- Which interested parties are involved in the evaluation?
- What should to be the role of the evaluator?
- How should the evaluation be conducted?
- How should the evaluation be utilised?

3.2 Step 2 – Deconstruction of previous evaluations

It is possible, and often also beneficial, to feed back previously conducted evaluations into the SUV model, and these results can then be sorted into and interpreted within the categories the model includes. Even if this kind of revised categorisation cannot be absolute, it may give indication on how accurately the previous evaluations have covered the entire problem area, and also on which previous questions may be important to follow up.

3.3 Step 3 – Problematisation and delimitation

The third step in the model is to make a more exact formulation of problems. By determining scope and level the system will be delimited, the choice of what to evaluate depends on the goal of the task. Anything outside the system boundaries is defined as the environment and will not be included in the evaluation. However, it is important to keep in mind that there is a continuous interaction between the system and its environment. The evaluator's list of what is considered to be most relevant to the evaluation (see step 1) can be refined and completed. Following steps can be used as guidelines in this phase:

3.3.1 The evaluation question

Formulate the evaluation questions as clearly as possible; both the overarching problematisation and more specific questions.

3.3.2 Delimitation of the system

Set clear boundaries for the system and specify what is to be included in the evaluation within these boundaries. It is also important to clarify which individuals are to be involved in the evaluation.

3.3.3 Safeguard the three levels

Make sure that the three levels, *individual, technology* and *organisation*, are represented in the evaluation.

3.3.4 Scope and timeframe

Clarify what resources are available in the form of individual effort, time, and access to respondents. Based on the evaluation's scope and available resources, make a time schedule for the evaluation process.

3.4 Step 4 – Method choice/Question construction

The SUV method allows for a range of different forms of data gathering, such as observation, focus group, interview, questionnaire and document analysis. In the same way as the system delimitation, the choice of appropriate methods depends on the evaluation goal and also on questions that are defined in step 3. Using various methods in the same evaluation usually makes the study more refined and valid.

The seven categories and three levels in the SUV circle are used as a framework for methodological design and also for question construction. Preferably, questions should be constructed for every single one of the fields in the SUV circle. Such full coverage will indicate that the evaluation takes into account all the important properties that characterize a complex system.

3.5 Step 5 - Dialogue, communicating questions

In SUV model, it is essential that the evaluation questions as well as obtained results are directly communicated to the participants. To enable a more formative and communicatively oriented evaluation, asynchronous and synchronous discussion conferences can be used to continually interpret and discuss findings; this approach can also contribute to the evolving evaluation method.

The web-based questionnaire tool contains several functions to convey questionnaires to respondents. Among the functions may be mentioned advance notice, automatic reminder and a report function for processing the gathered data.

3.6 Step 6 – Visualisation of data

In this step, the gathered empirical data is compiled and presented in suitable ways, both quantitatively and qualitatively. The previous categorization of questions into SUV categories and levels, can be used to structure the results in preparation for their subsequent analysis. One example of such structuring can be seen in figure 4, where the results from multiple choice questions are presented for each of the fields in the SUV circle (this functionality is included in the survey tool previously presented).

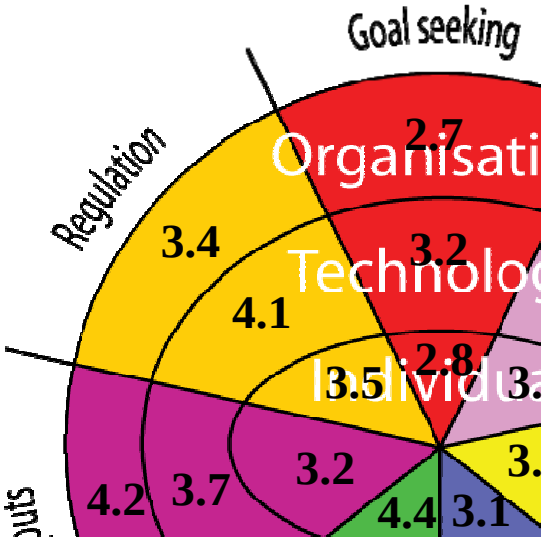

Figure 4: Visualisation of the mean values from multiple-choice questions.

3.7 Step 7 – Analysing the result

Assessment or analysis is the central task of evaluation (Karlsson, 1999). The analysis is often the least common denominator in different definitions of evaluation where the evaluation method constitutes the framework of the analysis. At the analysis/assessment of the result the evaluator may adopt a more or less objective or subjective attitude. The objective attitude includes general evaluation principles setting up norms, while the result in the subjective attitude is more dependent on the interpreter. It is also possible to choose the dialogically inter-subjective standpoint without getting stuck in objective or subjective reasoning (ibid.). The analysis may instead be shaped after a discussion between evaluators who try to get an insight into and understand each other's perspective. Consequently, this qualitative analysis takes place in a dialogue with different interested parties.

Comparison is one normative analysis approach used in evaluation contexts. There are basically three ways of viewing results. The *first* is to make a comparison with an explicitly defined reference point of some kind, such as a goal, idea or theoretical criterion. The *second* is comparing evaluation objects of the same kind, usually by ranking. The *third* is to compare transformation over time, in other words the development of the evaluation object. In SUV the starting-point of comparisons consists of the categories and levels stipulated in the model, which enables the use of all three approaches. A comparison emanating from its theoretically founded variables may be discussed using the following questions as starting-points:

- What qualitative differences exist within the various categories and levels?
- Which categories/levels received low and high values, respectively?
- Between which categories/levels are the differences the greatest or the smallest?
- How does the result change over time?

On the basis of this orientation any reasons behind the differences displayed are analysed. It is also possible to further refine the categories which have indicated deficiencies or those categories which may serve as positive examples.

3.8 Step 8 – Taking action

At the very planning of an evaluation the issue of how to utilise the obtained results should be discussed. On the basis of the analysis in step 7 an assessment can be made of which categories require further scrutiny or immediate action to be taken. The SUV model makes it possible to start this process at an early stage and to continuously adjust and attend to various phenomena and conditions, even during the evaluation process.

The fact that evaluation results are not always utilised for further development of the organisation and its activities may be due to various reasons. There may be external environment or background circumstances which affect the structure and implementation of the evaluation. It may also be due to how the evaluation is conducted and whose questions have been actualised. Further reasons may be that the results of the

evaluation are not considered relevant or that they are presented in a deficient way or that the intended receivers lack the knowledge, motivation and resources to operationalize the evaluation results (Karlsson, 1999).

There are various measures that can be taken in order to support the utilisation of the evaluation results. To start with, it is essential to create good representativeness and maintain a active dialogue with those who are in various ways involved in the evaluation process. Another important contribution to enhance the positive attitudes is to emphasise the importance of communication and interaction among the participants, The results should be reported in a way that makes them available and interesting. In some cases, it may be important to raise the general competence level in the organisation, for instance by educating personnel and by encouraging discussion and critical reflexion in practice. An important measure for enhancing positive attitudes is to set aside resources which enable the application of those evaluation results which are considered important (Karlsson, 1999).

4. Web-based questionnaire tool

The questionnaire application was constructed by the authors to create a tool for question generation, data gathering and analysis that would make the SUV method accessible and usable also for the non-systems theory savvy user. The questionnaire instrument includes the possibility of choosing between different question types and between different formats, as shown in Figure 5. The questions may be divided into background questions, SUV questions and open questions.

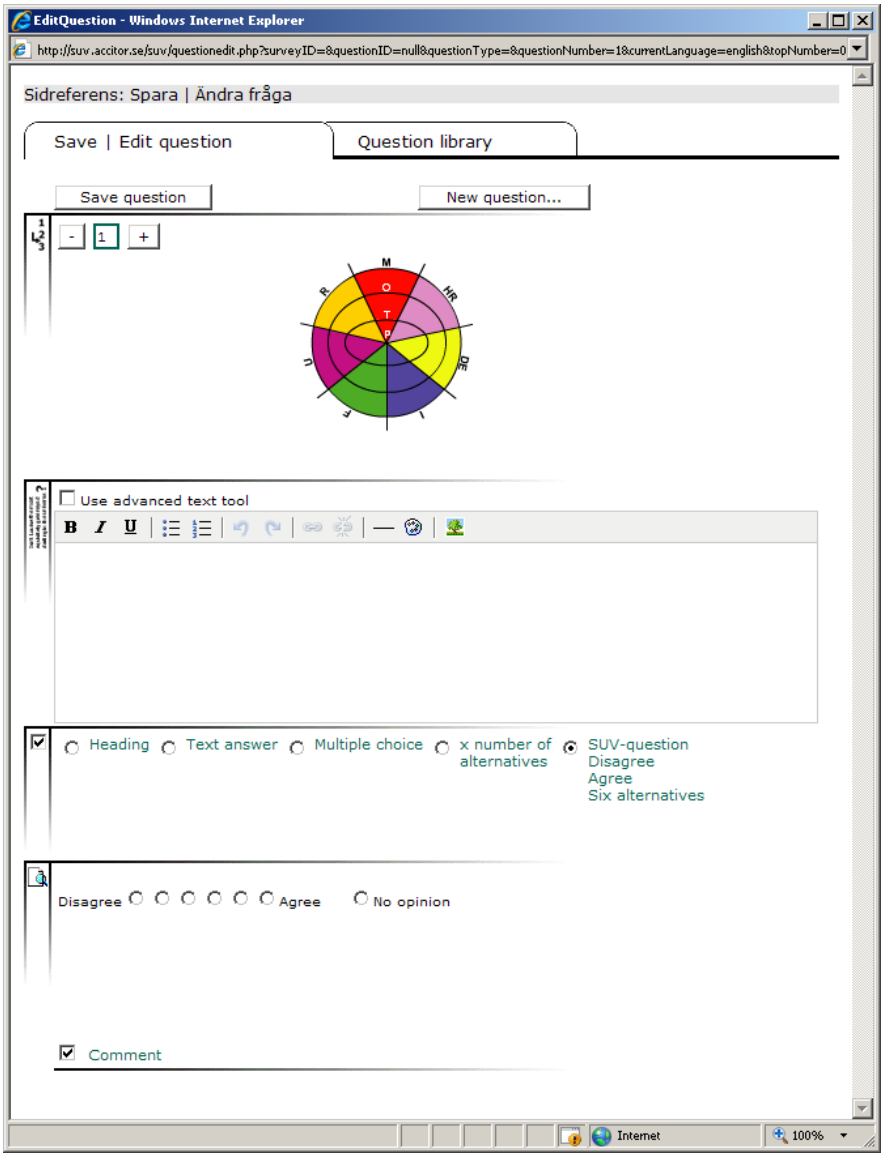

Figure 5: Different questions types and formats in the web-based questionnaire tool.

Background questions are used in the statistical analysis to see the differences between one group and another. When searching for possible explanations it is important to consider which variables are relevant to the current evaluation. It is common to ask questions on sex, age, education, computer experience etc., but every organisation has its own special conditions which may be of importance to use as background variables.

The multiple questions provided with a six-grade answering scale are termed SUV questions and are formulated so that each question belongs to a specific field in the SUV graph. The user marks in which field the question has been constructed and the tool states the frequency of questions in each field, as shown in Figure 6. This illustrates the distribution of questions among the different fields and also among the different categories and levels.

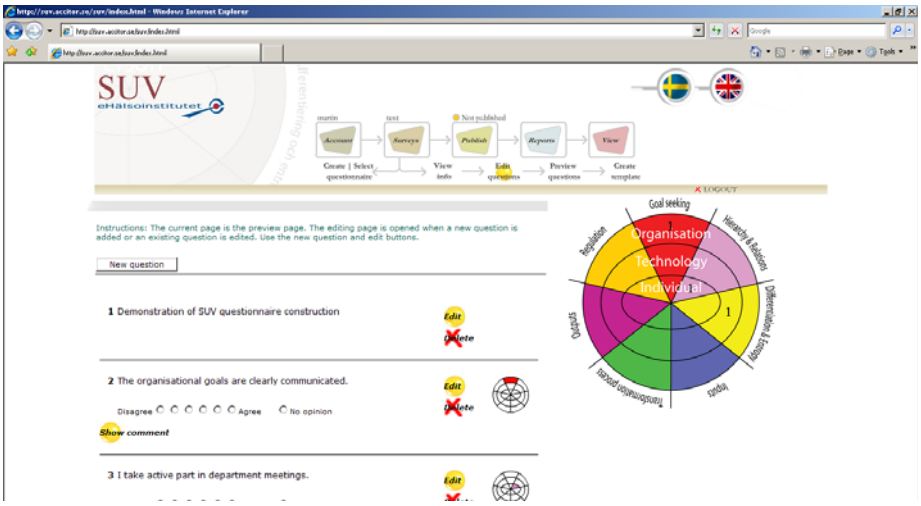

Figure 6: Examples of questions and of how the tool visualises their location and distribution within each question field.

There is also a possibility for the respondent to add further views on the SUV questions. In questionnaires distributed in paper form such a possibility is often offered at the end of the questionnaire, whereas the digital SUV tool makes it possible to attach comments immediately after each SUV question. The advantage is that this enables catching the spontaneous response, which may otherwise be lost if the respondent has to wait till the end of the questionnaire.

To create a clear structure for result presentation and analysis according to the model there is a specially adapted result presentation module in the program. The result presentation is based on the SUV graph and on the collected quantitative data which is used in the statistical calculations. The arithmetic mean as well as median value of the answers to the SUV questions can be shown in each in each field, category and level, respectively. The result is also presented in greater detail by including the frequency distribution within the various answer alternatives in a table and visualising them in a suitable graph. For background variables either a bar chart or a histogram is used, depending on whether it is a discrete or continuous variable.

If the number of respondents is sufficient the connection between background variables (groups) and SUV questions may be studied by means of a chi-two test (χ^2) (Curwin and Slater, 2004). It is important to emphasise that the statistical analysis is mainly used to looking for possible significant connections in the gathered material, the primary aim thus not being to draw general conclusions that are supposed to be relevant to all organisations.

The questionnaire tool also comprises a question archive where one may simply include and edit pre-constructed and pre-categorised questions. Usually, questions suitable for elucidating the various categories can be reused after a slight modification

5. Experiences and further elaboration

Constructing questions with the support of the SUV model does not automatically result in a holistic evaluation. It is certainly quite possible to use the principles of question construction to a smaller extent, but for SUV to achieve reliable reality the guideline is that all categories at all levels are well covered, as our pilot

evaluations demonstrate. It is also a clear advantage if the evaluation is conducted on a number of occasions and preferably complemented by several different data gathering methods. This complementation might suitably be performed after the first more explorative study. To continuously measure the status of the activity according to the same evaluation model enables a longitudinal comparison on the basis of a common theoretical reference frame.

Formulating questions that cover all categories and levels in the model may sometimes be difficult and tedious, and it certainly requires more effort of the evaluator. On the other hand, the proposed structure also provides that all parts of importance to the activity are included in the evaluation, and at the same time the holistic view of the problem area becomes the guiding principle. In our opinion, a specially adapted model and methodology are required to fully acknowledge the complexity of organisations where the use of information and communication technology forms an integral part of core activities.

As mentioned above, the SUV methodology can be used with several different methods. The same broad approach also applies to the analysis. The result may reflect a number of theories depending on the perspective/s adopted. A theory discussion is a necessary element if the method is to be used in a strictly scientific study.

Within the next few years the eHealth Institute is going to use SUV as its evaluation model in the projects conducted by the institute. The method will most likely be developed further in evaluations of e-related projects.

Acknowledgements

Thanks to Prof. Göran Petersson at eHälsoinstitutet for his continuing support for the SUV evaluation project. Our gratitude to the reviewers for their valuable comments.

References

Backlund, A. (2002) The Concept of Complexity in Organisations and Information Systems. *Kybernetes,* 31(1).
Barroso, J., Sandelowski, M. & Voils, C. I. (2006) Research Results have Expiration Dates. Ensuring Timely Systematic Reviews. *Journal of Evaluation in Clinical Practice,* 12(4)**,** 454-462.
Beer, S. (1979) *The Heart of Enterprise,* New York, John Wiley.
Bohm, D. & Peat, D. (2000) *Science, Order and Creativity,* London, Routledge.
Boulding, K. E. (1985) *The World as a Total System,* Beverly Hills, Sage Publications.
Brandt, P., Jokela, P., Karlsudd, P., Révay, P. & Wennberg, L. (2003) Evaluation model for e-phenomena – a case study. *Proceedings of the MicroCAD 2003 International Scientific Conference 6-7 mars 2003.* Miskolc, Hungary.
Checkland, P. (1981) *Systems Thinking, Systems Practice*, New York.
Churchman, C. W., Ackoff, R. L. & Arnoff, E. L. (1969) Analysis of the Organization. IN A, L. J. (Ed.) *Systems, Control and Adaptation Volume II (1980).* NewYork, John Wiley & Sons.
Cohen, S., Kaimenaki, E. & Zorgios, Y. (2007) Assessing IT as a Key Success Factor for Accrual Accounting Implementation in Greek Municipalities. *Financial Accountability & Management,* 23(1)**,** 91-104.
Cottam, R., Ranson, W. & Vounckx, R. (2005) Life and Simple Systems. *Systems Research and Behavioral Science,* 22**,** 413-430.
Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, D. C., Hornik, R. C., Philips, D. C., Walker, D. F. & Weiner, S. S. (1981) *Toward Reform of ProgramEvaluation: Aims, Methods, and Institutional Arrangements,* San Francisco, Jossey-Bass Publishers.
Curwin, J. & Slater, R. (2004) *Quantitative Methods,* London, Thomson Learning. de Shazer, S. (1994) *Words Were Originally Magic,* New York, W W Norton & Company.
Edwards, M. G. (2005) The Integral Holon – A Holonomic Approach to Organisational Change and Transformation. *Journal of Organizational Change,* 18(3).
Franke-Wikberg, S. (1992) *Utvärederingens mångfald: några ledtrådar för vilsna utbildare,* Stockholm, FoU-enheten, UHÄ.
Glaser, B. & Strauss, A. L. (1967) *The discovery of grounded theory: Strategies for qualitative research,* Chicago, Aldine.
Guba, E. G. & Lincoln, Y. S. (1989) *The Fourth Generation Evaluation,* Newbury Park, Sage Publications.
Hall, A. D. & Hagen, R. E. (1969) Definition of System. IN Litterer, J. A. (Ed.) *Systems, Control and Adaptation, Volume II (1980).* NewYork, John Wiley & Sons.
Hartford, K., Carey, R. & Mendonca, J. (2007) Sampling Bias in an International Internet Survey of Diversion Programs in the Criminal Justice System. *Evaluation & the Health Professions*.
Harun, M. H. (2002) Integrating e-Learning into the Workplace. *The Internet and Higher Education,* 4**,** 301-310.
House, E. R. (1980) *Evaluating with quality,* London, Sage Publications.
Jackson, E.T & Kassam, Y. (1998) *Knowledge Shared: Participatory Evaluation in Development Cooperation*, Ottawa, IDRC/CRDI.

Jacucci, E. & Hanseth, O. (2006) Taking Complexity Seriously in IS Research. *Information Technology and People,* 19(1)**,** 5-11.
Karlsson, O. (1999) *Utvärdering – mer än metod. Tankar och synsätt i utvärderingsforskning,* Stockholm, Norstedts tryckeri.
Katz, D. & Kahn, R. L. (1980) The Definition and Identification of Organisations. IN A, L. J. (Ed.) *Organizations, Structure and Behavior.* NewYork, John Wiley & Sons.
Koestler, A. (1967) *The Ghost of the Machine,* London, Arkana.
Langerth Zetterman, M. & Lindblad, S. (2001) Learning about e-learning. A strater about Internet discourses and borderless education. *NFPF:s 29. congress. Pedagogikens mångfald. Lärande inför och utanför institutionerna.* Stockholm, Sweden.
Litterer, J. A. (1973) *The Analysis of Organizations,* New York, John Wiley & Sons.
Madaus, G. M. & Kellaghan, T. (2000) Models, Metaphors and Definitions. IN Stufflebeam, D. L. (Ed.) *Evaluations Models. Viewpoints on Educational and Human Services Evaluation.* Hingham, MA, USA, Kluwer Academic Publishers.
Marine, R. J. (2002) A systems freamework for evaluation of faculty web-work. *New Directions for Institutional Research year,* 63-72.
Mulej, M. (2007) Systems Theory: A Worldview and/or a Methodology Aimed at Requisite Holism/Realism of Humans' Thinking, Decisions and Action. *Systems Research and Behavioral Science,* 24**,** 347-357.
Patton, M. Q. (1986) *Utilization-Focused Evaluation 2nd edition,* London, Sage Publications.
Patton, M. Q. (1989) A context and boundaries for a theory-driven approach to validity. *Evaluation and Program Planning,* 12 (4)**,** 375-377.
Patton, M. Q. (1996) *Utilization-Focused Evaluation: the new century text,* Thousand Oaks, CA, Sage Publications.
Petit, B. & Olsen, H. (1994) *Om svar anhålles. En bok om interaktionistiskt förändringsarbete Andra reviderade upplagan,* Stockholm, Mareld.
Sandelin, B., S & Sarvimäki, A. (2000) Evaluation of patient-focused health care from a systems perspective. *Systems Research and Behavioral Science,* 17:6**,** 513-525.
Schöjdt, B. & Egeland, T. (1994) *Från systemteori till familjeterapi,* Lund, Studentlitteratur.
Simon, A. H. (1969) The Architecture of Complexity. IN A, L. J. (Ed.) *Systems, Control and Adaptation Volume II (1980).* NewYork, John Wiley & Sons.
Skyttner, L. (2001) *General Systems Theory. Ideas & Applications,* Singapore, World Scientific.
Smith, N. L. (1994) Clarifying and expanding the application of program theory-driven evaluation. *Evaluation Practice,* 15(1)**,** 83-87.
von Bertalanffy, L. (1968) *General System Theory,* New York, George Braziller.
Zhao, Y., Yagi, Y., Nakajima, I. & Jujozi, H. (2002) IP Telephony – New Horizon for Telemedicine and e-Health. *Journal of Medical Systems,* 4(4)**,** 277-283.

