

# Information Systems Evaluation

## – adding process descriptions to six evaluation types

Stefan Cronholm  
Department of Computer and Information Science  
Linköping University  
E-mail [scr@ida.liu.se](mailto:scr@ida.liu.se)

### Abstract

All around the world there is a huge amount of money invested in IT. It is therefore important to evaluate the return on the investment. Evaluation is complicated and consequently there are a lot of proposals for how to evaluate IT-systems. Several of these proposals reside on the theoretical level. The aim of this paper is to be more concrete and discuss evaluation in terms of “how to do”. This paper builds on a evaluation framework presented by Cronholm & Goldkuhl (2003). In that paper, six generic types for evaluation are presented on an ideal typical level. In this paper, evaluation processes are suggested for how the generic types can be carried out in order to support the use of that framework.

**Keywords: Information Systems Evaluation, Evaluation Processes, Evaluation Strategies**

### 1. Introduction

All around the world there is a huge amount of money invested in IT (e.g. Seddon, 2001). It is therefore important to evaluate the return on the investment. Evaluation is complicated and consequently there are a lot of proposals for how to evaluate IT-systems.

Much of the literature on evaluation takes a formal-rational stand and sees evaluation as a largely quantitative process of calculating the likely cost/benefit on the basis of defined criteria (Walsham, 1993). These approaches are often developed from a management perspective and contain different measures that often are of harder economical character. One common criticism of the formal-rational view is that such evaluation concentrates on technical and economical aspects rather than human and social aspects (Hirschheim & Smithson, 1988). Further Hirschheim & Smithson maintain that this can have major negative consequences in terms of decreased user satisfaction but also broader organizational consequences in terms of system value.

There are also other evaluation approaches such as interpretative (e.g. Remenyi, 1999; Walsham, 1993) and criteria-based. Interpretative approaches often view IT-systems as social systems that have information technology embedded into them (Goldkuhl & Lyytinen, 1982). Criteria-based approaches are concerned with identifying and assessing the worth of programme outcomes in the light of initially specified success criteria (Walsham, 1993). The criteria used are often derived from one specific perspective or theory.

The papers discussed above are discussing the concept of evaluation on a theoretical level. The aim of this paper is to be more concrete and discuss evaluation in terms of “how to do”. In Cronholm & Goldkuhl (2003) evaluation is also discussed on a theoretical level. In that paper six generic types for evaluation are presented on an ideal typical level. That paper can be viewed as a framework of evaluation and this paper will take these ideas one step further forward. The aim of this paper is to suggest processes for how the generic types can be carried out in order to support the use of this framework. The processes generated are grounded in a case study of a knowledge-based system in a Swedish local authority. The framework of

evaluation is briefly presented in section 2. In section 3 the case study is briefly described. Section 4 consists of knowledge sources for evaluation. In section 5 the evaluation processes are described. Finally, section 6 consists of my conclusions.

## **2. Description of six types of evaluation**

The evaluation processes that will be described in section 3 are based on six generic types of evaluation (cf Cronholm & Goldkuhl, 2003 for a fuller description of the six generic evaluation types). These types are derived from two strategies concerning *how to evaluate* and *what to evaluate*. Strategies concerning how to evaluate are presented in section 2.1 and strategies concerning what to evaluate is presented in section 2.2.

### **2.1. Strategies concerning how to evaluate**

We distinguish between three types of strategy:

- Goal-based evaluation
- Goal-free evaluation
- Criteria-based evaluation

The differentiation is made in relation to what drives the evaluation. Goal-based evaluation means that explicit goals from the organisational context drive the evaluation of the IT-system. The basic strategy of this approach is to measure if predefined goals are fulfilled or not, to what extent and in what ways. The approach is deductive. What is measured depends on the character of the goals and a quantitative approach as well as qualitative approach could be used.

The goal-free evaluation means that no such explicit goals are used. Goal-free evaluation is an inductive and situationally driven strategy. This approach is a more interpretative approach (e.g. Remenyi, 1999; Walsham, 1993). The aim of interpretive evaluation is to gain a deeper understanding of the nature of what is to be evaluated and to generate motivation and commitment (Hirschheim & Smithson, 1988). According to Patton (1990) the aim of goal-free evaluation is to:

- 1) avoid the risk of narrowly studying stated program objectives and thereby missing important unanticipated outcomes
- 2) remove the negative connotations attached to the discovery of an unanticipated effect: “The whole language of side-effected or secondary effect or even unanticipated effect tended to be a put-down of what might well be a crucial achievement, especially in terms of new priorities.”
- 3) eliminate the perceptual biases introduced into an evaluation by knowledge of the goals and
- 4) maintain evaluator objectivity and independence through goal-free conditions.

The basic strategy of this approach is inductive evaluation. The approach aims at discovering qualities of the object of study. One can say that the evaluator makes an inventory of possible problems and that the knowledge of the object of study emerges during the progress of the evaluation.

Criteria-based evaluation means that some explicit general criteria are used as an evaluation yardstick. The difference to goal-based evaluation is that the criteria are general and not restricted to a specific organisational context. That means that they are more generally applicable. There are a lot of criteria-based approaches around such as checklists, heuristics, principles or quality ideals. In the area of Human-Computer Interaction different checklists or heu-

ristics can be found (e.g. Nielsen, 1994; Nielsen, 1993, Shneiderman, 1998). What is typical for these approaches is that the IT-systems interface and/or the interaction between users and IT-systems act as a basis for the evaluation together with a set of predefined criteria. More action oriented quality ideals and principles for evaluation can be found in Cronholm & Goldkuhl (2002) and in Ågerfalk et al (2002).

## 2.2. Strategies concerning what to evaluate

All of the approaches goal-based, goal-free and criteria based are different ways and their primary message is *how* the evaluator should act in order to perform evaluation. Besides this “how message” it is also important to decide about *what* to evaluate. When evaluating IT-systems we can think of at least two different situations that can be evaluated. We differ between evaluation of IT-systems as such and evaluation of IT-systems in use. IT-systems can be viewed from many different perspectives. The framework for IT evaluation presented in Cronholm & Goldkuhl (2003) is not dependent on any particular perspective.

Evaluating IT-systems as such means to evaluate them without any involvement from users. In this situation there are only the evaluator and the IT-system involved. The data sources that could be used for this strategy is the IT-system itself and possible documentation of the IT-system (see Figure 1). How the evaluation is performed depends on the “how-strategy” chosen. Choosing to evaluate “IT-systems as such” does not exclude any of the strategies of “how to evaluate”. The evaluator could use a goal-based, goal-free or criteria-based strategy.

The outcome of the evaluation is based on the evaluator’s understanding of how the IT-system supports the organisation. This strategy is free from a user’s perceptions of how the IT-system benefits their work.

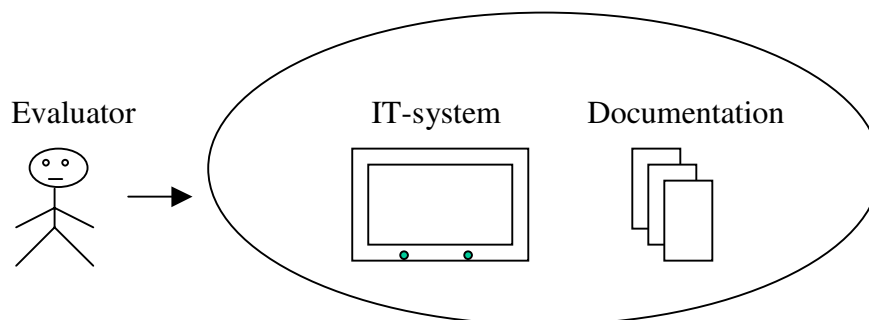


Figure 1. Two possible data sources for IT-systems in use

The other strategy of “what to evaluate” is “IT-systems in use”. Evaluating IT-systems in use means to study a use situation where a user interacts with an IT-system. This analysis situation is more complex than the situation “IT-systems as such” since it also includes a user, but it also has the ability to give a richer picture.

The data sources for this situation could be interviews with the users and their perceptions and understanding of the IT-system’s quality, observations of users interacting with IT-systems, the IT-system itself and the possible documentation of the IT-system (see Figure 2). Compared to the strategy “IT-systems as such” this strategy offers more possible data sources. When high requirements are placed on data quality the evaluator can choose to combine all the data sources in order to achieve a high degree of triangulation. If there are fewer resources to hand the evaluator can choose one or two of the possible data sources.

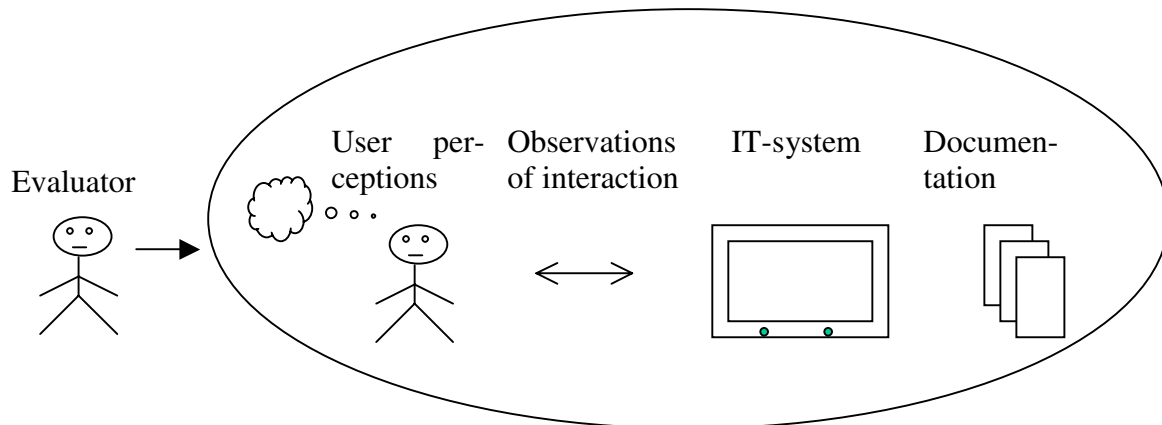


Figure 2. Four possible data sources for IT-system in use

An argument for choosing the strategy “IT-systems in use” is presented by Whiteside & Wixon (1987). They claim “... usability becomes a purely subjective property of the interaction between a specific user and the computer at a specific moment in time”. There are always subjective perceptions such as the user’s attitude towards an IT-system that are harder to measure.

How the evaluation of “IT-systems in use” is performed depends on the “how-strategy” chosen (see section 2). Ideally, it should be possible to choose any of the three strategies goal-based, goal-free or criteria-based when studying “IT-systems in use”. The outcome of this evaluation is not only based on the evaluator’s understanding of how the IT-system supports the organisation. It is also based on the user’s perceptions of how the IT-system supports their work.

### 2.3. Six generic types - a summary

From the strategies for how to evaluate and what to evaluate we derive a matrix consisting of the two dimensions “how” and “what” (see Table 1). The combination results in six different evaluation types. The aim of the matrix is to support different choices of how to perform an evaluation depending on the evaluation situation. In Goldkuhl & Cronholm (2003) each type is characterized. The characterization is done according to the criteria: main perspective, what to achieve knowledge about, data sources, deductive or inductive, who will participate and when to choose this type.

Table 1. The matrix of six generic types of information systems evaluation

	IT-systems as such	IT-systems in use
Goal-free evaluation	Type 1	Type 2
Goal-based evaluation	Type 3	Type 4
Criteria-based evaluation	Type 5	Type 6

### 3. Briefly about the case study

The case study has been performed in a Swedish local authority. The assignment was to carry out an evaluation of a knowledge-based IT-system. The aim of the IT-system is to support use and maintain information about methods for revision. In the case study all the six evaluation types have been applied, but not in a well-structured and articulated way. As an input to the case study, we had an idea of how to perform goal-free, goal-based and criteria-based evaluation. The evaluation processes presented in section 5 have been induced and reconstructed from the empirical data. Data have been gathered by studying the IT-system, reading documentation of the IT-system, observing users interacting with the IT-system and by means of interviews with managers and users.

### 4. Knowledge sources for evaluation

The question of how to perform the evaluation depends on which evaluation type is chosen. This choice has among other things to do with the available knowledge objects/subjects. Possible objects/subjects for gathering information are pictured in Figure 3. We have a business context consisting of users interacting with IT-systems. There is documentation of the IT-systems, such as requirement specifications, user manuals and technical descriptions. There are also business goals that normally exist in strategy documents or goal documents. In the business context there are also other actors, such as managers, system owners and other stakeholders. Outside the business context there are evaluation criteria. The reason for placing the criteria outside the business context is that they are general. This means that they are applicable for more than one specific organisation. Which of the objects/subjects is used depends on how the evaluation is to be performed (see section 2).

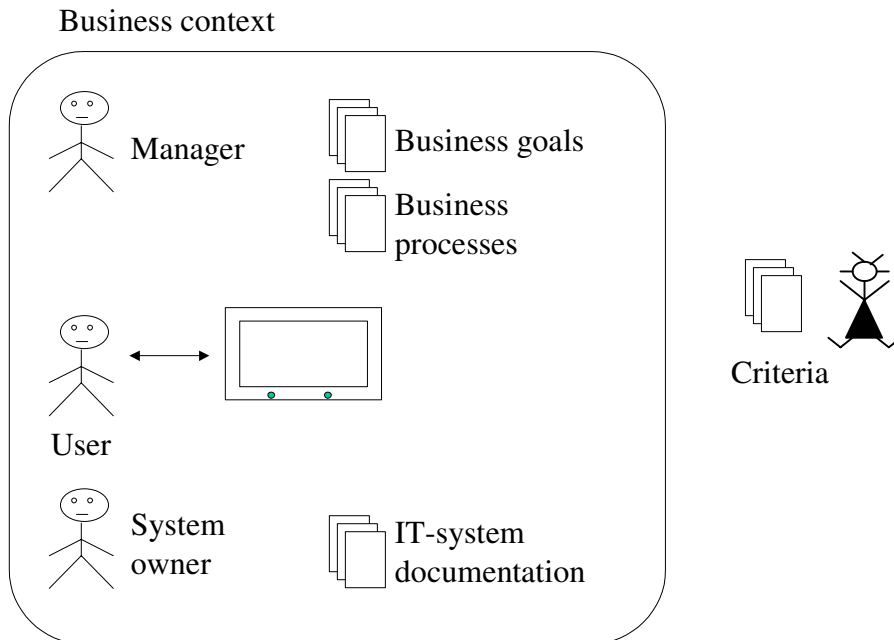


Figure 3. Possible objects/subjects in the evaluation context

For each object/subject identified there are one or several possible knowledge sources. A knowledge source is defined as where we can find information about an object or a subject.

For each knowledge source a data gathering method is proposed (see Table 2). For example, information about the object business processes could be found in the knowledge sources key personnel and business documentation. Since we could ask questions of key personnel and study them when they carried on business, the data gathering methods proposed are interviews and observation. We could also find information about business processes in business documentation. In this case we are gathering data by reading documents. In Table 2 there are for each of the object/subject proposed knowledge sources and data gathering methods.

Table 2. Overview of evaluation object/subject, knowledge sources and gathering methods

Object/ subject	Knowledge source	Gathering methods
Business process	Key personnel	Interview Observation
	Business documentation	Reading
IT-system	IT-system	Observation Exploration
	IT-system documentation	Reading
Business goals	Key personnel	Interview
	Strategic documents	Reading
Criteria	Criteria description	Reading
	The evaluation situation	Sensitivity, an open mind
IT-system in use	The user's interaction with the system.	Observation Interview
Users	Users	Interview Observation
	User documents	Reading

## 5. The Evaluation process

The processes of each evaluation type are depicted in Figure 4. If you follow the arrows you can see some of the activities are general for all the evaluation types and some are not. For example all evaluation types start with *plan the conditions* and continue with *describe the functionality* and *describe the users*. All evaluation types end with *draw conclusions*. The evaluation processes are described in a sequential way. Whenever needed, iterations can and should be undertaken.

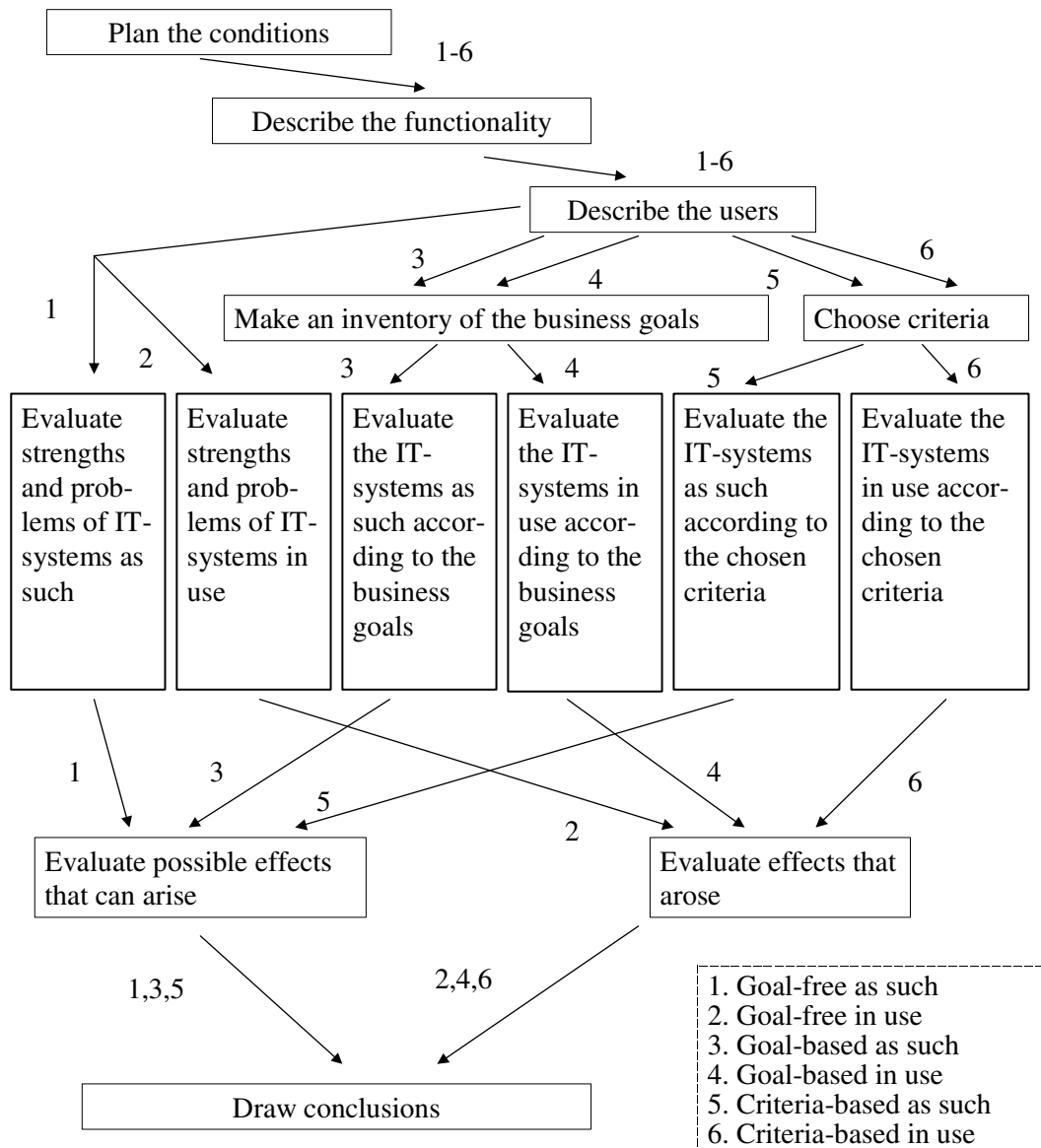


Figure 4. Process description

**Plan the conditions (evaluation type 1-6)**

For all the evaluation types we recommend to start with planning the conditions. Planning the conditions for the evaluation means deciding upon: the scope, the level of ambition, time, cost and resources. It also includes identifying the assigner, the performer, the data providers and the client of the result. Planning the conditions also mean to describe the business processes and an overview of the IT-system’s functionality. This stage is done in order to increase the evaluator’s pre-knowledge of the context. The planned conditions work as a base for choosing one or a combination of several evaluation types. Knowledge sources for this stage are managers and their assignees.

**Describe the functionality (evaluation type 1-6)**

After the conditions are planned, the next step is to describe an overview of the functionality of the IT-system. We suggest that the evaluator walks through the IT-system in an explorative way and describes its functionality briefly in order to be acquainted with the IT-system.

Besides studying the IT-system the evaluator should also gather information about the IT-system from reading documentation of the IT-system and interviewing the system owner.

**Describe the users (evaluation type 1-6)**

Choosing to study IT-systems as such (evaluation type 1,3,5) means that no users need be available. However, we think it is important if the evaluator even in these types can describe a user profile since the evaluation of the functionality of the IT-system is related to the users pre-knowledge. Describing the users means to describe the users' tasks, roles, experience, education and maturity. Information about users can be found in job descriptions or by interviewing managers and/or the system owner. Choosing to study IT-systems in use means that users are available. This means that the evaluator can gather additional information about users from observing user interaction and from interviewing them.

**Make an inventory of the business goals (evaluation type 3, 4)**

This stage is performed when a goal-based evaluation is chosen. The evaluator makes an inventory of the of the business goals. Business goals can be found in strategic documents or can be identified by interviews with key personnel.

**Choose criteria (evaluation type 5, 6)**

This stage is performed when a criteria-based evaluation is chosen. In order to evaluate the IT-system the evaluator chooses appropriate criteria. Which criteria and the number of criteria to be chosen depend on the conditions of the evaluation context. The selection depends also on the evaluation perspective (Cronholm & Goldkuhl, 2003). During the evaluation the criteria both support focusing and limits the evaluation. The criteria chosen are documented in a criteria list. Note, that the criteria can change over time. New insights into the evaluation situation could cause a need for new, rephrased or more precise criteria. The evaluator should therefore be sensitive and open-minded during the evaluation.

**Evaluate strengths and problems of the IT-system as such (evaluation type 1)**

In this stage the IT-system is evaluated in terms of strengths and problems. The evaluator approaches the IT-system with an open-mind and has no explicit criteria or business goals that govern the evaluation. The knowledge sources available in this stage are the IT-system and documentation of the IT-system. The identified strengths and problems are respectively related to each other during the final stage draw conclusions.

**Evaluate strengths and problems of the IT-system in use (evaluation type 2)**

At this stage, the evaluated in terms of strengths and problems, the difference between this stage and the stage above is that the strengths and problems are identified through observing interaction and interviewing users. During the observation the evaluator can use a technique called "think aloud" (Monk et al, 1993). The observation should be combined with follow-up interviews. The evaluation is made in an explorative way. We recommend that the users should be observed when they are working with real tasks in a natural setting. The knowledge sources available in this stage are the IT-system, documentation of the IT-system and users.

**Evaluate the IT-system as such according to the business goals (evaluation type 3)**

This stage means that the evaluator performs a goal-based evaluation of the IT-system. The evaluation can result in four outcomes when deciding if an IT-function meets a business goal. One outcome is that the IT-function supports a business goal. Another outcome is that an IT-function counteracts a business goal. A third outcome is that there is missing functionality and a fourth outcome is that there is functionality in the IT-system that is not related to any business goals. The knowledge sources available in this stage are the IT-system, documentation of the IT-system and business goals.



#### **Evaluate the IT-system in use according to the business goals (evaluation type 4)**

This stage means that the evaluator performs a goal-based evaluation of the IT-system in use. The difference between this stage and the stage above is that the evaluation also includes the users understanding if the IT-system meets the business goals or not. This means that the evaluator observes and interviews users. During the observation the evaluator could use the technique called “think aloud” (Monk et al, 1993). We recommend that the users should be observed when they are working with real tasks in a natural setting. As mentioned above the evaluation can result in four outcomes: the functionality supports the business goals, the functionality counteracts the business goals, there is a missing functionality for a business goal or there is functionality that is not related to any business goal. The knowledge sources available in this stage are the IT-system, documentation of the IT-system, business goals and users.

#### **Evaluate the IT-system as such according to the chosen criteria (evaluation type 5)**

This stage means that the evaluator decides if the IT-system’s functionality meets the criteria chosen. The evaluator simply uses the criteria as an instrument for judgement. Available knowledge sources in this stage are the IT-system, documentation of the IT-system and the criteria list.

#### **Evaluate the IT-system in use according to the chosen criteria (evaluation type 6)**

This stage means that the evaluator performs a criteria-based evaluation of the IT-system in use. The difference between this stage and the stage above is that the evaluation also includes the users understanding if the IT-system meets the criteria or not. This means that the evaluation result also includes observations and interviews of users. During the observation the evaluator could use the technique called “think aloud” (Monk et al, 1993). We recommend that the users should be observed when they are working with real tasks in a natural setting. Available knowledge sources in this stage are the IT-system, documentation of the IT-system, the criteria list and users.

#### **Evaluate effects that can arise (evaluation type 1, 3, 5)**

This stage is performed when evaluating IT-system as such is chosen. Since there are no users available, and therefore no opportunity to observe real interaction, the evaluator identifies possible effects that *can arise*. This means that the evaluator identifies the possible effects. There are two types of effects that can arise. The first type is called interaction effects and the other type is called business effects. Interaction effects are effects that arise from interaction with the IT-system. Interaction effects arise for example when the evaluator:

- doesn’t understand how to interact with system function
- doesn’t understand a message from the IT-system
- can’t find a system function

Interaction effects can cause business effects. A simple example, if a user can’t find the function for order registration and with that the registration is not completed. A business effect in this case is that the order stock is not updated. We recommend that the evaluator try to identify possible business effects caused by interaction effects.

If a goal-free evaluation is chosen the possible effects are related to the earlier identified strengths and problems. If a goal-based evaluation is chosen the possible effect is related to a specific goal. If a criteria based evaluation is chosen the possible effects are related to specific criteria.

#### **Evaluate effects that arose (evaluation type 2, 4, 6)**

This stage is performed for when the IT-system in use is evaluated. The difference between this stage and the stage described above is that there are users available. This means that the evaluator identifies effects that *arose* from observing users interacting with an IT-system.

The two types of effect are described above. If a goal-free evaluation is chosen the effects are related to the earlier identified strengths and problems. If a goal-based evaluation is chosen the effects is related to a specific goal. If a criteria based evaluation is chosen the effects are related to specific criteria.

### **Draw conclusions (evaluation type 1-6)**

Each evaluation type should end with conclusions. The conclusions should consist of key problems and key strengths. Key problems are identified through a problem analysis. This means that identified problems are related to each other in terms of causes and effects. Problem and problem relations can be documented in problem diagrams (see Goldkuhl & Röstlinger, 2003). Key strengths are identified in a similar way where different strengths are related to each other. Strengths on a higher level are related to strengths on a lower level (see Goldkuhl & Röstlinger, 2003). A problem analysis and a strength analysis are performed for all evaluation types. If a goal-based evaluation is chosen the IT-functions that support the goals are viewed as strengths and IT-functions that are missing, counteract the goals or are not related to a specific goal are viewed as problems. If a criteria-based evaluation is chosen, IT-functions that meet the criteria are viewed as strengths and IT-functions that do not meet the criteria are viewed as problems.

There identified key problems should be prioritised. If possible, the conclusions should end with formative statements. This means that recommendations for change requests should be formulated and if possible, proposals for a redesign.

## **6. Conclusions**

The empirical data gathered from the case study have generated six different processes that correspond to the six evaluation types. According to Hammer & Champy (1994) and Porter (1985) a process consists of several related activities. For each process several activities have been identified, labelled and related to each other (see section 5). This means that the initial idea of how to act in each evaluation type and the findings in the case study have contributed to a more articulated description of the evaluation processes.




The reconstruction of the case study has also make it clearer of how the evaluation types relate to each other. As showed, several of the processes overlap each other (see Figure 4). This means some activities are performed in more than one process. An interesting question to ask when deciding upon an evaluation strategy is how the different evaluation types can be combined in a meaningful way. Theoretically, there are a lot of possible combinations.

An analysis of the available knowledge sources in each evaluation type results in that the knowledge sources available in evaluation of “IT-system as such” is a subset of knowledge sources available in evaluation of “IT-system in use”. This means if any of the types of “IT-system in use” is chosen the corresponding type of “IT-system as such” is also chosen. Choosing to evaluate “IT-system in use” means to evaluate IT-systems more carefully than what would be done in “IT-system in use”. This counts for goal-free, goal-based as well as criteria-based evaluation.

A possible combination is therefore to start with any type of “IT-system as such” and then carry on with the corresponding type of “IT-system in use” (see Table 3). Another possible combination is to start with goal-free evaluation and then continue the evaluation with either a goal-based or a criteria-based evaluation. The reason for starting with a goal-free evaluation is that it is harder to discover something new if you ground the findings on predefined categories such as goals or criteria. The simplest type is goal-free evaluation as such. A recommendation is that this type always should be combined with another type. Continuing with an

evaluation type of “IT-system in use” means that strengths and problems that have been identified in goal-free evaluation as such can be verified through using other evaluation types.

Table 3. Possible directions when combining different evaluation types

	IT-systems as such	IT-systems in use
Goal-free evaluation		
Goal-based evaluation		
Criteria-based evaluation		

As future research we suggest a further elaboration of combinations of different types. There is also a need of proposing notational forms. Performing evaluation means asking a lot of questions. Questions generate answers that mean that there also is a need for proposing ways of documenting evaluation results.

## References

Cronholm S & Goldkuhl G (2002). Actable Information Systems - Quality Ideals Put Into Practice. In proceedings of *the Eleventh Conference On Information Systems (ISD 2002)*. 12-14 September, Riga, Latvia.

Cronholm S & Goldkuhl G (2003). Six Generic Types of Information Systems Evaluation. In proceedings of *the 10th European Conference on Information Technology Evaluation (ECITE 2003)*, (Berghout E, Remenyi D eds.). 25-26 Sept., Madrid, Spain.

Goldkuhl, G., & Röstlinger, A. (2003). The significance of workpractice diagnosis: Socio-pragmatic ontology and epistemology of change analysis, in proceedings of *the International workshop on Action in Language, Organisations and Information Systems (ALOIS-2003)*. Linköping University

Goldkuhl G & Lyytinen K (1982). A language action view of information systems”, *proceedings of third Intl. Conference on Information systems*, Ginzberg & Ross (eds), Ann Arbor.

Hammer M (1990). *Re-engineering Work: Don't Automate, Obliterate*, in *Harward Business Review*, July-Aug 1990, p 104-112.

Hirschheim R & Smithson S (1988). A Critical Analysis of Information Systems Evaluation”, in *IS Assessment: Issues and Changes* (eds N Bjorn- Andersen & G B Davis), North-Holland, Amsterdam.

Monk A, Wright P, Haber J, Davenport L (1993). *Improving your Human-Computer Interface*, Prentice Hall, New York.

Nielsen J (1993). *Usability Engineering*, San Diego, California, Academic Press.

Nielsen J (1994). *How to Conduct a Heuristic Evaluation*, [online], [http://useit.com/papers/heuristic/heuristic\\_evaluation.html](http://useit.com/papers/heuristic/heuristic_evaluation.html)

Patton M Q (1990). *Qualitative Evaluation and Research Methods*, second edition, Sage Publications, London. Remenyi D & Sherwood-Smith M (1999). "Maximise Information Systems Value by Continuous Participative Evaluation", *Logistics Information Management*, Vol 12 No 1/2 pp 14-31.

Porter M E (1985). *Competitive Advantage – Creating and Sustaining Superior Performance*. McMillan, New York.

Remenyi D & Sherwood-Smith M (1999) "Maximise Information Systems Value by Continuous Participative Evaluation", *Logistics Information Management*, Vol 12 No 1/2 pp 14-31

Seddon P (2001) "IT Evaluation Revisited: Plus ça Change", *proceedings of Eight European Conference on Information Technology (ECITE)*, Oxford, United Kingdom.

Shneiderman B (1998). *Designing the User Interface: Strategies for Effective User Interface Interaction*, 3<sup>rd</sup> ed., Reading, MA, Addison Wesley Longman

Walsham G (1993) *Interpreting Information Systems in Organisations*, Wiley & Sons.

Whiteside J, Wixon D (1987) "The dialectic of usability engineering", in *Human-computer interaction - interact '87*, Bullinger H-J, Shackel B (eds), Amsterdam, North-Holland

Ågerfalk P, Sjöström J, Eliasson E, Cronholm S, Goldkuhl G (2002) "Setting the Scene for Actability Evaluation: Understanding Information Systems in Context", *Proceedings of ECITE 2002*, 15-16 July, Paris, France