Modelling Risks in IS/IT Projects through Causal and Cognitive Mapping

Abdullah J. Al-Shehab, Robert T. Hughes and Graham Winstanley
School of Computing, Mathematical & Information Sciences, University of Brighton, UK
a.a.shehab@brighton.ac.uk
r.t.hughes@bton.ac.uk
g.winstanley@bton.ac.uk

Software systems development and implementation have become more difficult with the rapid introduction of new technology and the increasing complexity of the marketplace. This paper proposes an evaluation framework for identifying the causes of shortfalls in implemented information system projects. This framework has been developed during a longitudinal case study of a problematic project, which is described.

Keywords: causal and cognitive mapping, project evaluation, information systems project risk

1. Introduction

IT projects have been notorious for their proneness to fail for some time (see, for example, Flowers 1996). In the United Kingdom, recently reported problems include delays to an online hospital booking application in September 2004 (Arnott 2004a), to the national firearms database in October 2004 (Nash 2004) and to the implementation of a secure national radio system for ambulance and fire services in November 2004 (Arnott 2004b). A very well known supermarket chain reported a write-off of £260 millions associated with IT and supply chain systems (Knights 2004). The United Kingdom is not alone in suffering from these setbacks. One survey of over 13,000 IT Projects (Standish Group 2003) estimated that US corporations spent more than $255 billion per year on software development projects of which $55 billion was wasted on failed projects. The project success rate was just 34%, while the project failure rate was 15%, with 51% of projects suffering from cost overruns, time overruns, or a reduction in the features and functions delivered.

One reason that has been put forward for the prevalence of these failures is that information systems/information technology (IS/IT) applications are not constrained by physical laws (Brooks 1995). “Software is largely free of constraints and its potential is therefore unlimited”, according to the Royal Academy of Engineering (2004). Thus it is easy to embark on over-ambitious and ill-advised projects. This is perhaps exemplified by the supermarket project mentioned earlier: in that report (Knights 2004) an analyst is quoted as suggesting that the organization “was too ambitious on the business side, trying to over-segment its customer base which meant that it, in turn, created an overly complex IT system that wouldn’t scale.”

This paper proposes a method of presenting in a visual fashion the factors that have a bearing on project failure and their interrelationships. This allows the different stakeholders in a project to use the diagram to collaborate in the creation of risk models that can simulate the propagation and evolution of risks throughout the project life cycle.

2. Risk management

One textbook definition of risk management is “the identification of the hazards and possible problems, the evaluation of their importance and the drawing up of plans to monitor and deal with those problems” (Hughes & Cotterell 2002, p.134). Thus good risk management should be able to assist the reduction of the likelihood of project failures. There are numerous risk management paradigms, for example: Boehm 1991, Charette 1997, Dorofee et al 1996, McManus 2004, Robin et al 2002 and Yardley 2002.

The elements of risk management are summarised by Boehm (1991) as follows:

- Risk Assessment.
  - Risk identification
  - Risk analysis
  - Risk prioritisation

- Risk Control.
  - Risk-management planning
  - Risk resolution
  - Risk monitoring

Risk management paradigms tend to have a similar content, and most place an emphasis
on adding details of identified risks and their resolution to the lessons learned from previously implemented projects. The maintenance of historical records of risks and issues as they are closed allows an organisation to effectively learn from experience (Boehm 1991, Dorofee et al 1996, Charette 1997).

The definition of risks needs to be conducted with care. There are conditions that might cause loss, such as the inexperience of staff. Methods such as Riskit (Kontio et al 1998) identify these possible causes of risk as risk factors. However, a risk factor by itself does not identify a risk. There must, for example, be many projects where some staff are inexperienced in some way but where the project outcomes are successful. It also follows that the risk factor of inexperience could lead to several different unfortunate outcomes. For example, it could be that productivity is reduced in some cases, and in other cases the quality of the end products is reduced. On the other hand, a risk outcome like ‘delayed project completion’ by itself is incomplete because in this case a risk outcome has been identified, but not the risk factors. Thus Riskit also identifies risk events which, when triggered in an environment where certain risk factors are relevant, will cause certain risk outcomes.

3. Learning in the risk management process

It has been suggested above that an effective organizational learning cycle can increase the capability and maturity levels of the team, project and organisation. However, although risk management is focussed on identifying future problems, it is usually difficult for people to foresee future events and problems (Wiegers 1998). The study of past projects, however, can help to ‘sensitise’ project participants to the potential obstacles to a new project’s success. Top management should therefore support and sponsor the evaluation of implemented projects in a ‘no-blame’ culture. After the implementation stage it is time to learn from the collective experience of the project team and to retain that knowledge, (McConnell 1997). Pitagorsky (2000) sums this up neatly: “The most important step to improve the quality of decision making is the Post-Implementation Review.”

The processes by which an organisation can learn from past experiences are core elements of the concept of the ‘learning organization’ (Argyris 1999). As Argyris points out in his survey of the literature on learning organizations, some writers have identified obstacles that prevent organizations using past lessons as a basis for improving future performance. Leavitt and March (1988) for example point out that organizations often adopt strategies that have worked in the past but which do not work in new situations. The lessons may be based on a small number of cases that might not in fact be typical. The links between cause and effect in past projects may not in fact be obvious or can be subject to controversy. What exactly happened in a past project may not be clear, and judgements about the relative success or failure of a project may depend on the viewpoint of individual stakeholders.

These obstacles to effective organizational learning do not in themselves invalidate the argument that organizations need to learn from past experience. In fact it is argued that they underline the need to make the nature of the lessons learnt and the basis upon planning is based more explicit.

4. Project risk evaluation and documentation

Post-Project Reviews, Post-Mortems or Project Post-Evaluations all have one ultimate purpose: to learn from past projects, be they successful, challenged or failed. However, the terms used in such retrospective processes must be chosen with care. For example some people have a particular perception of the term ‘Post-Mortems’ where Spafford (2003) argues that “a failed project is one thing, but to have a post-mortem meeting after a great project just doesn’t sound good”. Even with acceptable shared perceptions that are intended to emphasise the positive aspects of learning, the actual processes involved need to be carefully designed. Dalcher (2003) states that:

"Information about each failure and the circumstances surrounding the failure are difficult to obtain, but there is also a general lack of knowledge about the ways, methods and approaches for doing so" (Dalcher 2003).

This author goes further in calling for new ways of “studying the failure phenomenon” supported by empirical validation (Dalcher 2003). Ikram (2000) observes that risk management itself has not benefited from rigorous research. In fact, the author is quite
critical about the claims made: “The current literature provides useful knowledge and guidelines on Risk Management, but many of the claims made in the literature have no empirical validation. According to the empirical findings, the application of Risk Management to Information Systems Development is not a common practice.” (Ikram 2000).

DeMarco and Lister (2003) suggest that by applying post mortems to half a dozen past projects we could have enough data as input for future risk management process.

5. Longitudinal case study in IS project failure

The proposed framework was developed as the result of a longitudinal study of a problematic system development and implementation project. An advantage of the case study approach was that it added richness to the detailed information collected. It was able to capture causal influences and interaction effects which would not have been detected by a more statistical approach (Garson 2004).

The longitudinal case studies (LCS) method requires that quantitative/qualitative data are collected a number of times from the area of study (Jensen and Rodgers 2001). In this instance, the case study, which was embedded in a government organization in Kuwait, was documented during several field trips/visits. The project started in 1998-1999 and raised many failure issues at the beginning of 2000. The project suffered from various setbacks during the following two years. At one point the project was stopped for a period of time, and many stakeholders thought that the project had failed and been abandoned. The project was reinitiated and went through much revision of the project design and management approaches. Many problems and issues remain with the project up to the current time.

5.1 First phase

The aim of the first phase of the study was to document the a) background for the project and, b) the shortfalls that led to project failure (Al-Shehab et al 2004). Data gathering was done through semi-structured interviewing (Kane & Brun 2001) of the stakeholders using a taxonomy-based questionnaire (Dorofee et al 1996). The analysis of the data collected in this way used a grounded theory approach (Dick 2003). This was used to identify qualitatively the concept variables relating to the relative success or failure of the project mentioned in the interviews. These concept variables were thus candidate risk factors.

The list of concept variables identified by the stakeholders is shown below:
- High level description
- Unclear project scope
- Insufficient budget estimation
- Poor performance contractor
- Unclear evaluation criteria
- Contractor has no previous experience in developing such projects
- Unrealistic schedule estimation
- Delays in project.
- Contract milestones were not clear towards payment
- Poor partnership relation between customer and contractor.
- Undefined user role
- No user involvement in the project
- All-in-one project
- Technical staff not appropriately skilled to tackle technical activities required
- No tech-team member involvement
- Wrong design
- Undefined project objective
- Lack of project management skills
- Lack of project control
- New technology introduced
- Poor database structure design.
- Lack of leadership
- Lack of communication
- Lack of control over contractor
- Poor product outcome
- Unfrozen requirement
- Poor design
- Poor documentation
- Unstructured design
- Project plan was not followed
- Lack of top management support
- Lack of project guide and objective
- Lack of user commitment
- Delays in acceptance testing.

What is noticeable with this list is that some are clearly risk outcomes, that is, the effects or consequences of preceding problems, as in the case of ‘delays in project’. It is also noticeable that the precise meaning of terms and expressions used by various stakeholders may not be obvious, and that problems associated with the lack of a shared and
explicit ontology may be quite important. Other identified variables do not themselves necessarily point to the probability of project failure, but might contribute to failure in conjunction with other factors, for example, ‘new technology introduced’. These would be identified as risk factors. What is required is the identification of cause and effect relationships between the risk factors and risk outcomes. Below is a list of the relationships between risk factors and risk outcomes that were identified by the stakeholders.

- High level design led to unclear project scope
- Insufficient budget estimation led to use of unqualified contractor.
- Unclear evaluation criteria led to not making sure the contractor has previous experience in developing such projects.
- Unrealistic schedule estimation led to delays in project.
- Contract milestones were not clear towards payment, which led to poor partnership relation between customer and contractor.
- Undefined user role led to no involvement in the project.
- An all-in-one application design led to large number of technical tasks needing to be completed compared with a small number of unskilled technical staff.
- No user involvement led to wrong design.
- Undefined project objectives led to wrong design.
- Lack of project management skills led to lack of project control.
- New technology introduced led to poor structure design.
- Lack of leadership led to lack of communication.
- Lack of control over contractor led to poor product outcome.
- Unfrozen requirement led to unstructured design.
- Poor documentation led to unstructured design.
- Project plan not being followed led to lack of control over project.
- Lack of top management support led to unclear project scope.
- Lack of user commitment led to delays in acceptance testing.

In this list it can be seen that one risk outcome could be linked to several risk factors. For example, ‘wrong design’ is caused by a combination of ‘no team member involvement’ and ‘undefined project objectives’. In other cases, the outcome from one risk becomes a factor that contributes to some other. For instance, ‘undefined user role led to no involvement in the project’ which in turn means ‘lack of user commitment led to delays in acceptance testing’. The relationships between the factors that lead to unsatisfactory project outcomes therefore seem to be more complex than what the simple list of cause and effect relationships above implies. In the next section the possibility of using causal maps to create a richer picture of these relations will be explored.

5.2 Second phase
Experience during data collection and subsequent analysis suggested the need for a method that could capture the often complex interactions between concept variables in a project environment. Causal cognitive maps (CCM) suggested themselves as such a method. The second phase of the case study therefore focussed on adopting CCM as a tool for, a) documenting the past experience of the project, and b) forecasting the outcome of the project.

CCMs were drawn individually by project team members in assisted sessions. These were then combined to produce a consolidated map that was presented to and further analysed in a group session. The participants were asked to provide their views on the map and whether they agreed or disagreed with any of the sub/individual maps. A large amount of data was collected, which is currently being analysed as part of a subsequent investigation into the application of quantitative model building. Currently a commercial tool, Decision Explorer by Banxia, is being assessed for the use in this case study.

5.3 Experimentation with and evaluation of relevant techniques
A particular aspiration of this work is to reach agreement on a common ontology of project risk factors and outcomes that could be applied to future projects. The project team members who are involved in the study can be clustered into a management group and a staff group, and the differences between the two need to be assessed. An intriguing research question is whether CCMs produced by the two groups will be different but complementary, that is focussing on different factors but not actually contradicting each other. An alternative is that there is actual contradiction where one group for example perceives a
positive influence between two factors when the other does not.

Elements of this part of the study are:
- Refinement of project maps in terms of the cause and effect chains that lead to undesirable project outcomes.
- Production of a consolidated map by the management group in a collaborative session.
- Production of a consolidated staff map.
- A comparison of the resulting management and staff maps.
- An investigation of limitations of the tool and techniques.

6. Causal and Cognitive Maps (CCM)

A causal map is a network diagram representing causes and effects (Bryson et al 2004). The diagram contains two basic elements: concepts, which are the nodes in the network and causal relationships, represented by the arcs between the nodes. Concepts are considered as the variables of the system and in some notations carry either a positive or negative sign implying the type of the causal relationship and effect (Tsadiras 1997). Cognitive maps use the concept to elicit and represent perceptions.

CCMs have been used frequently in the operations management discipline (Axelord 1976, Brown 1992, Bryson et al 2004, Eden 1988, Scavarda et al 2004, Williams 1995), commonly using them to support empirical research for building and communicating theory. The areas where the causal maps have been used include:
- Risk mitigation: anticipating unintended consequences.
- Diagnosis: identifying the possible causes of a problem.

Huff (1990) suggests that an advantage of causal maps is that they can portray information about a system more succinctly than a corresponding textual description.

Using CCM as the risk identification approach appears to have some advantages:
1. Group discussions guided by CCMs encourage the participation of all relevant stakeholders in the project.
2. CCM-facilitated group discussions tend to enhance communication between the project members (Gotterbarn, 2001).
3. Using CCM provides a clear picture of the project situation by creating a diagrammatic representation.
4. The diagram enables identification of the interrelations between risks.

![Figure 1: A causal map for the case study in Section 5.1](image-url)
7. System Development Life Cycle (SDLC)

In conventional project management, a project has to be broken down into physical activities to which resources can be allocated. It is argued by the authors that every IS/IT project follows some kind of life cycle that places structure on the temporal order of the project's activities. Admittedly in some projects this order is not very evident, but there is a generic development cycle that can be applied to most of IS/IT projects. This cycle typically contains the following phases:

- **Initiation**: system concept, management approval, funding approval.
- **Planning/Procurement**: identifying stakeholder, top-level project plan, feasibility study, high-level view of the intended system and the determination of its goals.
- **Design/Implementation**: requirement definition, analysis, design completion & approval, coding/unit test, integration/test, system/acceptance test.
- **Installation/Deployment**: installation complete, evaluation/acceptance, site preparation, training, business process re-engineering, rollout.
- **Completed**: review & maintenance.

These stages are reflected in the causal map of Figure 1.

The concept variables relating to risk that have been identified in earlier sections are indicators of some general conditions that could affect more than one activity and these activities could be in more than one project phase. A lack of project management skills could, for example, affect many different activities at many different stages of the project.

However, despite the ubiquity of some of the risk variables, others are specific to particular phases. For example, in the case study, the risk 'lack of user commitment led to delays in acceptance testing' clearly relates to one specific stage in a project. A typical approach to risk management is to maintain a risk register during a project. As the project progresses, some risks will cease to have an impact, while the threat of others may grow. It has been found helpful to produce a template separated into the generic stages of the development life cycle and to locate the occurrence of risk factors and their outcomes within these physical stages, as well as linking the risk factors through chains of cause and effect.

8. The way forward

It is widely accepted that efforts to identify and avoid potential risks in IS projects are commonly rewarded. Weigers (1998) states "anything you can do to improve your ability to avoid or minimize previous problems on future projects will improve your company's business success and reduce the chaos and frustration that reduces the quality of worklife in so many software organizations."

This author goes on to make specific suggestions about what needs to be done, and indeed how to go about it. His suggestions are not limited to the documentation of risk assessment, but also the actual perceptions of those involved in the assessment process. Both informal and formal assessments are important, as are the strategies employed in risk mitigation. Apart from the obvious need to evaluate, identify and mitigate, the overall aim is for an organisation to self-improve via the new knowledge gained, i.e. to maintain and update the corporate knowledge base. The Department of Defence (DoD) risk Management Guide for DoD Acquisition (2003) includes a step called "Analogy Comparison/Lessons-Learned Studies". This step uses lessons learned plus historical data to identify similar risks or risk areas in the current project. They argue that any new project is originated or evolved from existing projects or represent a new combination of existing components or subsystems.

The availability of a high quality historical knowledge base is therefore important. Causal maps can offer far more information than purely textual documentation formats and in a compact and highly-visual format (Huff 1990). Repositories of historical project data are common, but large collections of detailed information gathered during a project's lifecycle are difficult and time-consuming to analyse. There are many ways of gathering data relating to past projects. Some employ surveys, interviews, emails or reports to summarise the history of the project (McConnell 1997), (Collier 1996) and (Wiegers 1998). However, the need for a more structured method of documenting past projects in a more clear representation, is vital.
It is important to recognise that ontological issues are important in situations characterised by disparate groups of people, with different roles and with potentially different ways of expressing concepts. As Gruber (1993) points out, "One problem is how to accommodate the stylistic and organizational differences among representations while preserving declarative content." Stylistic differences relate in our case studies to the language structures used. In our opinion, a common ontology for risk is required, and the issues involved should be conveyed to potential users of causal mapping techniques. The measurement of risk factors and their interdependences also offers exciting possibilities for elicitation, modelling and simulation.

9. Further work

9.1 Risk modelling techniques

The use of CCMs in the management of IS/IT project risk is currently not widespread. In this context, the authors see the use of CCMs as a means of making visible the perceptions of project stakeholders with regard to the causes of shortcomings in completed IS/IT development projects. While no map can be justifiably claimed to be completely accurate in its identification of the source of unwanted project outcomes, they can promote debate and further investigation. This in turn should improve managerial decision-making.

Two major challenges are clear. The first is to move from the diagnosis of the sources of past problems to the prediction and forecasting of potential problems in new projects. While stakeholders may be able to identify such factors as 'insufficient budget' or 'inappropriate experience' as contributors to project shortfalls in completed projects, the identification of these risk factors in the proposals for new projects is not necessarily straightforward.

9.2 Measurement of risk’s cause and effect

The risk assessment process, guided by the process of causal mapping, introduces the concept of expressing the degree to which a risk factor exists in a project, and also the impact such a risk factor has (or is likely to have) on related risks. One approach of measuring the impact is by using a coloration approach (Aladwani 2002). In the authors' experience, stakeholders are commonly unable or unwilling to give precise values to these, but are usually more open to expressing them as an (expert) opinion, i.e. as a ‘fuzzy’ value (Negnevitsky 2002) that does not commit the participant to likely inaccuracy. In spite of potential inaccuracies, in the authors’ experience, expert judgment, tacit though it may be, is valuable and quite often close to reality. Fuzzy representation and reasoning approaches (Kosko 1986, Tsadiras and Margaritis 1997) and neuro-fuzzy and system dynamics (Rodrigues 2001) techniques may well hold the key to usefully capturing such risk data, and making the ensuing models functional.

10. Conclusions

The framework described in this paper can be used to document the knowledge that team members gained through experience. This knowledge can be presented in a single map, essentially visualising the ‘big picture’ of the past project as a pedagogical post-evaluation method. The elicitation process involves the identification of risk factors, their likelihood of occurrence, and their likely impact on other risks within the cause and effect relationships at the heart of the model. The use of CCMs throughout this process has been observed to facilitate ‘brainstorming’ and to achieve consensus in a more intuitive and effective way than with more conventional approaches. The availability of an agreed general risk model of a specific project is also useful, not only in evaluating the accuracy of past diagnoses, but also in providing the means to develop and challenge stakeholders’ mental models of perceived future events.

11. Acknowledgment

We would like to thank those of our colleagues who have contributed to this work, especially Tim Brady and Nick Marshal of CENTRIM at University of Brighton.

References

Arnott, S. (2004a) "Delays hit NHS bookings trial" Computing 2 September p1
Arnott, S. (2004b) "Emergency services face network delays" Computing 18 November p1
Huff, A (1990), Mapping Strategic Thought., Wiley and Sons, USA
Kane, E and Brun, O'Reilly-de (2001) Doing Your Own Research, Marion Boyars Publishers, United States and Great Britain.


Empirical Study on Knowledge Based Systems

Gabriela Avram
Centre de Recherche Public Henri Tudor, Luxembourg
gabriela.avram@gmail.com

Knowledge-based systems (KBSs) implement the heuristic human reasoning through specific techniques, procedures and mechanisms, in order to solve problems that do not have a traditional algorithmic solution. Research on this topic is being done in numerous organisations all over the world, from higher education laboratories to research institutes and software development organisations.

A first research project, aimed at gathering information about the State-of-the-Practice in building knowledge-based systems with practical applications, needed a preliminary study to ascertain if KBSs still exist today as a research topic, or the interest in them actually faded. The study was also required for finding organisations currently building KBSs for different domains. The project’s aim was to catalogue the software and/or knowledge engineering methods employed by the listed organisations, in order to draw a comprehensive image (State-of-the-Practice) of the field. The current paper contains the results of this preliminary study only.

A second research project re-used the results of the preliminary study, focusing on the study of KBSs’ successful implementations as a basis for building a method that would allow practitioners to choose the most appropriate KM tools for each organisation’s specific problems and situations. A trigger for this second project was the interest in studying the causes of KBSs rejection by the end-users. An attempt to map the identified applications of KBSs to different phases of knowledge management lifecycle is also presented.

Keywords: knowledge-based systems, taxonomy, success, failure, knowledge management tools

1. Introduction

Knowledge-based systems (KBSs) implement the heuristic human reasoning through specific techniques, procedures and mechanisms, in order to solve problems that do not have a traditional algorithmic solution. Research on this topic is being done in numerous organisations all over the world, from higher education laboratories to research institutes and software development organisations. During the ‘80s and especially in the ‘90s, a huge number of projects were developed and implemented in this field, and there was an important effort to streamline the development of KBSs by creating engineering methods and tools like Common-KADS and Protégé. But in the last ten years, this attention seemed to continually fade and KBSs almost disappeared from the scene, being mentioned less and less often.

A first research project, aimed at gathering information about the State-of-the-Practice in building knowledge-based systems with practical applications, needed a preliminary study to ascertain if KBSs still exist today as a research topic, or the interest in them actually faded. The study was also required for finding organisations currently building KBSs for different domains. The project was to proceed afterwards with an inventory and the classification of software and/or knowledge engineering methods employed by the listed organisations (if any), in order to draw a comprehensive State-of-the-Practice image.

The current paper contains the results of this preliminary study only (section 3), while sections 1 and 2 are intended to familiarise the reader with the domain of KBSs and to review previous research in the field.

Based on the results of the preliminary study, a second research project was developed, focused on the study of KBSs’ successful implementations as a basis for building a method that would allow practitioners to choose the most appropriate KM tools for each organisation’s specific problems and situations. A trigger for this second project was the interest in studying the causes of KBSs rejection by the end-users. It is well known today that even State-of-the-Art knowledge-based systems have failed in the past because of the lack of organisational concern for the adoption of the system by its intended users. Probably both research and developments emphasized too much the capturing, structuring and packaging of knowledge for reuse, neglecting the role of the human resource in the process. The findings added to the preliminary study during this second stage are presented in section 4.

ISSN: 1566-6379
Reference this paper as:
2. Knowledge Based Systems and their role

2.1 KBS definitions

The literature contains various definitions of this type of systems. From a strictly technical perspective, a KBS is:

“a program for extending and/or querying a knowledge base. A knowledge base is a collection of knowledge expressed using some formal knowledge representation language. A knowledge base forms part of a knowledge-based system (KBS)”. (FOLDOC, 2000), or

“A computer system that is programmed to imitate human problem-solving by means of artificial intelligence and reference to a database of knowledge on a particular subject.” (Computer User High-Tech Dictionary, 2004)

A description including both finality and functionality aspects belongs to the Elsevier Knowledge-Based Systems journal (Knowledge-Based Systems, 2004):

“Knowledge-Based Systems (the journal) focuses on systems that use knowledge-based techniques to support human decision-making, learning and action. Such systems are capable of cooperating with human users and so the quality of support given and the manner of its presentation are important issues.”

From the Artificial Intelligence perspective, KBSs are systems based on the methods and techniques of Artificial Intelligence. The knowledge base and the inference components are separated concepts. There are quite a wide range of opinions on what should and what should not be considered as being a knowledge-based system.

While Stelzer considers that expert systems, case based reasoning systems and neural networks are all three particular types of KBSs (Stelzer, 2003), there are other approaches considering that experts systems and neural networks are different and cannot be included in this category. Other authors also consider ontologies as belonging to KBSs. Davenport and Prusak speak of expert systems, case-based reasoning and neural networks when they give examples of Artificial Intelligence technologies used to support knowledge management, and they never mention the concept of knowledge-based systems (Davenport, 1998).

The Artificial Intelligence and the Organisational Learning perspectives on KBSs are quite different. It seems a narrow, more technical meaning coexists with a broader one.

While from the Artificial Intelligence point of view, KBSs are “hardware & software systems which aim at supporting a specific task by using a specific form of knowledge representation (rules, frames, neural networks) where knowledge is usually highly formalized” (excluding groupware and knowledge sharing mechanisms), the Organisational Learning point of view considers them as being organisations – “a basic support for different specific tasks which includes knowledge in different representation forms (such as experiences, software, procedures, databases, process descriptions) and formalization degrees, and including groupware and knowledge sharing mechanisms” (Stelzer, 2003).

The Organisational Learning point of view sees KBSs as a larger concept and obviously includes the hardware and software systems mentioned by the Artificial Intelligence perspective.

Figure 1 presents a classification of the research sub-domains in Artificial Intelligence, based on the topics list provided by the International Journal of Knowledge-Based and Intelligent Engineering Systems (IJ KBIES, 2004).
2.2 The relation of KBS to Knowledge Management

Compared to Knowledge-based systems, Knowledge management (KM) has a much broader scope, KBSs being only an enabler of KM.

We share Davenport and Prusak’s point of view considering:

‘Knowledge management … concerned with the exploitation and development of the knowledge assets of an organisation with a view to furthering the organisation’s objectives. The knowledge to be managed includes explicit, documented knowledge and tacit, subjective knowledge. Management of this knowledge entails all the processes associated with the identification, sharing and creation of knowledge. This requires systems for the creation and maintenance of knowledge repositories, and to cultivate and facilitate the sharing of knowledge and organisation learning.’(Davenport and Prusak, 1998)

Knowledge management being studied from several different perspectives, such as: organisational learning, artificial intelligence, business informatics, sociology, psychology, information science, informatics, and so on, there are a variety of approaches to this domain. Three ideas are important in this respect: (1) KM is multi-disciplinary, (2) people and learning issues are central to KM and (3) technology is a useful enabler rather than a central tenet at the heart of KM.

Knowledge management includes:
- Processes: knowledge acquisition, codification, storage, use, transfer and dissemination;
- Technologies: KBS, groupware, intranet;
- Knowledge: tacit and explicit, formalized or not formalized;
- People;
- Organisational culture.

Another concept often employed to label an integrated combination of IT tools used to support and enable KM is that of Knowledge Management System (KMS). According to Ronald Maier (Maier, 2004), a KMS “is an ICT system in the sense of an application system or an ICT platform that combines and integrates functions for the contextualized handling of both explicit and tacit knowledge, throughout the organisation or the part of organisation that is targeted by a KM initiative. A KMS supports networks of knowledge workers in the creation, construction, identification, capturing, acquisition, selection, valuation, organisation, linking, structuring, formalization, visualization, distribution, retention, maintenance, refinement, evolution, accessing, search and last but not least the application of knowledge the aim of which is to support the dynamics of organisational learning and organisational effectiveness”. Besides Artificial Intelligence technologies, KMSs also include intranets, document and content management systems, workflow management systems, business intelligence tools, visualization tools, groupware and e-learning systems.

Obviously, KBSs are just one of the applications of Artificial Intelligence included in the wide range of IT tools called Knowledge Management Systems and meant to support Knowledge Management initiatives.

2.3 The development of KBS

A KBS is a software application with an explicit, declarative description of knowledge for a certain application (Speel et al, 2001). There is no clear separation criterion between a KBS and an information/software system as almost all contain nowadays knowledge elements in them (Schreiber at al, 1999). Conventional software applications perform tasks using conventional decision-making logic -- containing little knowledge other than the basic algorithm for solving that specific problem and the necessary boundary conditions. This program knowledge is often embedded as part of the programming code, so that as the knowledge changes, the program has to be changed and then rebuilt. Knowledge-based systems collect the small fragments of human know-how into a knowledge-base which is used to reason...
through a problem, using the knowledge that is appropriate.

The development process of a KBS is similar to the development of any other software system; phases such as requirements elicitation, system analysis, system design, system development and implementation are common activities. The stages in KBS development are: business modelling, conceptual modelling, knowledge acquisition, knowledge system design and KBS implementation (Speel et al, 2001).

A KBS is nowadays developed using knowledge engineering techniques (Studer et al 1998). These are similar to software engineering techniques, but the emphasis is on knowledge rather than on data or information processing. The central theme in knowledge engineering techniques is the conceptual modelling of the system in the analysis and design stages of the development process. Many of the knowledge engineering methodologies developed emphasise the use of models (Common KADS, MIKE, Protégé). In the early stages, knowledge-based systems were built using the knowledge of one or more experts – essentially, a process of knowledge transfer (Studer et al 1998). Nowadays, a KBS involves "methods and techniques for knowledge acquisition, modelling, representation and use of knowledge" (Schreiber at al, 1999). The shift towards the modelling approach has also enabled knowledge to be re-used in different areas of one domain (Studer et al 1998). Ontologies and Problem-Solving Methods enable the construction of KBSs from components reusable across domains and tasks.

2.4 Utility of KBSs

The domain of application for KBSs is widening persistently, as new research topics emerge.

In the 90’s, their foreseen usage directions were: design (the embedding of design rules within applications), diagnosis, instruction, interpreting observed data, monitoring, prediction (by inferring likely outcomes of given situation), prescription of remedies for malfunctions (Swartout, 1996).

Since 1996, the applications of KBSs extended a lot, today proliferating in speech recognition, computer vision, cognitive systems, and many others.

The original promises of Artificial Intelligence were never fulfilled – robots taking over all physical work and computer systems replacing clerks - but AI is still considered by scientists to be „the next Big Thing in science“, while it is gradually moving more and more into everyday life. Today, KBSs are embedded in search engines that remember previous searches, legal software, social software – networking, automated pilots, medical diagnose, call centres, CAD applications, debugging tools.

Figure 2 contains a classification of Artificial Intelligence applications including KBSs.

3. Research on Knowledge Based Systems

The preliminary study demonstrated that during the 90s, there was a special trend of using the designation “knowledge-based systems” for not only computer applications implementing Artificial Intelligence concepts, but also organisational systems that paid a special attention to knowledge acquisition, storage and retrieval.

3.1 Previous surveys

In order to get acquainted to previous research done in this direction, we examined several studies dedicated to the topic or to broader topics including references to KBSs. The conclusion was that plenty of surveys and study cases were performed, and several collections of good practices were built in the field of KM. But just few of these were focused on KBS. The reason for reviewing these studies was the attempt to identify trends and a possible comparison base.
3.1.1 Previous surveys on knowledge management

There are a substantial number of KM-related empirical studies reported in the literature. Most of these also touch the problem of different IT tools used as enablers for KM.

Maier holds the merit of reviewing the most important surveys performed by research institutes, consulting groups and prestigious publications between 1996 and 2001 (Maier, 2004). Some of these studies included a few questions about information and communication technologies. Most of the surveyed organisations seemed to rely on more traditional ICT with no special focus on KM tools. Advanced KM-related technologies, such as AI technologies, were not used frequently (Maier, 2004).

Ronald Maier’s own study on KMS, performed between 1997 and 2001, was targeted at KM strategy, organisation, tools and economics. The results showed that groupware technologies were the most popular, while the rest of tools were not intensively used, mostly because they required substantial organisational changes. (Maier, 2004)

3.1.2 Previous surveys on KBSs

In 1992, Germond and Niebur performed a survey on the development and experiences in using knowledge-based systems (Germond and Niebur, 1992). The study is mainly focused on the role of computers in power systems in Europe, but it also discusses the characteristics of main application areas and forecasted developments in the field. The authors used the KBSs designation to speak mainly about expert systems.

A study performed by Swartout (Swartout, 1996) and dedicated to future directions in knowledge-based systems, identified several problems such as: insufficient understanding of the structure of knowledge-based systems, expensive knowledge acquisition, focus on complete (but narrow) solutions. Swartout’s study also contains notes on the solutions already in place at the time: separation of the different kinds of knowledge entered in the system, deployment of a knowledge engineering methodology for building KBSs, re-use of problem solving methods by using specific libraries, use of ontologies for supporting the building of knowledge bases, knowledge-acquisition tools meant for users (Protégé, Expert).

The 5-th Biannual Conference on Knowledge-Based Systems organised in Würzburg, Germany, published in its Proceedings four surveys dedicated to KBSs (Puppe, 1999). The surveys were focused on Knowledge Engineering and its future directions (Studer, Fensel, Decker and Benjamins), Knowledge-Based Diagnostics (Dressler and Puppe), Knowledge-Based Configuration (Günther and Kühn) and Case-Based Reasoning (Bartsch-Spörl, Lenz and Hübner). None of these studies takes into account the whole picture of KBS, focusing instead on either particular aspects of development and application, or on particular types of KBSs.

3.2 The current survey

The current paper aims at presenting the results of the preliminary study, i.e. presenting what is currently being done in different universities and research institutes under the knowledge-based systems designation, and showing that the interest in Artificial Intelligence in general, and in knowledge-based systems in particular, is still alive.

The trend is very well illustrated by the time span of “knowledge-based system(s)” mentions in research projects funded by the EU in the 1990-2004 period. From the total of 209 projects mentioning the term either in their title or in their description, 83 ended in 1993.

![Figure 3: Projects related to KBSs funded by the EU (Source: the Projects database on cordis.lu)](chart)

Our findings showed that interest in KBSs as research topic has not ceased, but the topic itself shifted toward a more secondary role - KBSs being today generally embedded in other types of systems. They are probably here to stay, but they are not hype anymore.

Our preliminary study intended to map the current situation, by identifying institutions all over the world still using this designation and finding out what are the topics they are working on - a kind of thorough picture of
research done under this name. As mentioned in the introduction, an extensive literature and Internet research was undertaken, attempting to identify organisations still mentioning KBSs in their research agenda.

The study was aimed primarily at organisations involved in designing and implementing knowledge-based systems for practical use, and three categories of such organisations were identified: academia, research institutions and businesses. While academic and research institutions with interests in this field were relatively easy to identify, once the hype passed, there were not many businesses left to label their products with this name. There are still high class scientific publications and prestigious conferences using this name, and all these made us conclude that knowledge-based systems are still a topic of interest for both scholars and practitioners.

At a first stage, we performed an exploratory review of academic centres performing research in this field. We considered the academic environment to be less exposed to commercial trends and more inclined to perform both fundamental and applied research compared to research institutions and businesses, where commercial motivations could be prevalent. But today an important part of the academic research is also commissioned and funded by the industry, so a bias exists there too. We are talking essentially about designations here (names that sell better or do not) and not about the very content of the research.

In the end, it proved to be difficult to separate university research units from research institutes operating under the authority or cooperating with universities, and we did not find it relevant either, so that eventually we decided to consider them together.

Research done in companies under this title was much more difficult to identify; probably due to the influence of market trends on the product names, “knowledge-based systems” was replaced with something trendier in most of the cases. Meanwhile, we are pretty confident that KBS research and development are currently entrenched in a lot of software products on the market, and there are a lot of software companies doing research and development related to this field.

The preliminary study was mainly based on extensive literature and Internet research. We understand the risks involved by such an approach, the World-Wide-Web containing lots of outdated pages of past projects originating in the glory years of KBSs. A lot of research projects on this topic were funded at the time, and almost all of them displayed information on web sites later abandoned. We tried to avoid this by carefully checking the last update date of the web pages taken into account and taking into account only the ones updated in the last two years. We are aware of the limitations of our study and of its lack of completeness. The survey does not pretend to be exhaustive. It was used performing literature review and Internet searches using Google in five languages only, and only the most popular pages where checked. The focus was on building an inventory of organisations involved in the design and implementation of KBSs, including their location, type of organisation (academic, research, business), URL and main research topics.

The original research project, now stalled, planned to continue with a survey including the organisations identified as being involved in KBS research and development, in order to refine and update the collected information and to gather data on the software and/or knowledge engineering methods employed.

The criteria used for defining the boundaries of our preliminary study were:
1. from the wide range of organisations dealing with Artificial Intelligence in general, only research units mentioning “knowledge base(d) systems” either in their name and/or in the topics of interest were picked up;
2. only institutions active in researching, designing and implementing such systems were selected – there are a lot of other universities offering courses related to the topic;
3. only sources updated in the last two years (2002-2004) were retained.

The study identified 47 universities and research institutes performing research in Artificial Intelligence located in six European countries, Australia, Japan, USA and Canada.

From the 47 organisations identified as performing research in the domain of Artificial Intelligence technologies, 16 either present KBS as one of their research topics, or include knowledge-based systems in their unit’s name (e.g. “Knowledge Based Systems Group”, “Knowledge Based Intelligent Engineering Systems Centre”, “Centre for Knowledge-Based Systems”). Even if the others were not included in the focus group at this stage, we
are aware that many of them perform research on topics where KBSs are applied, so it is very possible that KBSs could be embedded in a way or another in their research.

3.3 Taxonomy of the approaches

The interesting part is that using the same name of knowledge-based systems, different research entities focus on very diverse sub-domains and applications. The name of KBS seems to be a sort of general umbrella covering both particular types of KBSs - such as Case-Based Reasoning Systems- and very general KBSs named “Intelligent Systems” and that could, in fact, be based on any other Artificial Intelligence technology. We tried to catalogue the research sub-domains addressed by these organisations in 2 different categories (KBSs and Applications of Artificial Intelligence). The numbers in brackets indicate the frequency of appearance of these sub-domains in the list of research topics of the selected organisations.

As one can see, the applications of Artificial Intelligence included range from the highest degree of generality (“Applied Artificial Intelligence”), to theoretical sub-domains such as “Knowledge Representation and Reasoning” and well known complex application such as “Document Processing”. What we are trying to prove here is there are no precise limits between sub-domains leaving space for a lot of overlapping.

Table 1: Taxonomy of sub-domains mentioned in connection with KBSs

<table>
<thead>
<tr>
<th>A. KBSs</th>
<th>B. Applications of AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>KBS (7)</td>
<td>Knowledge Representation and Reasoning (1)</td>
</tr>
<tr>
<td>Case Based Reasoning (1)</td>
<td>Applied Artificial Intelligence (1)</td>
</tr>
<tr>
<td>Knowledge Based Intelligent Systems (1)</td>
<td>Automatic Programming (1)</td>
</tr>
<tr>
<td>Intelligent systems (1)</td>
<td>Automated Translation (1)</td>
</tr>
<tr>
<td>Fuzzy systems (2)</td>
<td>Cognitive Systems (3)</td>
</tr>
<tr>
<td>Multi-Agent systems (2)</td>
<td>Deduction and Multiagent Systems (1)</td>
</tr>
<tr>
<td>Neural Networks (2)</td>
<td>Document Processing (1)</td>
</tr>
<tr>
<td>Distributed KM (1)</td>
<td>Image processing (1)</td>
</tr>
<tr>
<td>Decision support systems (1)</td>
<td>Intelligent music processing (1)</td>
</tr>
<tr>
<td>Genetic algorithms (1)</td>
<td>Intelligent software agents (1)</td>
</tr>
<tr>
<td>Semantic Web (2)</td>
<td>Knowledge-based computer vision (1)</td>
</tr>
</tbody>
</table>

The fact that the concept of knowledge has different meanings for different specialists – it is application of data and information or information in context for the technologists, while some of the social and organisational scientists claim there can be no knowledge outside human heads – creates a lot of problems with names such as “knowledge base” and “knowledge-based system”. What is stored in knowledge bases is actually knowledge or simply data organised on a higher level of abstraction? Does robotics and computer-vision involve the use of knowledge-based systems, if we consider knowledge as being strictly related to humans?

Another difficulty we encountered was the translation of the title in different languages and the possibly different scope and understanding of these translations. The fact that institutions that perform research and development in the field did not publish information on it in English, French, German, Italian or Spanish made us unable to locate them.

Despite a number of scientific magazines that include KBSs in their title and/or topics, we were not able to locate any joint repository pointing to most of these resources. While research on KBSs is just a small part of the Artificial Intelligence research performed in the world, the fact that different research groups focus on different matters makes it extremely heterogeneous. In order to encourage the building of such a repository in the future, a page dedicated to KBSs was created in Wikipedia (http://en.wikipedia.org/wiki/Knowledge-based_systems) and part of the results were added there.

As of the target readership of the current study, we expect it to be of interest for academics and practitioners involved in both KBS research and in building KBSs.
4. The deployment of KBSs in organisations

This section of the paper refers to another research project aimed at building a method for identifying the most appropriate KM tools in general (and KBSs in particular) for different situations in the real life. From the perspective of the narrow definition of a KBS, that of a program for extending and/or querying a knowledge base collection, we are examining the concrete situation of Knowledge Based Systems’ deployment from the organisational point of view. Some KBSs successful implementations are briefly reviewed - followed by an analysis of the potential causes of KBSs rejection by their users. A first approach on the proposed methodology is presented and a presentation of the forecasted trends for the KBS domain is given in the end.

4.1 Successful implementations of KBSs

A number of application areas seem to profit the most from the deployment of KBSs. This type of systems, stand-alone or embedded in other tools, proved to be very useful in domains such as: natural-resource management, environmental monitoring and cleanup, construction, manufacturing, transportation, aerospace/defence, communications, electric-power generation, wholesale/retail distribution, financial services, logistics, law enforcement, medicine, pharmaceutics.

The experience of successful organisations showed that KBSs are likely to succeed when they focus on well-established, limited subdomains, where knowledge can be properly modelled. In order to ensure their success, Hanley suggests that such systems should only be implemented in places where they can solve an identified problem (Hanley, 2003). The project must be (a) “do-able”, (b) supported by management, and (c) accepted by people in the organisation.

4.2 Implementation failures and their possible causes

Several failures of technically sound knowledge-based systems in the past are today attributed to the lack of appropriate organisational measures to stimulate users in adopting the system. Probably both research and development activities put too much emphasize on knowledge capturing, structuring and packaging, neglecting the role of humans in the process.

One of the most difficult problems is to help users to employ the KBS and understand its advantages. If it doesn’t fill a direct need and if the use of the systems means supplementary work, it is very probable that users will reject it unless they perceive a clear and direct advantage to balance the extra-work.

Knowledge has to be usually captured shortly after the experience occurrence, as close to the source as possible and in a structured way. This operation requires dedicated time and skills, and many users are reluctant to invest in it. Imposing a structure enhances retrieval, but hinders users in contributing experiences, as contributing is perceived as complicated and time-consuming. If the content of a knowledge base isn’t properly filled and updated, there is a high risk of hampering its use after few unsuccessful attempts of getting advantage of it.

Lack of management sponsorship is another factor reported to be frequently leading to KBSs implementation failures. Appropriate training, a system of incentives or enforcement rules, together with the identification of possible champion users and the appointment of facilitators can prevent failure and avoid rejection. Another potential failure factor, in close connection with the previous, is the company culture - that should encourage knowledge sharing and clearly demonstrate its benefits. Feedback for contributing and reusing knowledge should be integrated in the organisational structure. There are numerous companies that while management declares it promotes knowledge sharing, actually encourages knowledge hoarding and do not pay proper attention to possible communication barriers.

Besides these human-related and organisational matters, user requirements being properly taken care of and usability are two other important aspects that if neglected could generate rejection.

4.3 Toward a methodology for selecting the most appropriate KM Tools

The technology has already undergone an adoption-rejection cycle, fed by initially unrealistic expectations and hype. A number of early adopting companies witnessed large-scale KBSs disasters, most of which occurred precisely because of the companies’ overly ambitious faith in the concept of Artificial Intelligence rather than in the reality of KBS technology. However, another class of users—
the companies that implemented the technology on a smaller scale and treated it as just another tool with its own unique assets and limitations—has seen significant benefits. Success stories are still largely in the shadow of early disappointments, but the list of systems with impressive return-on-investment numbers is growing. The point seems to be selecting the right technologies for solving specific problems, paying attention to parameters such as the problem’s scale, the risks involved, the objectivity degree of the involved knowledge.

Attempting to build a methodology for KM tools selection, we picked up the “regions of KM practices” model (Despres and Chauvel, 1999) and tried to map the different types of KBS technologies and applications on it.

![Regions of Practice in KM](source)

**Figure 5:** Regions of Practice in KM Source: Despres, Ch., Chauvel, D. (1999). “Knowledge management(s)”

The original model contained a third dimension, separating tacit and explicit knowledge, but we decided to give it up as being irrelevant for this particular case, because by focusing on KBSs, we implicitly take into account explicit knowledge only.

While attempting to map the different techniques and applications resulted from the preliminary study and connected to KBSs on this model, we realised that it is very difficult to locate the exact phase in the knowledge lifecycle and the level where they would fit. For example, Decision Support Systems (DSS) seem appropriate to be used on the individual level and during the scan/map and capture/create phases, but, depending on the implementation, they could as well support teams for making decisions and could contribute to knowledge transformation by proposing an alternative nobody thought of. Planning and workflow systems (PWS) are useful at all three levels and throughout the capture/create, package/store and share/apply phases. But what can we do about generic titles -such as Expert Systems or Document Management? They are too general for finding their place in that table. Certain techniques are never visible to the users, as they are embedded in search engines remembering our preferences, in automated translation tools we access on the Net or in educational software or computer games.

### 4.4 Current trends

The future appears to be bright for hybrid systems that derive their “expertise” by combining automated extraction of knowledge from data with human experts in specific knowledge domains. These hybrid systems will become increasingly popular as the increasingly digital world gives rise to massive amounts of data that require analysis and as people turn to experts to help them deal with greater complexity and uncertainty (SRI Consulting Business Intelligence, 2003). The signs show that the traditional marketplace for KBSs vanished. Nowadays, they are intrinsically integrated in various Knowledge Management tools, and there is a strong tendency of seeing them as accessories of knowledge workers, rather than a possible substitute for their role.

According to SRI Consulting Business Intelligence, some of the trends of the moment involving KBS deployment are: distributed Artificial Intelligence; real-time KBS; visualization software; standards development; the semantic web; open knowledge bases (SRI Consulting Business Intelligence, 2003).

### 5. Conclusion

As a result of the study, we can conclude that KBSs have not fallen out from the research agenda, but became a basic technique applied in various current research developments, such as ambient intelligence, artificial vision, pattern recognition etc. The study confirmed that the interest in KBSs as research topic has not ceased, but the topic itself shifted toward a more secondary role - KBSs being today generally embedded in other types of systems. They are probably here to stay, abut not holding the main stage anymore - the list of selected research entities and their associated topics attesting it.

Together with the article on KBS created in Wikipedia, we posted there our shortlist of research entities focused on this field, and we pointed at the most important journals dedicated to the topic, as a starting point for a central repository of information on KBSS.

A first step was made in building a method that would allow practitioners to choose the most appropriate KM tools for each organisation’s specific problems and situations. Further, the existing tools will have to be catalogued, the
alternatives for each situation have to be found, and the first guidelines drawn.

References


SRI Consulting Business Intelligence(2003), Knowledge-Based Systems, Technology Map Contents, April 2003 [online], http://www.sric-bi.com/Explorer/KBS.shtml

When Paradigms Shift: IT Evaluation in a Brave New World

Frank Bannister
Trinity College, Dublin
Frank.Bannister@tcd.ie

Over the years, there have been many foci in the search for IT value. However impending developments in information and other technologies may be about to change the nature of the quest entirely. For example, the prospect of technologically enhanced biological function raises new, difficult and disturbing questions about value that need to be explored. Longer term, developments areas such as cyborg technology, artificial intelligence and robotics could have profound, and potentially disruptive, implications for societies and even humanity as a whole. As of now, there is a rapidly diminishing window of opportunity in which to get our values and value systems clear before a combination of technological advance and market forces overwhelms our ability to make important value choices. This paper explores some of the possibilities that may be coming our way and asks some difficult questions about IT value in what may be a brave new world.

Keywords: IT value, emerging technology, artificial intelligence, robotics, cyborgs, nanotechnology, discontinuity.

1. Introduction: Thinking outside the box

1.1 A short parable

Consider for a moment the electric motor. The electric motor was invented in 1834 by Thomas Davenport, an impoverished, but self-taught, Vermont blacksmith (Wicks, 1999). At the time, the new invention was seen as a possible replacement for other forms of rotational power delivery. Davenport himself saw it as a way of powering the machines in his workshop. Later he came to view it as a possible form of engine for locomotives. Unfortunately for Thomas Davenport, despite patenting his brilliant invention, it did not become commercially successful in his lifetime. In fact its first major commercial success came when somebody had the idea of inverting an electric motor and turning it into a generator. Alas, Davenport never saw this possibility and consequently missed out on making a fortune. He died in 1849, still of modest means.

The relevance of this story in this context is that it illustrates the problem of paradigm limited vision, i.e. people’s tendency to overlook the transforming possibilities of new technology. For a long time, the electric motor’s potential was considered only in terms of those current technologies that it might replace, i.e. steam, wind and water. Apart from not seeing the potential for electricity generation, nobody at that time dreamed of refrigeration or vacuum cleaners or DVD players. Yet these technologies utterly depend on Davenport’s invention. The electric motor was going to change the world – but it was a generation or more before just how much it was going to change things became evident.

1.2 Three categories of evaluation

The time horizons within which IT evaluations are generally discussed, whether this be in the most abstract of theoretical expositions or in the most pragmatic of case studies, fall into three distinct categories:

- First there are studies that focus on the long-term historical economic impact of investments in IT. Brynjolfsson (with Hitt and others) have spent many years exploring the so-called productivity paradox and the cumulative effect of investments in IT on organisations. Brynjolfsson (Brynjolfsson and Hitt 2003) is now sufficiently confident of his findings to pronounce the so-called productivity paradox as near dead as matters and assert that it is now beyond dispute that almost all of the increase in US productivity in the past 20 years is due to IT. Other long term thinkers such as Strassmann (1985) have argued for many years that such effects are only really assessable over even longer periods, maybe as long as half a century.

- Secondly, on a less ambitious scale, there are studies of whether specific investments made over shorter periods have yielded value (or of ways of doing this). Usually such research is in the form of case studies and retrospective analyses. These vary from application of innovative methods to measure value realised to use of well established methodologies such as return on investment, comparison of how different metrics report or combinations of measures (such as the balanced scorecard (adapted from Kaplan and
the issue in this paper. What is of interest is Carr's thesis per se is not the issue in this paper. What is of interest is the nature of some of the attacks on it from scholars and professionals. Several of these argue that Carr's vision is hopelessly limited; that he does not understand the nature of technology nor does he take sufficient account of what is to come (see Stewart et al., 2003). An uncomfortable question that might also be asked is can the same criticism be made of IT evaluation? Is it time for IT evaluation researchers to start thinking about what is to come, i.e. outside the box?

It is, of course, well established that technology can have unexpected effects on both individuals and organisations. The way people use a technology may not be that for which it was designed, even where the mechanics of the technology itself are known and well understood. A simple example is the explosion in the use of mobile telephones for text messaging – something not anticipated by the telecoms companies. Another example comes from the early 1980s when there was great excitement about the 'home' PC and all the wonderful things for which people would use them. In practice, those that did not end up gathering dust, ended up being used for playing computer games. The ability of pundits, be they academics or business gurus, to forecast the impact of a technology is far from infallible. If this is true of established technology, how much more true is it likely to be of technology that is not yet with us? It is the conjecture of this paper that there are impending developments in technology for which the weapons currently stockpiled in the arsenal of evaluation techniques are not only inadequate, but are quite inappropriate. This is a bold conjecture and not easy to establish. Nonetheless, this paper will at least try to make the case that there may be a case. In so doing, it will ask some fundamental questions about the nature of evaluation itself in the context of paradigm shifts.

This will be done in two steps. First, although to support the argument it is not necessary to speculate on exactly what the future will be, it is necessary to look at some of the ways that potential developments in information technology could change individuals, organisations and societies over the next few decades. Five probable (or at least possible) developments will be considered. Secondly, it will be argued that contemporary evaluation tools either do not work at all with such developments, or at least do not work very well. Finally there are some reflections on what this might mean in terms of new opportunities for IT evaluation.

---

1 The word 'technologies' is used here and elsewhere to mean information technologies unless explicitly stated to the contrary.

Ironically, the much debated article IT Doesn't Matter (now a book less confrontationally entitled Does IT Matter?) by Carr (2003, 2004) reflects this same mode of thought. Carr argues that, from a business perspective, there is not much exciting left to come from information technology. It may be possible to build more efficient supply chains, make processes even faster, improve customer relationship management or even find out how to manage knowledge better using machines rather than people or organisations, but if everybody is doing these things, where is the strategic advantage? Carr's thesis per se is not.
2. Brave new worlds

2.1 Knee Points

“Prediction”, said Yogi Berra, “is very hard, particularly when it’s about the future”. When surveying the world of IT futures, it is not practical to pursue every possible avenue, so what follows is necessarily selective. There are many technologies which have the ability to radically alter the way people live and work, organisations operate and societies function. Many of these are closely interrelated; indeed all are interrelated at some point and it is frequently at the intersection of technologies that the most challenging issues arise. This section describes some of the more imminent possibilities.

An important point to bear in mind is that some of these technologies are approaching what Kurzweil (1999) calls the ‘knee of the curve’. By this he means that point on an exponential growth curve at which a technology which has been growing slowly for some time suddenly takes off. The concept is illustrated in figure 1.

![Figure 1: Knee point of a curve](image)

One of the characteristics of a knee point is that once a technology passes it, it becomes exceedingly difficult to control it thereafter. The technology takes on a life of its own and subsequent social or even legal constraint may be impossible. Beyond this point the technology tends, as it were, to find its own level. Whether social control is necessary or even desirable is a debate for another day (for those interested, McKibben (2004) discusses this issue at some length). The point to note here is that the IT/IS evaluation community may be faced with tricky questions about several of these technologies sooner rather than later.

---

2.2 Artificial Intelligence

Artificial intelligence is a broad church, encompassing technologies from rule based programming to image recognition. Some of these technologies are now at the knee of the curve. Some are still some distance away and yet others would still appear to be in the realms of science fiction though the latter have an increasingly disconcerting habit of become fact. While the areas in which AI is likely to have a radical impact over the next decade are robotics, control systems and localised intelligence/interfaces, the more remote possibilities of consciousness and self awareness need to be considered. A key step up on the AI tree is the adapting/learning system. At the most basic of levels this is now commonplace. Contemporary developments in robotics are one aspect of this. Anybody who has used a voice recognition system will be familiar with the concept.

For the purposes of this discussion, robots will be defined as differing from (mere) automation in two ways. First, within limits, they are capable of being programmed to do a wide range of non predetermined tasks. Secondly, robot systems have a capability to learn and more advanced systems can adapt to the problems they face and/or the environment that they are in. As a technology, robotics (as opposed to automation) has been slower to take off than many of its more misty eyed advocates anticipated although the market for robotics is forecast to rise to $16 billion by 2007 (Quality 2003). Despite major investments in robotics, today’s factories are still extensively dependent on human labour. Where humans have been displaced in the workplace or in the home it has tended to be by automation rather than by adaptive machines. Nonetheless the prognosis must be that robots will eventually be able to do many of the repetitive manual tasks currently done by humans. This may range from fitting a door on a car to serving a burger and fries.

Localised intelligence/interfaces are a more visible (or audible) development in AI. Such systems are already partially embedded in applications such as directory enquiries systems and help lines. When a customer dials a directory enquiry number today, there is a reasonable probability that she will be answered by a computer. Already experiments have taken place with fully automated telephone booking systems and help desks. Current research in Interactive Voice Recognition (IVR) includes emotion recognition so that machines can respond to users more
appropriately (Yacoub et al 2003). Such systems try to solve the problem themselves and only pass it to a human operator if they cannot resolve it or for the key decision(s). There is no reason why such systems should not be widespread by the end of the decade even at the level of domestic appliances or office equipment.

The ultimate goal of AI is the self aware or conscious machine. Currently this is still science fiction, the world of the film The Terminator or HAL in 2001. Whether a self aware, thinking computer is possible is hotly debated. Popular authors such as Kurzweil (1999) have long forecast this development. The late Nobel laureate, Francis Crick spent much of his later life trying to establish that the brain was nothing more than a biological machine (Crick and Koch 2003) and many of his fellow life-scientists hold similar views. One of the most noted proponents of this theory in the IT world is Marvin Minsky (1988). Philosophers and others have challenged the concept, one of the most forceful being Penrose (1989, 1995). However, if such a development were to be possible, the evaluation of the consequences would raise profound issues including, inter alia, for evaluation.

2.3 Cyborg technology

If AI is about machines behaving like humans, cyborg technology is about machines merging with humans3. This is the world of film The Matrix and the television series The Six Million Dollar Man. One of the most passionate advocates of this type of technology is Warwick (2003) who has gone so far as to have a chip implanted in his arm4. This chip enabled devices in Professor Warwick’s laboratory to detect his presence. Doors would open on his approach and greet him. Chip and bolus technology has been the subject of experiments with animals for tracing purposes (where it causes problems, not least from the tendency of such chips to migrate within the animal). In the Netherlands, customers at the Baja Beech club in Rotterdam have chips embedded in their upper arm in order to get automatic access to the VIP area of the club5. Humans already walk around with artificial limbs and pacemakers; having intelligent devices on board is only the next logical step.

Putting Radio Frequency Identification (RFID) chips into one’s upper arm is one the simplest applications for this type of technology. The ultimate goal is to link the body’s own central processing unit, the brain, to the computer. In theory, there are many ways in which this can be done. A crude method is to provide a link to external machines via electrodes attached to the skull. As fiction, this was the theme of Craig Thomas’ book Firefox (1990) in which the Russians had supposedly developed a fighter aircraft where the pilot’s brain was directly connected (via external electrodes in his helmet) into the aircraft’s control system. To fire a weapon, for example, the pilot merely needed to think of the required action. Enormous research is going into making this particular piece of fiction into fact. With fighter aircraft the race may be between this technology and aircraft that have no pilot at all. Warwick (2004b) describes just such an experiment. A further stage would be to implant such a communications system within the brain so that people could connect to a control system or the Internet via (say) WiFi anytime they desired. Imagine a house that responded automatically to one’s wishes, that would sense one was cold and turn up the central heating or automatically find the TV programme you wanted to watch or music you wanted to hear. In a business context, such a system would enable employees in an organisation to be in constant contact with each other without the hassle of bulky laptops or vulnerable mobile phones. In the longer term there is the possibility of clustering human brains in the way computers are clustered today. It would certainly give a whole new meaning to the concept of teleconferencing.

The ultimate development of cyborg technology would be with the implantation of enhancing technology within the body itself. To take a silly, but not entirely implausible example many people are poor at mental arithmetic. Suppose a company developed an arithmetic processor that could be connected to the brain in such a way that the user could pass any arithmetic problem to it and get an answer6. Such an enhancement could even be biological in nature (see below). Or what about the ability, as in the film The Matrix, to program in any mental skill7? Or how about a memory enhancing device? Many of the science fiction features first seen in the television series, The Six Million Dollar Man in 1975 are today close

---

3 Also a common theme in science fiction (e.g. the Borg in the Star Trek series or the film Demon Seed).
4 www.wordiq.com/definition/Kevin_Warwick
5 The author thanks Professor Egon Berghout for drawing his attention to this interesting example.
6 Readers who think this is far-fetched might wish to look at Warwick (2004a).
7 Programming in physical skills presents more fundamental problems!
to becoming realities. However even the bionic man was confined to increased physical powers; increased mental powers are a different matter entirely.

2.4 Miniaturisation

Another potential source of change, or it may be more accurate to say a catalyst for change, comes from developments in areas such as nanotechnology, biotechnology and quantum computing. At the moment, the first of these is nearer to large scale practical application (for IT) than either of the latter. Nonetheless the possibilities for radical change here are formidable. Biologically based computing, for example, offers the potential to put vast amounts of processing power into the tiniest spaces – a few spoonfuls of water according to one group of Israeli researchers. Quantum computing offers even more staggering possibilities in terms of power per cubic nanometre. As a result, the potential exists to build in, say, a high level of intelligence into any machine or device not to mention into humans. There are also implications in, for example, cryptography where quantum computing could lead to unbreakable codes (Singh 1999).

Miniaturisation is not, in itself, that significant from an evaluation perspective. After all, the world has lived with Moore’s law for three decades now. It is when miniaturisation is combined with developments in AI and cyborg technology that is acquires critical significance. To take but one aspect of this, contemporary surveillance systems still have a limited ability to process the amount of information they gather. The ECHELON system may be able to monitor most of the world e-mail, but it relies on slow and unreliable keyword searching to deduce information (Bamford 2001, Levy 2001). A quantum computing system could read every e-mail, listen to every mobile telephone conversation and digest the contents from every CCTV camera on the planet – a thought that gives a whole new meaning to the concept of a surveillance society. A particular case of miniaturisation is the concept of nanobots. Applications for nanobots are legion from cleaning up waste to keyhole surgery. Most of these applications are benign, but there are risks of undesirable consequences, in particular in control of such technology and in undesirable uses.

2.5 Networking

In many ways, networking is the most prosaic of trends. Unlike some of the technologies discussed in the preceding sections, this is a here and now. The means to connect and collect is already developed if not yet widely deployed. Although, unlike some of the other technologies discussed, the impact of networks is moderated by basic logistical constraints, when combined with artificial intelligence and miniaturisation then the potential effects are profound.

For example, in theory networks could connect everybody in the world or in a state to a central ‘nervous system’ of some sort. A crude version of this already occurs with the telephone system. Mobile phone systems mean that people can be contacted any time and (provided they leave their phone on) their movements traced. Already some nervous wealthy parents are putting RFID chips into their children so that they can trace them if they get lost or are kidnapped. Tagging of criminals is common. It is not difficult to envisage a world where the whereabouts of anybody is knowable at any time, and where everybody who does not make energetic efforts to avoid it is linked into some sort of universal communications matrix. There are considerable attractions in this for some people. The fact that people are prepared to leave their mobile phones on in places as diverse as bathrooms and lecture theatres suggests that many of us are all too willing to be on-call 24 hours a day. From an evaluation perspective, this presents quite complex challenges. What of the risks? What if a virus got into such a network, a possibility explored in fiction by Stephenson (1993). Like the other technologies, networks present opportunities and risks and it is this that causes problems for evaluators.

2.6 Electronic identification

A related issue electronic identification. Reference has already been made to the surveillance society. However there is a broader issue here than a simple fear of big brother looking over everybody’s shoulder. It is will shortly be within the reach of technology to issue a citizen identity card with unique identifying information such as a person’s DNA or retina pattern to all citizens. Identity cards have been a fact of life in many countries democratic as well as undemocratic for decades. Some countries, including the UK and Ireland do not have a citizen identity card

http://search.eurekalert.org/e3/query.html?col=ev3rel&qc=ev3rel&qt=drop+of+water
although the UK government is keen to introduce one (Stephens 2004).

There are obvious benefits from a universal identification system from countering terrorism to preventing fraud. Apart from the implications for privacy and civil liberties, there are also risks from identity theft. These risks differ by an order of magnitude when an identity system becomes electronic. In a society that is ever more dependent on carrying an increasing volume of personal identification, the downside risks to both citizens and society increase. To make matters more complicated, questions of freedom and the nature belonging to a society means are made more complicated. Can somebody remain a citizen without having a card? What is the position of those who refuse to comply? For evaluators, these questions raise difficult issues with soft benefits and disbenefits which have, to date, seen little discussion within the community.

3. Some problems in evaluation

3.1 Conventional methods

Traditional or conventional evaluation of IT assumes that IT has a cost (which may or may not be exclusively financial) and benefit(s). The central question in most of the literature is how to measure these benefits (although as Bannister et al (2003) have argued, measuring the cost is not that straightforward either). Benefits can be individual, organisational, economic or social or any combination of these, but, and it is an important ‘but’, the nature of the change wrought neither threatens the whole basis of society nor our social understanding nor our understanding of ourselves. Each of these is relevant in the context of potential future developments. If the evaluation community has struggled with valuing the impact of current systems and technologies, then the problems in evaluation discontinuous technologies are likely to present an even more formidable challenge.

3.2 Buying some brainpower

To illustrate the problems it is easiest to use examples. Taking cyborg technology as an example, consider the following thought experiment.

A computer company produces an electronic device that can be connected directly to a human brain in order to enhance a person’s mathematical ability. John, a dealer in the financial services industry, is contemplating having such a device installed. It is quite expensive, costing, say, €100,000 including implantation and after sales support. There are no known risks from this technology and running costs are minimal although its estimated working life is 20 years. How might this be evaluated?

In this instance, the financial cost is clear. What about the benefits? A crude approach might to be assess the increase in John’s lifetime earning power or by (say) his ability to make a killing on the futures exchange through a capacity to compute arbitrage rates faster than the market. Adopting this worldview, standard return on investment techniques can be used and value can be quickly assessed. There might also be slightly more subtle benefits from enhancements to John’s lifestyle. He may be able to impress his friends with his mathematical skill or keep track of what he spends in a supermarket as he fills up his trolley. The feeling of well being from this might be worth something although adaptation (Schwartz 2004) makes it likely that this effect will sooner or later wear off. Furthermore, the pleasure might be diluted by the vague feeling that he is cheating, that this ability is not really ‘his’, but that of a device to which he has access.

Another question is who pays the €100,000? If John pays it from his own resources, then there is one kind of evaluation problem. If his employer pays it, then the issues are different.

3.3 On being human

All of the above questions are easily to deal with when compared to the problems raised by the following awkward question: is John the same person that he was before the implant?

It could be argued ‘yes he is’ if one follows the line that John’s improved ability is not really ‘his’ at all. Apart from integration and speed, there is no essential difference between John and somebody with a good laptop and fast fingers. John’s integrity as an individual is not therefore compromised. On the other hand, it can also be argued that there is a difference between such an implant and a laptop. The laptop is out there. Anybody can use it. It can be switched off. The implant is uniquely attached to John, it may even be personalised or tailored to his physiology, and the question ‘does John still exist apart from the implant?’

9 In answer to the obvious question why would an employer pay for such a device, the short answer is that employers currently pay to improve their employees in various ways from putting them on training courses to giving them free medical check-ups.
especially if the implant were biological, is a valid one. Furthermore, would John be able to differentiate between the implant and his ‘normal’ brain. And if one answers no to that question, then some part of John has changed: certainly his ability has been enhanced, but his integrity as an individual has been diminished or at least altered by the fact that he is now partially constructed. The question then is, what is the value of this change/loss?

To make this case more vivid, suppose that instead of a computational implant, John has a memory implant. As the computational implant contains powers John does not have, so the memory implant can contain information that John has not ‘learned’. It might contain a dictionary or an encyclopaedia or two or a language module. And how about some pre-programmed happy memories? It is alleged (it is a much debated topic) that the brain can generate its own false memory (Stanton 1997); imagine what it could do with a little technical help. This raises uncomfortable questions about the meaning of reality, at least from John’s perspective. He may be unable to distinguish his real past from an artificial past created for him. Again this issue has been explored in science fiction. One of the most famous examples is by Dick (2002). It needs to be explored in the evaluation community.

The question that arises in both these scenarios of what it means to be human? Is John more or less human because of these enhancements, or is the nature of his humanity unaffected. He now has increased powers, but less integrity, because he is now, in part, designed by somebody else. However John also has an artificial hip and two dental crowns. Is the brain implant qualitatively different? Furthermore, his perception of reality is in part constructed by outsiders and is not part of his authentic experience. Pushing this to extremes, it is possible to envisage a whole range of IT-powered enhancements that might improve John’s ‘performance’ at the progressive expensive of altering who John is. Is this a cost and if so should it be added to the financial cost of the implant? Even the counter argument that as long as his personality is unchanged, John is unchanged, does not stand up to close scrutiny. How is it possible to disconnect personality or behaviour from memory for example? Finally, how are such benefits and ‘costs’ to be evaluated?

3.4 What is a life?

Another type of problem (if not several such problems) is/are posed by AI. To take an extreme case, suppose that a self-aware, conscious machine is developed. By any definition this would be a new life form. Furthermore, given the dynamics of the situation, unless its design specifically inhibited it, it is likely to be a form of intelligence that will rapidly surpass the mental capacity of its inventors (provided that is they keep supplying it with the extra processing power and memory it needs) 11. The impact of such a development on humanity, not to mention business, is hard to assess. Like cloning and genetic engineering, a conscious machine may subtly (or maybe not so subtly) change the meaning of what it is to be human or at least our understanding of ourselves. This poses major problems for evaluation of such technology. How would one value such a creation? Philosophers, politicians and insurance companies debate the value of a human life. What value would one put on such a machine life (even if it did not have any feelings)? Even avoiding this question, what are the business implications? Such a machine might soon make most managers redundant. This may be cost effective, but the implications for those making the investment could be redundancy and given that many large corporations these days seem to be run mostly for the benefit of managers (Galbraith 2004), who would want to make such a decision?

3.5 Joining the matrix

Another example is the potential of new technologies to disrupt society in too short a time for humanity, be it individuals, organisations or societies to react, is the ability of computers to intrude into people lives. One of the most difficult of evaluations is the trade off between personal security and privacy. With modern technology, it is possible to make people’s lives more secure in a variety of ways. Identity cards with biometric information, closed circuit television, RFID tagging, mobile phone tracking and other technologies can be employed to ensure that citizens are not defrauded, mugged, kidnapped or lost. But this is achieve at a price to privacy and a risk of misuse by private organisation or by the state.

Balancing risks in this situation is not simple and again represents a challenge to the evaluation community. A large part of the

10 This short story was the basis of the film ‘Total Recall’.

11 Dick (1968) also wrote a famous book on this subject. It was made into the film ‘Blade Runner’. In the story, the lifespan of androids is deliberately kept short.
problem in evaluation is the asymmetry in risk perception. People perceive the risks from, say, injury in a terrorist attack to be much higher than it is in practice. In parallel, they fail to perceive potential risks to freedom, privacy and even democracy from technologies ostensibly designed to prevent terrorist attacks. Recently two US commentators, Dash (2004) and Rosen (2004) have looked at this problem from a broad perspective. However it is also an IT evaluation problem and one which, to date, the evaluation community has given little or no attention.

3.6 Problems for evaluators

The problem for the evaluator of the type of technological developments outlined above is which evaluation techniques are appropriate? Financial and economic methods hardly seem appropriate. Consumer satisfaction, organisational improvement, information economics, balanced scorecards; none of these seem adequate to cope with the philosophical problems raised. Questions of risk, judgement, bias, uncertainty, humanity, meaning and society are all impacted by these developments. The English poet John Donne once wrote that ‘no man is an island entire of itselfe’ and this has never been more true than it is today. An investment by any business in these technologies affects far more than the business. Evaluation cannot be ring fenced in these circumstances.

The more transformative an IT induced change, the more difficult is can be to evaluate. There are several reasons for this. First, as was pointed out above, the metrics applicable in the before and after situation may not be the same because, to use a well known phrase, apples are not being compared with apples. This may not be a problem when looking at the investment in retrospect, but it is a definite difficulty in ex ante evaluation. The problem here is that the type of change that might happen with these technologies is more than merely transformative, it is discontinuous. It represents a potential radical departure from the known.

Secondly, the more transformative the change, the more likely it is that evaluation will depend on subjective metrics, i.e. metrics which depend on the views or judgments of actors and in particular ‘experts’ and so-called gurus. It is clear that such views and judgments are not just problematic from a psychological and judgemental perspective, but are time dependent, so the impact may depend on when the measurement is taken (Schwartz 2004, Myers 2002).

Thirdly, the more transformative a technology, the greater the degree of uncertainty about the outcome and the greater the scope for errors of judgment. The law of unintended consequences will inevitably apply. There are many cases of IT investments where the result has not been quite what was expected. In evaluating technologies such as AI and cyborg technology, IT evaluation may therefore have to make much more extensive use than heretofore of general risk evaluation tools such as scenario analysis (see, for example, Wright 2001) and sophisticated risk analyses methods.

In summary, the challenge presented by some of the technologies now evolving is that they are more than radically transformative, they are discontinuous technologies. The evaluation toolset currently available is not capable of providing meaningful assessments of such technologies. It is necessary to reach not just for new methods, but for deeper philosophical tools.

4. Some reflections

In his keynote speech at the 10th European Conference on IT Evaluation (ECITE) in Madrid in 2003, John Ward, reviewing the state of IS evaluation ten years after the first ECITE, suggested that maybe after three decades of academic attention, the topic of IT evaluation was running out of steam. There were many hundreds of papers, books and articles now in print on the topic, but there was still no one agreed approach or agenda. Instead there was a large toolkit of techniques, none of which was entirely satisfactory and which users therefore mixed and matched as the circumstances required.

A more accurate reflection (as a glance of the proceedings of that particular conference shows) would have been that there is no shortage of new ideas, interesting case studies, new applications and new ways of combining techniques around. Research and thinking in the field remains healthy and active; special issues of journals on the topic of IS evaluation are still being published. But it is probably accurate to say that there have been no big conceptual breakthroughs in quite some time (although there are some approaches that may hold promise, for example Halpin and

12 For an amusing, but insightful, collection of examples see O’Boyle (2000) or Harvey-Jones and Tibballs (1999)
Stapleton’s (2003) application of complexity theory to post implementation evaluation). Pluralism is increasingly the name of the game. Approaches to evaluation tend to be holistic and reflective (Cronk’s (1999) concept of ‘holistic construal’ is a good example of this school of thought) of the complex realities that are involved in all but the most simple of situations. A good example of this type of rounded approach can be found in Curley (2004).

This paper proposes a radical new line of enquiry for evaluation of IT, one that looks beyond the known and the short term future and contemplates the impacts of discontinuous technology. In so doing, this paper has raised questions rather than attempt to provide answers. One possible way forward is for the world of IT evaluation to engage much more with current thinking in decision sciences and in risk analysis. Nonetheless there remains a major intellectual challenge in evaluating discontinuous developments. Evaluation here has to move beyond the financial and economic, beyond the conceptual toolkit of the current literature and even beyond questions of risk and uncertainty into fundamental questions about the nature of organisations, humanity, meaning and society. For a subject that is engaged with technology, IT evaluation has tended to be fairly tame in its remit.13 Dealing with the types of challenge and potential for discontinuous change which IT may present over the next few decades will require new tools and engagement with philosophers from Wittgenstein to Midgley who have wrestled with this problem. What may emerge from this may be quite different from the type of discourse that has dominated the field since these issues were first broached in the 1960s. To paraphrase another piece of science fiction, it may be evaluation Jim, but not as we know it.

References


13 There are, of course, some exceptions to this general rule!
Seven Ways to get Your Favoured IT Project Accepted – Politics in IT Evaluation

Egon Berghout¹, Menno Nijland² and Kevin Grant³
¹Centre of IT Economics Research (CITER), University of Groningen, The Netherlands
²M&I/Partners, Amersfoort, The Netherlands
³Glasgow Caledonian University, Caledonian Business School, Scotland

e.w.berghout@eco.rug.nl
Nijland@mxi.nl
kevin.grant@gcal.ac.uk

IS managers are being put under increasing pressure to justify the value of corporate IT/IS expenditure. Their constant quest for the ‘holy grail’ continues, as existing methods and approaches of justifying IT/IS expenditure are still failing to deliver. The decision making process is not as objective and transparent as it is claimed or intended to be. This paper discusses seven commonly used tactics used by business managers to influence IT appraisals. The paper takes a ‘devil’s advocate’ position and adopts some irony when looking at the area of power and politics in IT evaluation. Rather than promoting the use of these techniques, this article aims to raise awareness that IT evaluation is not as rational as most IT evaluation researchers/practitioners would want it to be or indeed claim it to be. It is argued that rationalisation or counter tactics may counteract influence techniques in an attempt to get behind the cloak and dagger side of organisational power and politics, but politics and power in decision-making cannot and should not be filtered out. Due to dissimilarities of objectives, limitations of time and information, influence techniques will always be used. However, rather than being counterproductive, these techniques are essential in the process of decision making of IT projects. They help organisations reach better decisions, which receive more commitment than decisions that were forced to comply with strictly rational approaches. Awareness of the influence and manipulation techniques used in practice will help to deal with power and politics in IT evaluation and thereby come to better IT investment decisions.

Keywords: IT Evaluation, IT Decision Making, IT Assessment, Information Economics, Decision Making, Organisational Power & Politics. Information Management.

1. Introduction

IS managers are being put under increasing pressure to justify the value of corporate IT/IS expenditure. Their constant quest for the ‘holy grail’ continues, as existing methods and approaches of justifying IT/IS expenditure are still failing to deliver what they were intended to deliver and the decision making process is not as objective and transparent as it is claimed or intended to be. This can be attributed to the following reasons; benefits are difficult to assess, measure and manage; costs are high and difficult to predict; large uncertainties and major risk are involved; communication problems and stakeholder politics (Renkema, 2000). It is anticipated that this paper may contribute to managers’ attempts to make sense of power and politics that occurs when evaluating IT investments (Weick, 1995). Sense making is a way to make explicit what is implicit during the IT evaluation process, as sense making seeks to structure the unknown by placing stimuli into some kind of frame of reference, which allows one to “comprehend, understand, explain, attribute, extrapolate and predict” (Starbuck and Milliken, 1988).

The concept of power and politics has been a central theme of study in academia for many years. Within the social sciences, political science and organisational behaviour disciplines it has been widely discussed and reported. Over the years there has been a growing interest and discussion within the IS community to help explain and explore the subjective nature of information systems and how they are integrated and used within organisations.

With the growing recognition of the social and political nature of systems development (Keen, 1981; Markus and Bjorn-Anderson, 1984; Knights and Murray, 1991) and IT evaluation (Hirschheim and Smithson, 1988; 1998), the concept of power is playing a growing and prominent part in IS research. Awareness is growing among researchers that the scientific rational view of evaluation has to be expanded. Perhaps being replaced by a perception of evaluation as a social and political phenomenon, where power is an essential element, as power has a major impact on the decision making process and on the actual decision itself. Many of the existing frameworks developed for IT evaluation identify and discuss the concepts of power and politics, yet they never fully unpack nor assist managers to...
see and respond to the highly complex and highly political dimensions that are embedded within the IT evaluation whole process.

Within the academic literature, there is an abundance of diverse, diffuse and available studies, including ‘n’ step guides on how to do IT Evaluation successfully. While these models and frameworks are often simple, attractive and illustrative they tend to view IT evaluation as rational and objective, with a strong technical focus, and ignore the incremental, serendipitous, and subjective nature of information systems.

Many researchers have tried to understand the nature and essence of what power may be and how it may be defined, however, no widely held consensus of these exist.

Several researchers claim that power is a ‘primitive term’ that needs to be clarified using such terms as ‘influence’, ‘authority’, ‘control’, and so on (Ryan, 1984). Power by many researchers and practitioners is clarified as the resources available to a person to make another person do something that the person would not have done otherwise. Moreover, influence techniques can be viewed as the actual use of power (Raven et al., 1998), whereas power is the potential to influence. Manipulation can be viewed as a form of influence, where the person influenced is unknowingly made to do something, which he or she not had done if he or she knew the perceived results.

Influence tactics and power tactics to influence people have been topics of research by various researchers (e.g. Raven et al., 1998; Yukl and Falbe, 1991; Kipnis, 1990). By drawing upon their work, different bases of power can be identified, on which influence tactics rely on.

Table 1: Bases of power according to French and Raven (1959)

<table>
<thead>
<tr>
<th>Bases of power on which influence tactics rely on</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Coercive power – threat or punishment</td>
</tr>
<tr>
<td>2. Reward power – promise of monetary or non-monetary compensation</td>
</tr>
<tr>
<td>3. Legitimate power – drawing on one’s right to influence</td>
</tr>
<tr>
<td>4. Expert power – relying on one’s superior knowledge</td>
</tr>
<tr>
<td>5. Referent power – based on target’s identification with influencing agent as model</td>
</tr>
<tr>
<td>6. Informational power – convince by (rational) argumentation</td>
</tr>
</tbody>
</table>

Politics is the process by which differing interests reach accommodation. It is the accommodation of these interests that is the business of politics and the accommodations that are generated, modified or dissolved by politics, which ultimately rests on the disposition of power (Checkland, 1990). Knights and Murray argue: power can be seen to infuse all organisational relationships such that rather than being an exception or aberration from the norm, political activity is the focal process through which organisations are sustained, reproduced and transformed (Knight and Murray, 1994). Power can be understood and to some extent made sense of via four dimensions discussed by Horton (1998) when discussing the work of Lukes (1974) and Hardy et al. (1996) and their contributions of looking at power from a sociological perspective. Lukes developed three dimensions of looking and thinking about power. Hardy et al further developed the work of Lukes and added a fourth dimension (Hardy et al., 1996).

1.1 Dimension 1

Actual behaviour in making decisions i.e., concrete instances of power where one agent intentionally influences the behaviour of another one by a certain expressed behaviour (‘A gets B to do something’). Dimension 1 involves a focus on behaviour in the making of decisions on issues over which there is an observable conflict of (subjective) interests (Lukes, 1974).

1.2 Dimension 2

Dimension 1 is restricted to the making of “concrete decisions”, disregarding the so-called “non decision-making” by which potential political issues are kept covert or unvoiced. In non decision-making a decision on such an issue is never considered. Dimension 2 of power allows consideration to be given to the ways in which decisions are prevented from being taken on potential issues over which there is an observable conflict of (subjective) interests, seen as embodied in express policy references and sub political grievances (Lukes, 1974).

1.3 Dimension 3

Moving the concept of power beyond a link with conflict, that is both decision making and non decision making are concerned with issues where there are at least two parties seeking different outcomes. This dimension acknowledges the ways in which issues could be prevented from arising at all, i.e. avoiding the potential conflict altogether.
Luke tries to give some guidance to the conceptualisation of power by distinguishing between instrumental power and symbolic power (Luke, 1974). Instrumental power addresses dimensions one and two, and symbolic power addresses dimension 3, i.e. an unobtrusive use of power occurs in order to secure an outcome by preventing conflict.

However, this indicates that there is a direct belief that power determines choice and changes as if the intention of the powerful were directly coincident and continuous with their effects. Therefore, Hardy adds another dimension to make sense of this issue (Hardy et al., 1996).

1.4 Dimension 4

Entitled the conceptualisation of power that addresses the power of the system, i.e., where power is neither given nor received but exercised and it only exists in action (Foucault, 1980). This is more aligned to the notion of a web like structure of power within an organisation. This view lies in the unconscious acceptance of the culture and subcultures, roles, norms and values of the how the organisation do things with a combination of structural and non-structural mechanisms of the system.

Clegg (1989) builds on this Foucauldian perspective in his framework of circuits of power. The framework considers three circuits: an episodic circuit, social integration circuit and system integration circuit (Clegg, 1989). In the episodic circuit, power is manifested by agents capable of producing their intended outcomes by use of controlled resources and established alliances. The social integration circuits comprise norms, rules and meanings that give identity to particular groups, and system integration circuit consists of technical means and techniques for production (Introna, 1997). The three circuits provide a basis for understanding power as a network of relations, which can be applied to understand situated decisions and outcomes.

Rather than theorising about the concept of power in IS research, this article takes a practical and ‘on the ground’ view of how power is exercised in the process of IT evaluation. It focuses on the manifestation of power in IT evaluation and the description of political actions engaged by managers to try to influence IT decision outcomes. Recognising such actions could aid managers to assess the political issues within the organisation. However, to ensure an enriched and deeper understanding is obtained then more interpretative frameworks are needed, perhaps drawing upon and building upon the existing body of knowledge developed by academics such as Clegg, 1989, Luke, 1974 and Hardy, 1996, who have provided frameworks to help understand why politics is exercised within situated contexts. IT evaluation is one such situated context.

2. IT evaluation

IT evaluation can have many objectives, throughout the life cycle of an IT project (Swinkels, 1997). One important objective of IT evaluation is to facilitate an ex ante appraisal of IT investments. The value of an IT investment proposal is determined to facilitate decision making about a proposal. The discussion on IT evaluation in this article refers to this ex ante appraisal of IT investment proposals.

The traditional formal-rational ideal view on evaluation assumes that it is possible before an IT project commences, that managers and evaluators can determine the outcomes of an IT investment project proposal. Knowing the outcomes, an objective decision about whether to allow a project to go ahead can be reached. Many evaluation methods and decision aiding tools are based on this model. For example, based on a Net Present Value (NPV) technique, one can tell immediately whether to invest in a proposal (e.g., a NPV number that is higher than zero, says the investment is worthwhile). Many criticisms have been raised to this formal-rational view on evaluation and the evaluation aiding methods it employs (Weill, 1993; McKeen and Smith, 1993). Some key criticisms to formal-rational evaluation methods are:

- They neglect the qualitative aspects of investments;
- They favour short-term views on investments and thereby disfavour long-term infrastructure investments;
- They neglect the establishment and discussion of risk factors in investment determination;
- They are susceptible to manipulation, and the inappropriate scientific use and historical ways of working, rather to address and respond to the social view of evaluation (Hirschheim and Smithson, 1999).

On the bases of these criticisms new evaluation perceptions and methods for IT evaluation have been constructed. Methods have been constituted that include the
intangible aspects of the investment (e.g. Parker et al., 1988); that include the notion of an investment lifecycle (Willcocks, 1996, Swinkels, 1997); that assess a portfolio of IT investment proposals (Farbey et al., 1993, Berghout, 1997); that include risk-assessment, and so on. Serafeimidis and Smithson constructed an interpretive methodology that besides the content of the method, also considered context and process of the evaluation (Serafeimidis and Smithson, 1998). However, none of these methods explicitly address the political aspects of IT evaluation.

3. Seven influence tactics

This section details seven influence tactics that are common in the practice of IT decision making. They are presented from a ‘devil’s advocate’ perspective. The findings are based on research relating to the comparison of IT evaluation and decision making within the Benelux and Scottish financial sectors (Nijland and Berghout 2000, Stansfield et al. 2000). These seven influence tactics are by no means exhaustive, however, they show the main strategies commonly used by ‘project champions’ to get their ‘pet’ projects approved.

Table 2: A devil’s advocate perspective on IT evaluation influence tactics

<table>
<thead>
<tr>
<th>Influencing tactics</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Designate a project ‘strategic’;</td>
</tr>
<tr>
<td>b. Designate a project ‘must-do’ or ‘going-concern’;</td>
</tr>
<tr>
<td>c. Slice-up a project (the salami-technique);</td>
</tr>
<tr>
<td>d. Be creative in your cost/benefit/risk analysis;</td>
</tr>
<tr>
<td>e. Find a problem for your solution;</td>
</tr>
<tr>
<td>f. Claim you do not have time to make a proper proposal;</td>
</tr>
<tr>
<td>g. Retry your proposal over and over again.</td>
</tr>
</tbody>
</table>

3.1 Designate a project ‘strategic’

Call your project a strategic initiative, which the organisation must have, so you will not have to commit to verifiable quantitative estimates. With the recognition of recent years that IT investments can have strategic impacts for organisations, many projects have acquired a strategic status. The strategic impact is by definition hard to quantify accurately, as it deals with the future. With the argument that the benefits therefore cannot be calculated, the underpinning of the investment will be at best, a text or narrative discussion, interpreting and describing how the investment will contribute strategically to the success and or the attainment of corporate objectives, e.g. to enhance the existing product portfolio; to develop new products; to give improved and or access to new markets (e.g. by using the Internet), etc. Estimates about what the secondary effects (e.g. increased number of customers, increased number of sales, etc.) of the investments will be absent or a wild guess. By claiming that the project is strategic, it allows the project champion to press ‘hot words or hot buttons’ in senior managers that trigger them to support the project.

3.2 Designate a project ‘must-do’ or ‘going-concern’

Argue that your project is necessary for the organisation to stay operationally efficient and effective, or argue that your project is not an investment, but rather costs for increased maintenance and system management. With investments in the Y2K and in the EURO, together with competitive forces, regulations, and governmental laws, the growth of the Internet and other information communication based technologies, many projects acquire a status of ‘must-do’. Also from a technical perspective, many legacy systems are kept operational, requiring large investments, because the business operations depend on them. One of the more common arguments is that ‘if we do not, our competitors will’, which plays on the mindset that ‘we are (or should we) the best’. Often these projects have highly technical elements, which are unfamiliar to the senior business managers. Under the guise of must-do, a more comprehensive evaluation is said to be redundant. However, even in so-called must-do projects, many alternatives exist. For example, and alternative to investing in legacy systems, a new standard solution package could be acquired to substitute the legacy systems.

3.3 Slice-up a project (the salami-technique)

Divide a bigger project into smaller projects, which stay below the organisational investment threshold for capital projects. Many organisations have a certain financial threshold, before an official IT appraisal is required. As long as the projects stay beneath this threshold, no efforts have to be put into evaluation. Another similar technique entails getting commitment for the first part of the project (salami) and then adds the other parts while the project is on its way asking for additional investments at a later point. Few organisations actually stop a project, which has already started. From this perspective, it is explainable that some projects are started without formal justification, using the argument that waiting for a formal justification would take too long and that the business would loose out if no action were taken quickly.
3.4 Be creative in your cost/benefit/risk analysis

Include only the costs that affect you, stress the benefits of the project to others and be favourable in your assumptions under which conditions your project will be a success. Despite the need and the explicit assumptions within IT evaluation methods on including all lifecycle costs, practice shows that many IT investment proposals only include the initial investment costs, leaving operational costs out. Benefits often are very positive and arise in multiple places in the organisation. Due to an uncertain future, estimates will vary and many assumptions will have to be made which usually are more evangelical than realistic in nature and focus.

3.5 Find a problem for your solution

Look for problems in your organisation that can be resolved by applying your ideas. Many technical investments are justified because they have a champion who has a strong belief in the new technology, not because there is a great organisational need for the investment. The project proposers will look for opportunities in the organisation where the new technology will fit, and build a case for that. It is however by no means certain that the new technology will actually be beneficial to the organisation, or that it is the best choice among the different alternative investments.

3.6 Claim you do not have time to make a proper proposal

In the hectic of a day’s work, people will not mind if you spend your time on keeping the business running, rather than working out investment proposals. Many people in organisations are more inclined to go for short-term goals, rather than long-term goals. If one claims that there was not enough time to make a proper proposal, due to being busy working on keeping things operational, one might get away with a minimum detailed proposal.

3.7 Retry your proposal over and over again

If you do not succeed at first, resubmit your proposal the year after. In the dynamic market of IT, people change positions and jobs frequently. Since IT evaluation can be highly subjective, it is likely that other people will like your proposal if repeated. Many organisations use an unstructured approach to IT evaluation and do not keep records of previously discarded proposals. Reconsider the title your proposal to make the tactic less obvious.

Though we do not encourage the use of these techniques to manipulate outcomes, since they could cloud the objective of IT evaluation to come to an informed IT investment decision, we acknowledge the fact that in real life these influence techniques exist and may even be very useful in terms of persuasion, participation, commitment, bargaining, voting, delegation, generating incentives, etc, on which decision making processes rely heavily on. Indeed it is a widely held believe that it is these influence techniques in fact which make the decision making process (in the case of ex ante IT evaluation) happen (Merkhofer, 1984).

4. Counter tactics

In the political arena of IT investment evaluation, it is equally possible to deploy counter tactics against the seven previously described tactics. Managers typically employ two forms of counter tactics: one of rationalisation and one of political counter tactics. The first is aimed at reducing ambiguity by clarification of objectives, processes and information, working towards some common understanding of what is right and wrong. The second increases political behaviour in the organization, and might even create a battlefield where decision makers have to outwit or overpower their opponents by political counter tactics.

Rationalization as a counter tactic will be described first, followed by more political orientated counter tactics.

4.1 Rationalizing a decision making process

Conflicting objectives are a typical source of political behaviour in organizations. Consequently, political behaviour can be reduced when these underlying conflicts are resolved.

Daft summarizes the differences between political and rational behaviour according to eight characteristics (Daft, 1986). Depending on certainty, uniformity of goals, and tightness and centralization of control, decision-making will be rational, political, or mixed (Daft, 1986). This is illustrated in table 3.
Table 3: Organizational characteristics of rational and political decision making (Daft, 1986)

<table>
<thead>
<tr>
<th>Organisational characteristic</th>
<th>Rational model</th>
<th>Political model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goals, preferences</td>
<td>Consistent across participants</td>
<td>Inconsistent, pluralistic within the organisation</td>
</tr>
<tr>
<td>Power and control</td>
<td>Centralised</td>
<td>Decentralised; shifting coalitions and interest groups</td>
</tr>
<tr>
<td>Decision process</td>
<td>Orderly, logical, rational</td>
<td>Disorderly, characterised by push and pull of interests</td>
</tr>
<tr>
<td>Rules and norms</td>
<td>Norm of optimisation</td>
<td>Free play of market forces; conflict is legitimate and expected.</td>
</tr>
<tr>
<td>Information</td>
<td>Extensive, systematic, accurate</td>
<td>Ambiguous, information used and withheld strategically</td>
</tr>
<tr>
<td>Beliefs about cause-effect relationships</td>
<td>Known, at least to a probability estimate</td>
<td>Disagreement about causes and effects.</td>
</tr>
<tr>
<td>Decisions</td>
<td>Based on outcome maximisation choice</td>
<td>Result of bargaining and interplay among interests</td>
</tr>
<tr>
<td>Ideology</td>
<td>Efficiency and effectiveness</td>
<td>Struggle, conflict, winners and losers</td>
</tr>
</tbody>
</table>

The characteristics described by Daft refer to bringing more clarity and consensus in the organisation and they all require additional time. Consequently, when you are already in favour of a decision or uncertain whether you favour the outcome of the more rational decision making process, you will be less willing to support a more rational decision making process.

Elements that need to be assessed when rationalising a decision making process are:
- Stakeholders: who is responsible for what?
- Decision criteria: which criteria are used and how did each project score on each criteria?
- Project descriptions: which projects are considered, what do they consist of and what should be achieved when?
- Process: how is the entire decision making process organised?

Daft’s simplification of decision-making, though appealing to many managers, may be misleading due to the complexity that surrounds the issues and concepts of power and politics discussed earlier. Daft’s requirements to accurateness of information and consistency of all goals of all participants, it is not logical as to assume there is such a thing as rational decision-making is problematic. However, organisations that employ IT evaluation methods often do this with the aim of rationalising decisions and more ‘objective’.

4.2 Political counter tactics

Given the above problems associated with rational decision-making, managers often apply a political counter tactic to challenge non-favourable proposals. Typical examples are (based on Harrison, 1981):
- Blaming or attacking others. Compare a proposal with similar projects that went wrong. This nicely focuses the discussion on well-known weakness of the other projects.
- Present your own (selective) information. Finding an academic or consultant that supports your arguments should always be possible.
- Use your previously established favourable image (credits). Do not argue, just oppose. Suggesting this should be sufficient to cancel the proposal and move to the next point on the agenda.
- Develop your own base of support. This is preferably done before the actual meeting takes place.
- Associate with other influential persons.
- Create obligations. Agree with the non-favourable proposal, however, do include your own pet project (include success in your established defeat).

Combinations of all counter tactics are being used against all seven influencing tactics.

4.3 Desirability of counter acting the influence tactics

To think that either by rationalising or politically counteracting all possible influence tactics will by definition improve IT investment decision making, is an illusion. To start with the latter, the political counteracting of the influence tactics themselves call for counter-counter tactics to be employed, thereby clouding rather than aiding the IT investment decision process. The former, fully rationalising the decision-process, might seem ideal, however it is also detrimental. Decision-making processes are characterised by political aspects. These aspects are an important reason why such
decision-making processes actually work. A crucial part of decision-making is that decisions are not predetermined, but one has the freedom to reach a decision based on arguments. These arguments do not have to be solely rational; people develop creditable and plausible arguments or frames of reference from their own knowledge and experience, norms and values and their personal worldview. To reduce the freedom in decision making by restrictions reduces the participation of the decision makers who at the end of the day have to execute the decision. Moreover, politics in decision-making helps to solve differences of opinions, and thus potentially help to make compromises. It keeps the organisation stay dynamic and to avoid stasis (Keen, 1981) and is a source of new ideas, which enables radical and innovative changes to happen, i.e., changes that would not fit the format of formal decision making.

An important thing to remember is that decision making most of the time is not about zero-sum situations, but often a solution can be found to create a win-win situation for all parties involved. Such solutions can count on broader commitment, thereby significantly increasing the chance of success.

5. Practical implications for decision makers

Influence tactics and manipulations will always form a substantial part of IT decision making. In the theoretical ideal situation, all necessary information will be available to make an informed decision and be free from personal bias, interpretation and personal gain. All alternative outcomes can be discussed and calculated to determine the most favourable investment opportunity.

However, due to the bounded rationality of decision makers only limited information will be considered during decision making (Simon, 1960). The decision makers are informed (by definition) with partial information, which has been communicated through restricted means of communication (e.g. written or spoken). The outcome of the decision primarily depends on this limited information and the way this is presented. The influence tactics can be seen in this perspective, as a means to either cope with or misuse these limitations.

The way the information is presented to the decision makers can be twofold:
- Influencing: the information is shaped in such a fashion that the final decision matches closest to what the decision makers actually wanted. This means, that although the information was not complete, the decision makers will afterwards be satisfied with the predicted results.
- Manipulation: the information is shaped in such a fashion that the decision makers are tricked into taking a decision, that afterwards turn out to not match their wishes.

In practice this entails discovering the boundary between influencing and manipulation. The decision maker should ask himself if the information on hand gives enough insight and trust in future outcomes to make the decision.

This article aims to contribute to the discussion about managing the politics in IT evaluation by arguing that:
- The decision makers should be aware of the possible manipulation and influence tactics that can occur, starting with the seven described in this article;
- Trying to eliminate politics in IT decision making by rationalisation is undesirable, since decision making is not a strict rational process, but a political process;
- Trying to eliminate possible manipulations beforehand by employing counter tactics is regarded pointless, since counter tactics trigger more manipulations (e.g. counter-counter tactics);
- The decision makers should reserve time to inform themselves sufficiently about the backgrounds of the investment proposal and to further explore unclear and questionable aspects via the persons responsible for the given information.

6. Conclusions

Politics is often referred to in a negative fashion. However, given limitations of time and information and differences of interest, organisational decision-making will always include politics. Mintzberg observes the following benefits of politics (Mintzberg, 1985);
- Politics in organisations may correct deficiencies in other more legitimate systems of influence and provide for certain necessary forms of flexibility not otherwise available. Politics can act in a Darwinian manner to bring the strongest members of the organisation into managerial decision-making roles.
- Politics can often help to promote necessary organisational change blocked by more traditional means.
Politics can often facilitate the decision making process, particularly the effective implementation of choices to serve particular interests.

Politics in decision-making is inevitable. Elimination is both impossible and undesirable. Impossible, because counter tactics themselves are forms of politics. Undesirable, because they are crucial to come to decisions with commitment that keeps the organisation dynamic and flexible to respond to internal and external developments. The bounded rationality of people makes influence tactics and manipulations an integral part of IT decision making processes. Instead of fruitlessly trying to employ counter tactics, it is argued to be aware of the techniques, in which this article has elaborated on, and to inquire further into aspects where manipulation is suspected or where more information is needed to become more confident in the proposal at hand.

By unpacking and making sense of power and politics in a more interpretative manner, then perhaps the knowledge base of IT evaluation can move on from the plethora of post financial methods and techniques of Multicriteria methods; Ratio methods and Portfolio methods (Renkema and Berghout, 1997). The shift towards enriched and interpretative frameworks, which are able to help IT evaluators ‘wrestle’ with the complex social phenomena that exists within the decision making process is needed.

The impact of power and politics in IT evaluation is significant. However, a deeper understanding of the politics of IT evaluation in specific managerial contexts could be reached by making a complete political appraisal of the organisation aided by more interpretive IT evaluation frameworks, by utilisation and developing the work of Clegg as one suggested starting point in the journey for ‘true’ transparent IT evaluation (Clegg, 1989).

References


Hirschheim, R., Smithson, S. (1999). Evaluation of Information Systems. In L. Willococks & S. Lester (Eds.), Beyond the IT productivity paradox (pp. 381-409), West Sussex, UK: John Wiley & Sons Ltd.


McKeen, J.D., Smith, H.A. (1993). The relationship between technology use and organizational performance,
Chapter 20 in: Strategic information management, Banker et al. (eds.), Idea Publishing Group (Harrisburg), pp. 405-444.


Ryan, M. (1984). Theories of power, In A. Kakabadse C. Parker (Eds.), Power, politics and organizations (pp. 21-45), London: John Wiley & Sons Ltd.


Internet Banking in Brazil: Evaluation of Functionality, Reliability and Usability

Eduardo Diniz, Roseli Morena Porto and Tomi Adachi
Escola de Administração de Empresas de São Paulo (FGV-EAESP), Brazil
ediniz@fgvsp.br
Roseli@gvmail.br
tomi@gvmail.br

Evaluating the performance of business Web sites has been a constant concern of researchers in different fields. This article presents an approach that contributes to the development of a methodology to assist researchers, developers and managers to establish criteria to evaluate and build digital business environments. Based on a multiple case study in three large banks in Brazil, this article proposes and tests a model of three dimensions to evaluate virtual business environments from the user's point of view: functionality, evaluates the offered services profile; reliability, investigates the security of a transactional site; and usability evaluates the quality of user interaction with the site.

Keywords: internet banking; banking technology; usability; security; Internet

1. Internet banking in Brazil

Internet banking has been reported as the most important innovation for banking services deployment of the last years. In Brazil, like other countries, Internet banking has grown very fast. Since the early adopters in 1996, banks offering Internet banking services jumped from only 9 banks in 1997 to 75 banks in 2003 (Diniz 2004). These numbers are even more impressive if one considers that these banks represented 83% of the total assets of the whole banking system of the country at that time.

In terms of transactions, Internet banking has kept the highest growth rate among all of the banking channels. Between 1998 and 2002, while the whole banking system had increased 20% a year in the total number of transactions, the ones made through the Internet had a growth rate of 180% a year. The figures (CIAB 2003) also show that Internet banking transactions represented more than 10% of all banking transactions in Brazil, in 2002. In order to make a comparison, ATMs, the most used channel for transactions, held at the time something around 30% of the total number of transactions.

Also the number of Internet banking users in the country has skyrocketed. According with studies carried out by the bankers federation (CIAB 2003), the number of clients banking on the Web almost doubled between 2000 and 2003, from around 8 millions to 15 million, less than 10% of them being from company accounts.

In 2003, most of the banking services offered in regular branches and ATMs were also found on the Internet. As shown at the table 1, this phenomenon of large adoption of Internet banking can be observed by all the main banks operating in Brazil. It is also important to notice that the use of Internet banking use in Brazilian banks reaches higher rates than other countries. While, in Europe, Internet banking use rate is around 18% of the clients, in Brazil, similarly with USA, the same rate is around to 23% (Hessel 2003). However, as one can see in the table, some of the main banks in Brazil have surpassed this percentage.

<table>
<thead>
<tr>
<th>Banks</th>
<th>% clients banking on the Web (2002)*</th>
<th>% clients banking on the Web (2003)**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banco do Brasil</td>
<td>20%</td>
<td>33%</td>
</tr>
<tr>
<td>BankBoston</td>
<td>42%</td>
<td></td>
</tr>
<tr>
<td>Bradesco</td>
<td>33%</td>
<td></td>
</tr>
<tr>
<td>Caixa Econômica Federal</td>
<td>15%</td>
<td></td>
</tr>
<tr>
<td>Citibank</td>
<td>16%</td>
<td></td>
</tr>
<tr>
<td>HSBC</td>
<td>15%</td>
<td></td>
</tr>
<tr>
<td>Itaú</td>
<td>22%</td>
<td></td>
</tr>
<tr>
<td>Real ABN Amro</td>
<td>15%</td>
<td></td>
</tr>
<tr>
<td>Santander</td>
<td>26%</td>
<td></td>
</tr>
</tbody>
</table>

Source: Business Standard*; Gazeta Mercantil**

Although the numbers mentioned above show the growing importance of Internet banking, they do not say anything about how the services are offered. That is exactly the point this article intends to focus on. With many banks offering different on-line services to such a large amount of clients, the question is to develop ways to evaluate the quality of these services. In the next session it will be

Reference this paper as:
presented a discussion on Web site evaluation, aiming the development of a framework for evaluate the quality of Internet banking services.

2. Web site evaluation

Website quality has been evaluated from many different approaches. A number of models derived from the Technology Acceptance Model – TAM (Davis 1989) are used to create a Web evaluation framework (Heidjen 2001; Horton et al. 2001; Lederer and Maupin 2000; Moon and Kim 2001; and others). These models highlight website usability and utility. The WebQual (Barnes and Vidgen 2000; Loiacono et al. 2000) and the e-ServQual (Zeithalm et al. 2002) models add additional analysis dimensions, such as reliability, security and privacy (among others).

In addition to these customer-focused models, the quality of the Web sites can be evaluated from the software development side. Pressman (1997) advances the FURPS – functionality, usability, reliability, performance and supportability – quality factors for software development, and Mendoza et al. (2002) create an ISO/IEC 9126-based portal environments evaluation model.

Certain Brazilian studies on Web evaluation have concentrated on set of on-line services offered (Soares and Hoppen 1998; Diniz 1999; Angulo and Albertin 2000; Nogueira et al. 2000; Steil et al. 2001; Cunha and Reinhard 2001; Mariano et al. 2002). Others studies address browsing- and design-related aspects (Ramos and Costa 1999; Nogueira et al. 2000; Steil et al. 2001; Cunha and Reinhard 2001; Sartori et al. 2002; Oliveira 2002; Ferreira and Leite 2002). A few also investigate transaction security and reliability aspects (Soares and Hoppen 1998; Sartori et al. 2002). A significant share of these works also try to establish criteria and attributes that may be used to evaluate and compare virtual business environments (Cano et Becker 1999; Freitas et al. 2001; as well as others in previous references).

Considering all these models and studies, as well as the authors’ experience with the evaluation of business websites, three concepts appear common:

1. **Functionality**: defines the set of services offered, focusing on business opportunity and strategy;

2. **Reliability**: defines the level of transaction security, focusing on the elements that may contribute to user trust;

3. **Usability**: defines user interaction with the website, focusing on the ease of browsing and performance of tasks leading to completion of the transaction.

The purpose of this article is to propose and test a model based on these three dimensions and capable of making a contribution to the development of methodologies to evaluate and compare virtual business environments from the perspective of website users.

For each one of the above dimensions, this article presents a coherent set of criteria that may be useful to virtual business environment evaluation and design. We propose an analytical framework for business environments in the Web based on: (i) the set of services offered; (ii) investigation of security requirements adopted in business Websites; and (iii) Website usability.

**Figure 1**: Evaluation model of digital business environment

In the proposed model, each of these three dimensions is further subdivided into three other levels (Figure 1), according to the level of difficulty involved in implementing the relevant resources at the corresponding dimensions. The outermost layer contains factors more easily implemented on a Web-based business site, the intermediate layer contains medium-difficulty factors, and the innermost layer contains factors most difficult to implement. Below, we discuss each of these factors in greater depth and consider each dimension's characteristics.
3. The functionality dimension

Functionality, meaning the online services offered, is the first dimension to consider when studying virtual business environments. Prior to anything else, when a company decides to offer its services on the Web, it must consider which services may best result benefits for the company and its customers and partners. This dimension varies according to the industry and company strategy, as well as the profile of the customers for whom a particular set of services is intended.

Diniz et al.(2002) categorize Web services based on direction of the information flow between users and the corporation over the Web. The authors identify certain technical and organizational requirements for on-line services implementation. One major element of this categorization is its focus on the services offered, which divide into the three categories below:

Table 2: Categorization of website services

<table>
<thead>
<tr>
<th>Basic</th>
<th>Intermediate</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissemination</td>
<td>News, Institutional Promotion, Publications</td>
<td>search tool, downloadable documents and forms, links</td>
</tr>
<tr>
<td>Transaction</td>
<td>requests, sign-up</td>
<td>inquiries, payments</td>
</tr>
<tr>
<td>Relationship</td>
<td>e-mail, forms</td>
<td>cookies, calculators</td>
</tr>
</tbody>
</table>

(Source: Diniz, 2002)

4. The reliability dimension

More than a purely technical matter, many experts regard security as a matter of customer perception (Albertin, 1999). As in conventional transaction environments, the digital world too security is never an absolute (Camp, 2000). Transaction errors may occur, be it out of bad faith, user naiveté or lack of expertise, misconduct, fraud, theft, aggression or trespass by third parties. In addition to straightforward human acts, security problems can arise from systems failures or even from the organizational process itself.

Schneier (2000) states that Internet is probably the most complex system ever developed, since it is a public network with millions of computers connected to a highly complex physical network. Each of these interconnected computers has hundreds of software running, and some of these interact with other programs, either in the same computer or in another one connected to the network. As a result of its complexity, Internet security relates with the prevention and detection of, and reaction to, trespass, fraud and loss to prevent financial and moral damages.

The best information security management practices code was developed by BS ISO/IEC 17799 (2000), and is characterized by the preservation of:

- Integrity: data can not be corrupted during handling or transmission.
- Confidentiality: data can not be handled or read by unauthorized people.
- Availability: communication between computers must take place whenever requested.

For the purposes of the Internet, other information security checks have been added (Camp, 2000):  
- Privacy: concerning disclosure of or access to information through electronic means.
- Non-repudiation: ensuring that a completed transaction cannot be denied.
- Authenticity: ensuring that the signatory of a document is really who they purport to be.

Considering the use of computer science itself (Meirelles 1994) and the increasingly disseminated use of technology (Schneier 2000), the security process may also be divided into three layers: physical, logical and human.

The physical layer is characterized by the location of the hardware. It is the physical space where users interacts with the Internet, where the means of communication (cable and waves) are, and where the data is physically stored.

The logical layer is characterized by software-based protection. This layer is defined by encryption and decryption solutions, presence or lack of digital certification of computers and users, application development language, databases, communication middleware with both legacy systems and new ones.

The human layer is characterized by human behavior in the use or maintenance of information systems. According to Schneier (2000), this is the weakest link in the security chain, and chronically responsible for most security system failures. The most important aspect of this layer is people’s perception of risk: how they deal with events that rarely occur; whether they are skilled users or not; the hazard of malicious or unwitting trespassers; and the social engineering by means of which hackers can legally obtain information (Schneier 2000).

Table 3 summarizes the six security items as they manifest in the three layers described earlier. Items that can be identified by simple browsing are highlighted in bold typeface. This table will be the basis to understanding the items that have to be considered for the

<table>
<thead>
<tr>
<th>Table 3: Security Items X Security Layers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Logical Layer</strong></td>
</tr>
</tbody>
</table>
| **Integrity** | - Hash function *(message digest)*  
- system trespass patrolling system | - access passwords  
- secure environment hardware  
- renowned access provider | - proper handling of information  
- prevention of intentional or contingent action of third parties |
| **Confidentiality** | - encryption  
- system trespass patrolling system | - access passwords  
- secure environment hardware  
- renowned connection | - legislation  
- contract  
- enrollment and access procedure |
| **Availability** | - immediate scaling system  
- patrolling and decision system against mass trespass | - redundancy of environment and connection  
- contingency of prompt action  
- balancing of accesses | - training and qualification of the infrastructure team of the company responsible for the Website and of the customer’s web access provider |
| **Privacy** | - encryption  
- corporate and personal firewall | - access passwords  
- secure environment hardware  
- virtual keyboard  
- renowned connection | - legislation  
- contracts  
- privacy policy  
- prevention of intentional or contingent third party action |
| **Non-Repudiation** | - asymmetric encryption  
- digital signature  
- digital certificate | - access passwords  
- secure environment hardware | - contract  
- process of transaction validation |
| **Authenticity** | - digital certificate of participants  
- digital signature | - access passwords  
- secure environment hardware  
- safekeeping of digital certificates | - monitored access to certificates  
- composite/fragmented password  
- documentation |
purposes of evaluating the reliability of a business transaction environment on the Internet.

5. The usability dimension

Since computers are available to a great number of people and support a great variety of applications, information systems projects increasingly require accessibility, usability and user involvement (Laurel and Mountford 1990). The priority for interface designers must be to improve computer usage, conceiving ever more intuitive and user-friendly systems (Rheingold, 1990; Lees 2002).

Usability has become even more important in the Internet age (Nielsen 2000) since it takes place before customers spend any money on potential purchases: on the Web, users experience usability first and pay later. A study of e-commerce usability (Nielsen et al. 2001) found a success rate of only 64%, meaning that more than one-third of all attempted e-commerce transactions fail to reach completion. Evaluation of the causes for this revealed that about two-thirds of the problems in online transactions can be traced back to poor usability projects. Figure 2 shows the more frequent causes of failure found at the sites studied.

![Figure 2: Reasons for transaction failures in digital environments (Source: Nielsen, 2001)](chart.png)

Non-functional usability requirements include interface quality, user-friendliness and human factors often overlooked by software engineers (Ferreira and Leite, 2002). The usability evaluation model used here was developed based on three categories (Pressman, 1997):
- Layout: concerns display of information on the Web page
- Data entry: concerns fields for user information collection
- User on command: concerns how much control users are given over the page

For each of these categories, a group of items was defined to enable ascertaining the Website's usability level. As for 'Layout', three items were selected as being helpful towards Internet banking sites evaluations: consistency, visibility and clarity. Three items were also selected for the 'Data entry' category: forgiveness, feedback and error treatment. And a further three items were selected for the 'User on command' category: search tools and links, constant text size, and site map. The explanation for these items can be found in systems design literature (Microsoft 1991; Apple 1992; Tognazzini 1990; Pressman 1997; Ferreira and Leite 2002; Shneiderman 1998; and others). Table 4 summarizes these three categories and their items.

Table 4: Categories used to analyze the Usability Dimension

<table>
<thead>
<tr>
<th>Usability Dimension</th>
<th>Items to be analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layout</td>
<td>Consistency</td>
</tr>
<tr>
<td></td>
<td>Visibility</td>
</tr>
<tr>
<td></td>
<td>Feedback</td>
</tr>
<tr>
<td>Data Entry</td>
<td>Error prevention</td>
</tr>
<tr>
<td></td>
<td>Forgiveness</td>
</tr>
<tr>
<td></td>
<td>Error treatment</td>
</tr>
<tr>
<td>User on command</td>
<td>Search</td>
</tr>
<tr>
<td></td>
<td>Different ways to do the same action</td>
</tr>
<tr>
<td></td>
<td>Interface</td>
</tr>
<tr>
<td></td>
<td>Customization</td>
</tr>
</tbody>
</table>
6. Evaluation of Internet banking sites

Three banks in Brazil were selected for Internet banking services testing on the model’s three dimensions: Bradesco, BankBoston and Banco do Brasil. The choice is justified by the specific importance each has for the use of Internet banking in Brazil. Bradesco was the first Brazilian Website and the fifth in the world to offer banking services over the Internet (Gates 1999). Banco do Brasil (BB) is the largest bank in the nation and competes with Bradesco for leadership in absolute number of Internet banking users (Business Standard, 2002). BankBoston (BKB), although not as large as the other two, has the lead Internet banking usage ratio, with 42% of its customers making use of the bank’s Web-based services (Hessel 2003). Furthermore, this sample includes one bank from each segment of the industry: one private Brazilian bank (Bradesco), one state-controlled bank (Banco do Brasil) and one foreign bank (BankBoston) with operations in Brazil, therefore encompassing different experiences and Internet banking strategies.

Besides the fact that these are three of the most important institutions in the Brazilian banking sector, this sample is also convenient because since each one of the authors is a customer of at least two of these banks, they have full access to the Internet banking part of the Web sites.

Based on a check list previously prepared with all features described in the three dimensions, each author/customer individually accessed the Internet banking services of two different banks, to ensure a double checking process for all of them. After the data gathering process, a meeting was organized to discuss the individual findings, adjust and refine the data, and deal with any possible doubts relative to application of the criteria. Following are the findings of the evaluation done on the three Internet banking sites.

6.1 Functionality

The functionality dimension is usually identified with the offer of bank services, and the three banks at hand (Bradesco, Banco do Brasil and BankBoston) offer a large roll of services. To consider the functionality dimension for these three banks, we may use the results published by Business Standard Magazine (2002 and 2003), showing a ranking of Internet banking services in Brazil for the past two years, with criteria based on the model illustrated in Table 1. Bradesco, second best in individual services in 2002, maintained its position in 2003. Banco do Brasil, fourth place in individual services in 2002, dropped to fifth in 2003. BankBoston, with the lowest functionality ratio of the three banks considered, ranked 14th in 2002 and climbed to 12th in 2003.

Internet banking functionality is similar at the banks researched: all three post information at the basic and intermediate levels, and customization is already a standard at the advanced level. As for the transaction category, Bradesco offers a greater list of services, followed by Banco do Brasil. On the relationship category, Banco do Brasil does a better job of using the Web to communicate with customers, offering even an on-line consulting channel.

6.2 Reliability

At the logical layer, item ‘integrity’, BankBoston and Banco do Brasil, the certified access option include the hash function, which generates an algorithm that checks for document integrity and enables determining whether or not the document was corrupted during its transit over the network.

Confidentiality and privacy are ensured at the three banks by means of encryption, which varies in degree according to the technology used. The simplest one is Bradesco’s and the most complex is Banco do Brasil’s under the certified access option. As for the non-repudiation and authenticity item, BankBoston and Banco do Brasil, under the certified access option, use digital certification by the customer, and the entire process takes place through asymmetric encryption. Concerning these two later items, the technology used is determined by reading the explanatory texts displayed upon access to the online services and by installing a security component. The difference in use of asymmetric encryption between BankBoston and Banco do Brasil is that BankBoston is the issuer of the certificates, and access to the security component is not controlled, while Banco do Brasil’s certificates are issued by Certisign and download of the security component is restricted to those customers who submitted a signed agreement at the bank branch.

At the physical layer, Bradesco was the first bank in Brazil to use a virtual keyboard to ensure secrecy of the access password, but err by not making use of the virtual keyboard mandatory. Banco do Brasil also uses virtual
keyboards, and only through these can the access password be entered.

At the human layer, confidentiality, privacy and non-repudiation are ensured at all three banks by means of a terms of use agreement or by a letter of agreement under which customers are bound to perform the registration procedures. At Banco do Brasil, issuance of the certificate is only authorized after confirmed signature of the letter of agreement at the bank branch. Banco do Brasil and BankBoston disclose their privacy policies at their Websites. Bradesco’s security policy is limited to the procedures to gain access to the online services, and a session called “security information” that provides information on on-line security and protection. Confidentiality and privacy are also guaranteed by a bill currently under legislative appreciation.

Authenticity at the human layer is ensured by registering access passwords, in addition to the type of data and information requested at the time of registration. Each bank has its own flow and criteria, based on the ATM card passwords and and call center access PIN. The complexity of registration is not necessarily related to the quality of the security. The complexity or simplicity of the process may or may not simplify adhesion to and use of the service, and may be remedied by a clear, easily understood communication at each bank’s Website.

6.3 Usability

As regards layout, at Banco do Brasil the various services are organized hierarchically, and can be found easily by placing cursor upon the tool bar listing the main service groups (also called ‘subsites’): balances; transfers; payments; etc. At Bradesco the second bar (subsites) only shows the services available via Internet Banking and its information is hierarchically structured. The service in use is, however, somewhat difficult to locate, as the subsite title is not highlighted. At BankBoston, services hierarchy is displayed by running the cursor through the links of the ‘Search’ section.

Banco do Brasil’s Internet Banking offers links to BBresponde (help desk) and to FAQ in all pages. BankBoston offers a ‘Menu’ link that drives users on to Internet Banking services. All of Bradesco’s pages offer a site map.

‘Account statements’, the most popular Internet banking service, was chosen to analyze data entry usability at the three Banks. Banco do Brasil asks for the relevant statement month but, in some situations, the current month’s statement is provided in the absence of this information. This is inconsistent with user needs on the first day of any given month. At Bradesco balances can be seen at a specific subsite (‘Account Balances and Statements’) in the upper-left corner which covers checking and saving accounts with the possibility of choosing from five different periods. At BankBoston, customers are only allowed to enter the initial date of the statement, which makes balancing difficult because the covered period is non-customizable. BankBoston Internet Banking uses the metaphor ‘calendar’ for all services that require the filling in the ‘date’ field (payments, transfers, etc.) and follows the consistency principle documented by Tognazzini (1990), according to which mechanisms must be used in the same manner, no matter when and where they occur.

Banco do Brasil’s and Bradesco’s Internet Banking Websites offer no open-ended search, only a menu with a list of predefined options. At the sites of all three banks, when an error occur, a warning appears, in most cases advising how to correct it.

Banco do Brasil offers a customization option at the ‘My page’ link, allowing users to define how they prefer to visualize the more relevant information based on a minimum configuration. The ‘change text size’ option has no effect on the Internet Banking pages. Bradesco Internet Banking text size can be changed by the customer, by means of browser commands. This is an important aspect to meet the visual needs of different customers using a service whose major feature is numbers layout.

7. Conclusion

This article sought to show the use of a model for evaluating a virtual business environment based on three dimensions: Functionality, Reliability and Usability. We believe that these dimensions represent fundamental elements for the conduction of transactions, and that they should also be considered for the evaluation of virtual business environments.

As stated under the ‘Functionality’ dimension, most of the Basic services are available at the sites of Banco do Brasil, Bradesco and BankBoston. However, the adoption of Intermediate level services demands a higher level of user-Website interaction, compelling organizations to regard the digital channel as an environment with specific needs that
requires reorganizing internal procedures. Since this involves resources that demand a specific Web technology skills, a better mastery of the technology and control of its consequences will be required. At an Advanced level, the use of technology will only yield results where perfectly integrated with the organization’s structure. The demand created by the level of interaction also has an impact on other areas, necessitating strategic redefinition and affecting internal organization flows.

Reliability of a service may be incremented by using a larger number of resources to make sure the six security items (integrity, confidentiality, availability, privacy, non-repudiation and authenticity) are present at the three layers (physical, logical and human). There must be a cost/benefit analysis of adopting a larger number of resources versus the complexity of the process and the implementation and maintenance costs of the security architecture. The services analyzed at the three banks boast security architectures that guarantee reliability at a level sufficient to satisfy a large and heterogeneous public. But Internet Banking still has some weaknesses as regards the ‘non-repudiation’ and ‘client authenticity’ items.

Concerning Usability, the ‘Layout’ and ‘Data Entry’ categories meet most of the requisites proposed by those involved. In certain cases, details are overlooked by those responsible for the website, making browsing less intuitive and the interface not as user-friendly as it might be. On the other hand, all three banks should make their Internet Banking services more flexible, so that customers might customize it according to their tastes. Adding user-friendliness, reliability and additional services to the Website is the best way to turn customers who check their balances and account statements into Internet Banking-traders.

Due to the dynamic nature of the Internet business environment, this 3-dimensional evaluation model the corresponding criteria do not purport to be final, and will require constant improvement. In the future we expect to apply this model to other virtual business environments in order to reach a refined model that may effectively serve as an evaluation tool.

References


Edited by Brenda Laurel. Addison-Wesley Publishing Company.


Mariano, Sandra Regina H., Mayer, Verônica F., Ferreira, Patricia T.F. et Gomes, Luiz Flávio A M. 2002 “Aplicação de métodos de apoio à tomada de decisão multiatributo na avaliação de sites de varejo brasileiros” Anais ENANPAD 2002


Exception-Based Approach for Information Systems Evaluation: The Method and its Benefits to Information Systems Management

Heikki Saastamoinen
University of Jyväskylä, Finland
Jyväskylä Polytechnic, School of Information Technology, Finland
heikki.saastamoinen@jypoly.fi

Exceptions are events that cannot be handled by an information system by following normal processing rules. Exceptions arise for two main reasons: flaws in system design and post implementation changes in the system domain. Only few exceptions should arise in an information system serving its user community well. In practice, this is rarely the case and exceptions are sometimes rather common even with routine processes. In this paper, an exception-based approach to evaluate information systems is presented together with practical examples of its use. The benefits of the analysis to information system management are elaborated on.

Keywords: Information Systems Evaluation, Exception Handling, Information Systems Management

1. Introduction

Information systems are implemented with an inherent assumption that they – or the users using them – can handle the associated events the organization faces. This assumption, however, is not completely relevant. The implementation process inherently includes two factors that are hard or impossible to control: system’s design and system’s environment (Berki et al 2004). In a large number of cases, the system covers its domain area only partially (cf. Wand and Weber, 1995), that is, all requirements are not incorporated in the system, the requirements are conflicting, they are misinterpreted, or will become badly implemented. The just finished system or software is not what was expected.

Even in ideal cases where all requirements are met by perfect design, the system starts to outdate from the very day of its completion, due to the changes in its use environment, that is, the domain area the system was designed to cover changes after the system was implemented. Regardless of the technical or organizational environment, the system is supposed to serve, the organization and the world around it are in constant change, eventually causing changes in the technical environment as well.

The extent the above mentioned two factors influence the usability of the systems varies greatly. When an information system is not able to handle all events of its domain, exceptions arise.

Exceptions can formally be defined as cases for the handling of which no applicable rules exist (Auramäki and Leppänen 1989). For most standardized organizational processes, information systems supporting them form the majority of these rules. Thus, exceptions are in every-day life observed as events that cannot be handled by a system. Incoming invoices not matching with the purchase and stock data, and engineering orders with insufficient data are just a couple of examples of such exceptions.

The approach presented in this paper uses the number and kind of exceptions to analyze the operational usability of information systems. It is claimed that a system associated with a high ratio of exceptions versus normally handled events is not serving the organization. Various characteristics of exceptions are briefly discussed to provide means for more thorough analysis of the system and the process in order to find the most crucial points to be improved. Four main benefits of such analysis to information systems management are suggested. Case examples are provided to illustrate how the evaluation can be done and to demonstrate the value of such exception-based analysis.

2. Rules, exceptions and information systems

Before we can formally analyze the concept of exception, the concept of rule must be discussed. The simplest method of coordinating interdependent sub-tasks is to specify their behavior before their execution in the form of rules or programs (March and Simon 1958). Rules can be viewed as instruments of policies aiming to solve problems (Twining and Miers 1976). The primary virtue of rules is that they eliminate the need for further communication among
organizational units (Galbraith 1973). It is sometimes said that the main function of rules is to guide behavior (Twinning and Miers 1976). According to Galbraith (1973), rules thus perform the same functions for organizations that habits perform for individuals - they eliminate the need for treating each situation as new. In addition, rules provide stability to the operations of an organization. When people come and go through an organization, the rules provide a constant for handling routine situations. Thus rules not only transfer past learning, they also control behavior within the organization. These two roles permit the transfer of past learning, and provide a unique solution when a task itself does not provide it (Cyert and March 1963).

We here use the term of "rule" as a generic concept that refers to various types of norms, prescriptions and directives. In a broad sense, a rule can be defined as a general term that includes precepts, regulations, rules of thumb, conventions, principles, guiding standards and even maxims (Twinning and Miers 1976). Good business practice, standard industry practice, and ethical business practice have been seen as rules as well (Cyert and March 1963). In addition, habits and other structures that guide actors' actions are kinds of rules (Williams and Lochofsky 1989). These, however, are usually not precise enough to be used as the basis of event handling or exception handling.

In general language, the term "exception" refers to an abnormal event. That is also the case with information system exceptions. Before the concept of exception or any other concept related to it can be defined, the concept of a "normal event" has to be clarified. A normal event can be defined as an event with the event handling rules necessary for identifying as well as for handling it (cf. Auramäki and Leppänen 1989, Saastamoinen 1993, Saastamoinen and Savolainen 1992). The term "event" here refers to both internal and external events (cf. Wand and Weber 1995). An organization might, for example, have a domain area of order fulfillment for which it has implemented an ERP system including manufacturing, financials, planning and other such modules, to handle events associated with that domain. Examples of events triggering system processing include 'customer orders', 'incoming invoices', and other such external events but also internal events such as 'item stock too low', 'product ready for shipping', etc.

It is these normal events, the information systems are built to process. Kunin (1982) talks about main line as a procedure for the most predictable normal events of a certain type. However, when one is working with information system modeling, one eventually must deal with the problematic details caused by flaws in the main line. These details can become variations and exceptions. Kunin (1982) gives the following definition of the concept of variation: A variation is work that is added to the main line, i.e., a variation is a procedure for less predictable but still known events of a certain type. An exception is an event for the handling of which no applicable rule exists (Saastamoinen, 1995a).

Information systems are formal representations of rules for processing certain events. The systems do not exist alone including only software and hardware and other such "firm" things, but they are always associated with the context they are used in the domain. In any given case, certain processing rules within the domain area are formally implemented in the form of (hardware and) software; the rest is to be handled manually by the users of the system. Thus, from this perspective, an information system is able to process certain events belonging to the categories of main line and variations, as discussed above, but is not able to handle exceptions. The larger the portion of events that can be handled by the system, the better the match between the system and its domain. The higher the number of exceptions to be handled fully or partially manually, the weaker the match is. Thus, the number and kind of exceptions can be used as an approach to evaluate information systems.

### 3. Characteristics of exceptions

The nature of exceptions is negative – even though they are sometimes claimed to have positive impacts similarly (cf. Auramäki and Leppänen, 1989). Most organizations are built to perform in a planned and ordered manner around their core processes and main functions. Exceptions are not a part of those plans and require additional attention and work causing processing delays and additional costs. The costs can be significant even with presumably routine processes (Saastamoinen 1995b). Even though all exceptions share this negative virtue, they are different in many other ways.

Auramäki and Leppänen (1989) discern three elements of exceptionality: acceptability, frequency and degree of difference. Inspired by their initial work, a more comprehensive taxonomy of exceptionality was developed.
The dimensions of the taxonomy are (Saastamoinen 1995a):

- **Exceptionality**: The difference between an exception and a normal event based on rules.
- **Handling delay**: The time between the appearance of an exception and when it can be handled.
- **Amount of work**: The amount of extra work caused by an exception when compared to a normal event.
- **Organizational influence**: The number of people the exception involves.
- **Cause**: The reason for an exception.
- **Rule impact**: The change an exception causes to an organization's rules.

In addition to the above dimensions, **frequency** is a noteworthy characteristic of exceptionality from the viewpoint of system evaluation. Even though almost all exceptions occur infrequently (Saastamoinen et al. 1994), there are some kinds of exceptions that happen more often than others when their frequencies are observed over a period of time.

There are several studies (e.g., Auramäki and Leppänen 1989, Saastamoinen et al. 1994, Saastamoinen 1995a) focusing on classifying exceptions and offering detailed analysis of different dimensions of exceptionality. From the viewpoint of information system evaluation its necessary to master those classifications only in the level of details to understand which kinds of exceptions are more harmful than others and should be addressed with higher priority when the system or process are further developed. The longer the delay, the higher the amount of work, and the more people influenced by the handling, the higher the frequency, the more severe an exception is. For the purposes of information system evaluation, the classification is discussed in more detail in Saastamoinen (2004).

The above discussion regarding the concepts of exception and its characteristics is merely an introduction to the phenomenon. There are few key papers addressing the issue in a general level. In addition to the work already referred to in the above, Suchman (1983), Ellis (1979 and 1983), and Strong and Miller (1989 and 1995) report the first studies concentrating on the real nature of exceptions.

4. Exception handling as an approach for system evaluation

As exceptions are undesired events indicating a mismatch between an information system and its domain, analyzing the number and kind of exceptions provides valuable information not only about exceptional events themselves, but also about the entire system. For example, chances in the number and kind of certain exceptions are a good measure when the value and benefits of a new system in place are evaluated and the new and old systems are compared. Furthermore, the number and kind of exceptions can also be used to evaluate system adaptation, as the number should decrease and severity diminish when the users learn the new process and system. The same analysis can be used even when the successfulness of various rollouts of the same packaged software are compared or evaluated.

4.1 Overview of the approach

Using the number and kind of exceptions for the above mentioned or other purposes is rather an approach than a method. However, practice has proven that certain steps are to be taken for the evaluation to be reliable and to form a solid basis for further development activities. The most important steps to be taken – or issues to be considered – are listed in Table 1:

<table>
<thead>
<tr>
<th>Table 1: The steps of the approach</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main steps</strong></td>
</tr>
<tr>
<td>Preparatory tasks</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Developing the system and organization</td>
</tr>
<tr>
<td>Development actions</td>
</tr>
</tbody>
</table>

The system to be evaluated with this approach needs to be selected with caution. This is, however, in many cases a step that has already implicitly been taken – there is an information system with presumably high number of exceptions. The approach is not inherently limited to certain kinds or types of systems only, however, for optimal results certain precautions have to be made. Quite many information systems are much too complicated, involve huge amounts of users and process so large a number of different kinds of events, that it is not feasible to evaluate the entire system. One can also assume, that it is not the whole system, e.g.,
an entire ERP-package, that would need to be evaluated, but rather a part of it, for example, pay-roll processing, order entry and confirmation, or invoice processing.

Before the evaluation starts, one should carefully study the system's history: are there some factors external to the system itself that are likely to cause the high number of exceptions? For example, a system might have been implemented as a part of the corporate policy even though it was not really meant for this specific line of business, the system is already a couple of decades old, the users did not receive proper training for the system or were not able to participate in its design, recent changes in the organizational structure or processes - just to mention a few of such factors. These issues need to be noted in advance to be able to explain the results.

One should be able to clearly define what are normal cases and how they are processed. Knowing more about exceptions has value of its own, but for the purposes of developing the system, process, or organization, the information is much more usable when it can be reliably compared to respect normal events. Furthermore, if one cannot determine what a normal case is and how it is processed, there is hardly a way to analyze the exceptions. This is not necessarily a problem with an organization's ability to define its processes, but can also be an indication of the system's nature. In the beginning of their famous and widely used textbook "information system management in practice", McNurlin and Sprague (2001) distinguish two main types of information work: procedure based and knowledge based. It is of capital importance to note, that analysis of exceptions is likely to be more beneficial with systems supporting procedure based work. Knowledge based work is often based on ill-structured procedures and its output measures are less defined. With such work one can hardly define a normal event or the task might even have been set to create new ideas, to make decisions, or to create something new.

One of the main benefits of the evaluation is the information of the real causes of the exceptions. Of course, during the course of an analysis one can ask with every exception what caused it. However, in many cases, a majority of the causes are already known, only their real number, frequencies, and relative portions are not.

When a data collection form or a system is designed, it might be useful to utilize the list of known causes as the basis of the form. When certain parts of a system and process are analyzed, the exceptions are likely to be very much alike. For example, for some reason an incoming invoice does not match with the system's data about the corresponding order or goods received. This kind of exceptions - as the cases described at the end of this paper demonstrate - can occur quite frequently. The idea of the evaluation is then not to collect information about exceptions in order to be able to classify them by using taxonomy but rather to gain understanding on why the exceptions take place and how the organization handles them.

When an evaluation period is selected, extraordinary periods of time, such as holiday season, ends of reporting periods, and short peak seasons, should be avoided as the information gathered from such periods is hard to generalize and the results of the study would be too vulnerable for critique. Experience has also shown, that the period should be fairly long to be able to form a picture of how the organization normally handles exceptions. If the evaluation period is short, e.g., one week only, and people know their actions are monitored, they are likely to try to work more efficiently and more precisely to prove their own skills and value. However, a bit longer period, e.g., a few weeks or a month, seems to eliminate that problem by being already too long for the employees to work with other than their normal pace and accuracy. One more factor influencing the period to be selected is the desired sample size, which also speaks for a longer period.

The employees participating in the evaluation have to be well informed. Unless the information system to be evaluated can by itself be used to collect the data, or there is another system associated with it that can be used for the purpose, the collection of the data relies completely on employees. If the evaluation period is of sufficient length and the number of exceptions is high, this can result to a significant amount of additional work. The importance of filling in every form completely and accurately - either on screen or in a sheet of paper - has to be stressed out. One cannot overlook the fact that many exceptions are internally caused by staff inexperience or staff carelessness, and some of those people causing the exceptions initially might be the people also partially handling them on the later stage. With that in mind, it has to be
Heikki Saastamoinen

emphasized that one is evaluating the information system and the process it supports, not the people using it.

As the above is unlikely to be fully absorbed by all participating employees, it is vital for the researcher to be active especially in the beginning of the evaluation period. Should there be forms incompletely filled or with inaccurate data, the problem needs to be addressed immediately to avoid further such problems and to correct the data perceived inaccurate.

After the data is collected, it can be analyzed with normal statistical methods. It still has to be noted, that part of the people participating might have been the people causing the problem initially and that others might be the ones the work of whom is slowed down or complicated by the problems. The first group might have had an objective to diminish the problem while the latter group might want to emphasize it.

4.2  Cautions regarding the use of the approach

There are a few general cautions that have to be made about the use of this approach. First of all, this approach does not provide its applier with a holistic view to the information system evaluated; it only focuses on the part of the system not performing as planned. One should not judge the entire system or the organization based on the results of this analysis. Even though the number of exceptions observed might be relatively high, a conclusion that either the system is wrongly implemented or totally outdated, or that the organization is incapable of using the system, might be premature.

Secondly, this approach does not vary from any other approaches: the researcher using the approach needs to understand well what he or she is doing, and the one implementing the approach should know the organization well enough to be able to interpret the results correctly.

If an overly high number of exceptions are observed, definite conclusions need to be drawn and corrective actions have to be taken, however, a more throughout analysis is just required before them. A large number of exceptions observed can be an evidence of fatal mismatch between the information system, its users, and the organization. It can also be an indication that the process underlying the system and the one assumed by the organization do not match - or it can be due to the fact that the target system of analysis or the period of the analysis were not well selected.

4.3  Using the approach for information systems management

An analysis based on exceptions can be very valuable from the viewpoint of information systems management. The analysis provides an information system management team with the four main benefits:

• Structures prioritization of system development and maintenance activities
• Shifts the focus from technology to processes
• Increases communication with the user community
• Provides feedback about the performance of information system operations

In the following, each of these benefits is briefly discussed with some emphasis on the first one of them.

4.3.1  Prioritization of system development and maintenance activities

Information systems are to serve the organization. In terms of the procedure-based systems (cf. McNurlin and Sprague, 2001), exceptions should not exist in large numbers. If that is the case, an organization has a system in place that fails to fulfill its purpose. The flaws revealed by using the approach should be corrected – on the system or on the organization that is using it – for the organization to perform well.

For fulfilling the ultimate goal of any given organization, information systems are unlikely to be equally important. Some systems have a more crucial role while the others are merely supporting some secondary activities. We can call this relative importance of the systems to their impact on organizations’ performance. The impact depends on a number of issues varying from organization to another: relative amount of revenue flowing through the system, number of users using the system, system’s visibility to external customers, the system being or not being a part of e-commerce solutions or portals – just to name a few.

By combining the two, the number of exceptions observed and the impact of the systems, an organization can outline a graph to pinpoint the systems that most urgently need to be either further developed or replaced.
It needs to be noted here, that this prioritization only covers systems already existing; new systems possibly required by the business are not included.

Figure 1: Prioritization of IS development activities

As the impact of the system is a fairly vague concept, it is strongly recommended here that a formal approach to classify systems by their impact should be used. For example, a split to operational control (including process management and asset management) and to organizational effectiveness (including growth and increase in market share, restructuring of the organization, and restructuring of the industry) presented by Primozic, Primozic, and Leben (1991) can serve as a neutral framework over multiple personal opinions regarding the importance and impact of organization’s information systems.

4.3.2 Focusing on processes

Exceptions arise as a result of a lack of applicable rules needed to handle events. As discussed earlier, those rules can be incorporated into information systems, or they can exist in the minds of the user community. When exceptions are analyzed, it is found out that they are virtually never caused by technical malfunctions (Saastamoinen et al, 1994). On the contrary, the majority of them seem to be caused by people; either by the users of the system or closely connected parties such as suppliers (Saastamoinen, 1995b).

Keeping the above in the mind, the activities taken to decrease the number of exceptions are unlikely to be only software development projects or other such technical undertakings. The focus naturally shifts on the issue how the system is used. The process the system serves and how the process and the system match will be analyzed first, hardware or software implementations are to follow on the later stage.

4.3.3 Increased communication with the user community

As is apparent in the above, exceptions cannot be avoided by the actions of information system professionals only. The issue of exceptions has to be thoroughly addressed by the users and IS professionals together. This has few obvious but remarkable benefits: it helps IS professionals to understands the business, fosters relationships between the IS professionals and the users, and assists in creating a common vision with the users.

Furthermore, as briefly discussed earlier, exceptions have two root causes: flaws in systems design and changes in the systems environment. The first of the two is highly related to the communication. It is claimed here, that the systems are not necessarily overly difficult to design to match the reality more closely. In large number of cases the requirements of the users are simply not expressed well enough and are understood even more poorly resulting to the requirements not being implemented at all. To avoid the same recurring while systems are updated, all communication assisting in real exchange of information and contributing on mutual understanding on the system and its requirements is of essence for the entire organization.

4.3.4 Feedback about the performance of information system operations

Using the approach to evaluate information systems is not only valuable as such, it furthermore can server as a way to evaluate the performance of information system operations. A multitude of models and scoring systems exist to evaluate the performance of information systems and technology. Some of the models are widely used in practice, some can even be found in the literature. What appears to be the most difficult factor when IS/IT operations are to be evaluated objectively, is the performance of the information systems, that is, the match between the systems and the processes it serves.

This evaluation can be greatly facilitated by systematically using the kind of an approach reported in this paper. It is suggested here, that the results of the same analysis that is used to prioritize the development and maintenance activities, as suggested earlier in
this chapter, could be used also to evaluate the performance of the information system operations of the organization. Evaluating IS/IT operations of an organization without proper emphasis on the support the systems give to crucial processes would clearly be less than adequate.

5. Case example: A large engineering shop

This approach has been used extensively in a multitude of cases, however, a set of two case studies performed in a large engineering shop in 1993 (Saastamoinen, 1995b) and again in 2001 are the best examples for the purposes of this paper. Even though the study carried out in 1993 involved also other systems, both of the studies focused on unmatched incoming invoices.

An invoice does not match if the data stored from the invoice does not match the data of the corresponding order or the data of the corresponding deliveries. For example, if the unit price in the order database does not match the unit price in the invoice database, the corresponding invoice is unmatched. Likewise, if an item invoiced cannot be found in the storage database or the number of delivered items does not match the number of invoiced units, the invoice does not match.

These unmatched invoices are exceptions for transaction verifiers. When they verify an invoice, they are supposed to be able to decide whether it is correct or not. If it is unmatched, they may make inquiries as to whether it is really incorrect or not. For example, they can make a query to the storage personnel to find out whether items that seem to be missing have been delivered but are not yet in the storage database. Or they can compare an unmatched invoice to other orders from the same supplier in order to find out whether an invoice contains items from other orders not mentioned in the invoice. However, even though these methods are sometimes applicable, there are no rules that indicate which method might produce a solution. Thus, unmatched invoices can often be classified as established exceptions. Sometimes there seems to be no method for verifying an invoice at all, in which case an unmatched invoice is an otherwise exception for the transaction verifiers.

This problem was well recognized, however, its real extent was not known and purchasers tended to diminish the problem and even blamed transaction verifiers about making inquiries for no or minor reasons.

The department of finance had initially analyzed and listed the most typical causes for the invoices not to match. This listing was used as a basis of a research form to be used in the detailed study. The form was constructed to gain more profound understanding of the problem: what are the real frequencies of the kinds of exceptions, what causes them, what does their handling cost, and how much delay do they cause before a proper payment can take place.

A period of four weeks was carefully selected – there were no major holidays, no closing of books, no special reporting, and the business was generally assumed to run as usual. Throughout the period, every unmatched incoming invoice was inspected individually by physically attaching the research form with each of them and by making sure that the information requested in the form was carefully filled in at each stage of the process.

Before the actual study started, all the employees that would have to fill in the form were briefed in departmental briefings where also their managers were present to stretch the importance of the study. Individuals missing from the briefings were informed individually. The first two weeks included a lot of individual discussions with the employees who failed to correctly fill in the form. They were all purchasers, who often found themselves as the causers of the exceptions and did not want to report that they had ordered the goods with outdated prices or had not specified the terms of delivery, just to name a few of the problems they had caused.

The study covered a total of 2687 invoices of which 902 unmatched and were thoroughly analyzed. There were a total of 1367 causes for the invoices not to match. These causes where classified to 21 categories. Also resulting delays in the handling of the invoices were calculated. Possible unnecessary money transfers caused by various exceptions were evaluated as well. Various management actions taken as a consequence of the study have been reported in (Saastamoinen 1995b) and (Saastamoinen 1995a).

The same study was repeated eight years later. The engineering shop had implemented a new ERP-system replacing its old legacy system still in use on the time of the previous study. From the viewpoint of transaction
verification, the new system had certain major advantages and disadvantages that were often brought up to general discussion inside the engineering shop - even though their factual impact to the work of the people had not been studied. The core problem itself – unmatched incoming invoices – again appeared as a major problem in the order fulfillment process.

The study in 2001 also covered all the invoices during a period of four weeks and utilized a very same kind of a research form. This study not reported in detail in public provided similar information as the previous study, but also provided the organization with a longitudinal view into the system and process, resulting in a more comprehensive understanding of the situation.

For example, as a great majority of the exceptions were caused internally as a result of staff carelessness or incompetence, the kind of exceptions had increased dramatically as a result of decentralizing the Department of Purchase. Many exceptions caused by the suppliers had not received the attention of the purchasers as the same supplier now received orders from dozens of part-time purchasers instead of a few full-time professionals. To reduce the number of invoices that had to be sent to purchasers prior to payments, the engineering shop had also given the transaction verifiers an authorization to approve exceptional invoices if the monetary error was within a certain limit. Accompanied with the decentralization, this had caused the company to lose a view of supplier performance outside of the real hard performance factors, such as on-time delivery and quality issues.

The study also provided factual information about how the new system served transaction verifiers when compared to the old system. This information was also, like in the previous study, converted to salary related costs as a result of certain tasks taking more or less time.

The exception based approach was in both cases found to be a proper approach to evaluate the system. If the system would have been evaluated just by external measures, one could have claimed that there was hardly anything to improve – goods were purchased, they were received in time and invoices were paid. Neither the new nor the old software collapsed, the data was not corrupted, the systems produced desired reports, etc. However, a more internally focused exception-based analysis could, without doubt, point out major flaws in the system and its use.

6. Summary

Exceptions are events for the handling of which no rules applicable as such exist. Exceptions are inevitable and common and they are often associated even with routine tasks and processes. Even though all exceptions cannot be avoided, most of them could be handled normally by information systems, if the systems and processes and the people using the systems as a part of the processes were aware of the exceptions they cause and the match between the system and its domain would be tighter.

Information systems can be evaluated by focusing in exceptions. The approach field tested in various industrial organizations over the past eight years and presented in this paper can be used to analyze systems functionality in the user's level, to compare systems, and to provide factual information on the system and process flaws.

Focusing on exceptions can be very beneficial for information systems management in practice. Such analysis formalizes prioritization of system development and maintenance activities, shifts focus from technology to processes, increases communication with the user community, and provides valuable feedback about the performance of information system operations.

References


Peer Assessment: A Complementary Instrument to Recognise Individual Contributions in IS Student Group Projects

Elsje Scott¹, Nata van der Merwe² and Derek Smith¹
¹Department of Information Systems, University of Cape Town, South Africa
²Xiplan (Pty) Ltd, Kenilworth, South Africa
escott@commerce.uct.ac.za
nata@xiplan.com
dsmith@commerce.uct.ac.za

This paper discusses peer assessment as a component of the assessment strategy used for Information Systems student group projects at a South African university. The value of peer assessment and the contribution to the real-life experience offered by group projects, will be discussed. It will also illustrate how this process adds value by enhancing deep learning. Its value as a complementary assessment instrument in a multiple assessment strategy and how the results of peer assessment are used to recognise individual contributions to group performance will be illustrated. The use of peer assessment as an instrument for both informal formative assessment and formal summative assessment will be described. To perform the peer assessment specific instruments were designed and used throughout the lifecycle of the course.

Keywords: Peer assessment, group work, assessment, self-assessment, IS Project.

1. Introduction

The recognition of individual contributions to group performance in student group projects is essential. Assigning the same grade to all group members implies equal effort which is often not the case. When assessing group projects equal contribution by all group members cannot be assumed, and the focus cannot be on the product alone (Cooke et al, 1997). It is therefore imperative that the assessment of group projects includes an assessment instrument that will allow the measurement of individual contributions to the group project by peers (the group members) based on specific predefined criteria. The use of peer assessments allows group members to voice their perceptions of their contributions and those of the other members in the group (Feigenbaum and Holland, 1997). As such it can be implemented to discourage social loafing, a term used by Smith (2004) to identify students who under perform in project teams.

The systems development group projects in both the 3rd and 4th year Information Systems (IS) courses at the University of Cape Town are capstone courses that bring together hard and soft skills and closely emulate professional practice. A multiple assessment strategy is used, comprising of formal summative assessment, formal continuous assessment and an informal formative assessment. Various methods and instruments to accomplish these assessments are used, e.g. tests and exams, checklists, questionnaires, mark sheets and scoring rubrics. In both courses, strong emphasis is placed on giving the students a real-life experience that encompasses the full systems development life cycle. Both Cooke et al (1997) and Miller (2003) point out that peer and self-assessment are skills required in industry and have become necessary for professional practice. Since students enjoy carrying out peer assessments and find it beneficial to their learning (Sluijsmans et al, 2002); the classroom provides excellent opportunities to acquire and hone these skills.

This paper will discuss the implementation of peer assessment as an assessment instrument in the 3rd and 4th year IS courses. The use of groups and peers within the context of this assessment strategy will be explained and the aims and benefits of peer assessment will be highlighted. The paper will detail the peer assessment process and the method used in the two courses. The way peer assessment is used as an informal formative assessments tool as well as how it is applied to make a formal summative assessment, will be described and motivated. The development and implementation of a numerical algorithm to quantify the results of the peer assessments, and the use of the quantified results to derive individual ratings will be discussed and illustrated in detail. The paper will illustrate how these ratings were used to adapt the individual marks to more adequately reflect the different individual contributions within each group. It will report on the practical implication thereof over the past two years and how it has affected individual performances within the groups.
2. Background

Group participation and interaction is part of the real world and in most cases group participation is evaluated by peers (Cooke et al. 1997). Thus the implementation of both formative and summative peer and self-assessment to assess student performance in group projects is becoming more common in higher education (Miller, 2003; Sluijsmans et al., 2002). Assigning the same grade to all members of a group gives reason for concern about equity as there is the underlying assumption that all contributed equally (Feigenbaum and Holland, 1997). Moreover, in environments where educational inequity exists and where there is significant cultural diversity, this concern is aggravated and the process of learning should be guided carefully.

Correct assessment practices can motivate students to achieve the desired learning outcomes and can be used as a valuable and effective teaching tool that ought to have its place in the classroom (Schmidtke, 2001). Evidence exists that peer and self-assessment can trigger a greater responsibility among students for independent learning (Dochy et al., 1999). Other benefits of this active involvement of students in the learning process are the obtaining of personal and academic meaning in their studies (Denicolo et al in Orsmond et al, 2002). Cooper (2000) confirms this viewpoint by stating that self and peer-assessment when used formatively incorporate feedback which facilitates student learning. For assessment to be formative, the assessment and feedback should initially be separated from grading. This allows students to develop their own judgments before being presented with the grades from other assessors (Taras, 2002).

Peer evaluation can be done by designing an instrument to objectively grade individual performance in groups. Peer assessment when used summatively provides the opportunity to grade individual contributions in a final product. Taras (2002) advocates greater emphasis on student participation through peer and self-assessment, particularly in summative assessment that "counts".

Several methods to quantify the peer review process are discussed in the literature (Conway and Kember, 1993; Goldfinch, 1994; Li, 2001; Miller, 2003). Despite the fact that choosing between alternative methods may be a subjective process, the authors support Conway's (1994) approach of implementing a simple, yet sufficiently fair method to assign a mark to a group member that will reflect individual effort. Goldfinch (1994) implemented a two-part peer assessment form where the first part determined which tasks were performed by each member during the course of the project. Scores allocated for the second part, where members were assessed on their group-working skills, contributed to individuals' final mark. Conway (1993) implemented a one part form which seeks to determine the extent of each student's participation in the tasks that make up the project.

2.1 Aims and objectives of the peer review process in general

According to the Peer Review Handbook by Christine Bruce (1997) the general aim of the peer review process is to contribute to the professional development of participants. More specifically in a systems development group project it can assist to provide equity to grade distribution (Feigenbaum and Holland, 1997).

Furthermore, according to Netpro project and project based learning (retrieved 2004), the main aims of using the peer assessment process are to:
- Develop students’ judgment and understanding of quality.
- Enable students to assess their own work, resulting in improving the quality of their products.
- Enhance students’ responsibility and accountability.
- Foster constructive communication.

2.2 Benefits

The peer review process allows students to enter into the Action Learning Cycle, a cycle that promotes continuous planning, reflection, observation and action amongst participants (Bruce, 1997). This could encourage group members to adapt to change more easily and share their learning with others. It might also act as an incentive for participation as team members are accountable to each other for individual performance (Cook et al., 1997). Amongst the benefits are also improved designs, real world experience and the development of crucial skills (Feigenbaum and Holland, 1997). Specifically, these skills in a systems development environment include both soft and hard skills.

2.3 Definitions

Within the context of the IS student group projects, groups will be used to refer to self-chosen teams of four to five members. Each
group has its own organisational structure and is responsible for its own administrative tasks and project deliverables. There are typically between twenty and thirty-five groups per course.

*Peers* will refer to the members of a group, and the peer assessment will be limited to the particular group. During the 3rd year, the final peer assessment will also include self-assessment.

*Assessment* will be used to include both *measurement* and *evaluation* as components (Scott and Van der Merwe, 2003). Du Toit et al (2001) describe *measurement* as a qualitative and/or quantitative grading process, and *evaluation* as a value judgement concerning quality – in essence an interpretation of the results through measurement.

*Peer assessment* is conducted by the members of a group with reference to the work done within the group. Each group member will rate the contribution to the group project of every other member in the group.

*Self-assessment* is done as a personal judgement by an individual group member of the value of their specific contribution to the group project relative to the contributions of other members of the group.

This paper will now proceed to discuss peer assessment as a component of the assessment strategy used for Information Systems student group projects. The value of peer assessment, the contribution to the real-life experience offered by group projects and how this process enhances deep learning, will be illustrated.

### 3. The group systems development project

The complex organisational and dynamic software development environment establishes the need to equip students with a diverse set of competencies to face the challenges of the work place more effectively. The systems development group projects in both the 3rd and 4th year courses integrate hard and soft skills to prepare students for professional practice. This reflects the viewpoint of IS2002: An Update of the Information Systems Model Curriculum (IS2002) that the characteristics of an IS professional evolves around three major areas namely:

- A broad business and real world perspective
- Strong analytical and critical thinking skills
- Strong interpersonal communication and team skills

The main deliverable for the 3rd year course is a comprehensive web-based management system with a concise and clear business focus. Students are provided with a generic business problem, detailed specifications and functional guidelines. To expose students to the complexities of interacting with users in real organizations, they are required to use these specifications to find a “best fit”, i.e. an appropriate business problem and sponsor in industry. The sponsor supports the students with expert knowledge and experience, and performs assessment from an industry perspective. The 4th year IS project is a carefully scoped real world project and builds on the experience gained in the similar but smaller 3rd year project. Unlike the 3rd year, these students have to go out in industry, and identify a need to translate a business problem, efficiently and creatively into an automated computerized system.

Both these projects equip students with crucial problem-solving abilities using object-oriented techniques and business process re-engineering; as well as with the insight and understanding required to capture business processes programatically. The competitiveness of project teams encourages creative solutions and necessitates that students acquire advanced technical skills as well as implement the most recent technologies.

Collaboration and communication form integral parts of both these projects. 3rd Year project groups report to a member of faculty acting as project manager, and they also have regular meetings with the business sponsor. The development process is guided by several interim deliverables and milestones, culminating in a final shrink-wrapped product and project presentation. The 4th year groups manage their own projects and have fewer deliverables, with the project also culminating in a shrink-wrapped product and a project presentation. The final event in the calendar of both 3rd and 4th years is an exhibition to showcase their expertise and professionalism to industry, learners from various schools and the public.

The assessment strategy and peer assessment component as implemented in the IS group projects, are discussed in the following sub sections.
3.1 Assessment strategy

Scott and Van der Merwe (2003) advocated the use of an assessment strategy that involves multiple assessment approaches to enhance student learning and aid the objective assessment of group performance. It was argued that an assessment strategy must be adopted that will give credit to the complexities and challenges of IS group projects. This strategy is rooted in those proposed by Shepard (2000) and Pellegrino et al. (2001) to reflect comprehensiveness, coherence and continuity.

The IS group projects in both the 3rd and 4th years are very practical in nature and have a number of deliverables requiring a wide range of abilities and skills that must be assessed. In the development of the assessment strategy, these skills and abilities were identified and assessments tools were chosen to effectively evaluate and measure them.

Table 1 on the next page, adapted from Scott & Van der Merwe (2003), summarises the strategy used for the 3rd year IS projects. The 4th year course uses a similar strategy. It reflects the comprehensive assessment strategy, using multiple approaches to ensure coherence, and providing continuity through regular review-points, feedback and opportunities for improvement. This strategy supports the active learning process and aims to coach students into a deep learning approach thus maximising their learning experience.

One of the challenges mentioned in Scott and Van der Merwe (2003) is to give recognition to the contribution of individual members to the group project. The inclusion of peer assessment as part of the assessment strategy becomes important when meeting this challenge.

Table 1: Assessment strategy

<table>
<thead>
<tr>
<th>Component</th>
<th>Occurrence</th>
<th>Key assessment strategy (Shepard)</th>
<th>Group / Individual</th>
<th>Contribute to final mark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mid-year exam</td>
<td>Once - 3hr exam</td>
<td>Prior Knowledge</td>
<td>Individual</td>
<td>Yes</td>
</tr>
<tr>
<td>Interim deliverables</td>
<td>8 - approx every 2 weeks</td>
<td>Dynamic Feedback</td>
<td>Group</td>
<td>Yes</td>
</tr>
<tr>
<td>Milestone deliverables</td>
<td>3 - approx every 6 weeks</td>
<td>Dynamic Feedback</td>
<td>Group</td>
<td>Yes</td>
</tr>
<tr>
<td>Technical workshops</td>
<td>Weekly – first 10 weeks</td>
<td>Prior Knowledge</td>
<td>Individual</td>
<td>Yes</td>
</tr>
<tr>
<td>Weekly reports</td>
<td>Weekly</td>
<td>Feedback</td>
<td>Group</td>
<td>No</td>
</tr>
<tr>
<td>Weekly project management meeting</td>
<td>Bi-weekly</td>
<td>Feedback</td>
<td>Group</td>
<td>No</td>
</tr>
<tr>
<td>Sponsor meetings</td>
<td>When required</td>
<td>Feedback</td>
<td>Group</td>
<td>No</td>
</tr>
<tr>
<td>Sponsor evaluations</td>
<td>Twice</td>
<td>Dynamic Feedback</td>
<td>Group</td>
<td>Yes</td>
</tr>
<tr>
<td>Course evaluation</td>
<td>Twice</td>
<td>Evaluation of teaching</td>
<td>Individual</td>
<td>No</td>
</tr>
<tr>
<td>Peer assessment</td>
<td>When required and once</td>
<td>Student self-assessment Feedback</td>
<td>Individual</td>
<td>Yes (Final assessment)</td>
</tr>
<tr>
<td>Self-assessment</td>
<td>When required and once</td>
<td>Student self-assessment Feedback</td>
<td>Individual</td>
<td>Yes (Final assessment)</td>
</tr>
<tr>
<td>“Mock” presentation</td>
<td>Once</td>
<td>Dynamic Feedback</td>
<td>Group</td>
<td>Yes</td>
</tr>
<tr>
<td>Final presentation</td>
<td>Once – 3hrs</td>
<td>Explicit Criteria</td>
<td>Group</td>
<td>Yes</td>
</tr>
<tr>
<td>Code review</td>
<td>Once</td>
<td>Explicit Criteria</td>
<td>Group</td>
<td>Yes</td>
</tr>
</tbody>
</table>
3.2 Peer assessment component

3.2.1 Use of peer assessment in IS student group projects

As noted by Miller (2003), it is becoming more common to use peer assessment for the assessment of individual student performance in group work, as it can be used for formative as well as summative purposes. Boston (2002) describes formative assessment as the diagnostic use of assessment to provide feedback, while summative assessment is described as taking place after a period of instruction and requires making a judgment about the learning that has occurred by grading or scoring a deliverable, test or exam.

Taras (2002) believes too much emphasis is placed on grading and too little on what is being learned and on student participation. For assessment to be formative, Taras (2002) asserts that assessment and feedback should initially be separated from the grading process. To create a balance between assessment, feedback and grading, she advocates student participation through peer and self-assessment in formal summative assessment.

The authors agree with this approach, and therefore peer assessment is performed at different stages in the group IS projects. During the development stages it is used as informal formative assessment, to facilitate student learning and develop skills that are needed in industry. It also prepares them to make accurate and fair assessments of their group members in the final formal summative assessment. These views are shared by Orsmond et al (2002) and Sluijsmans et al (2002). In the 3rd year course, self-assessment is used in tandem with peer assessment for the final assessment of the group projects.

3.2.2 Instruments

The main instrument used in the assessment of the group projects, is scoring rubrics. In the development of the assessment strategy, scoring rubrics were found to be the most effective way to align the different parties involved and to stimulate discussion and debate, as well as to limit bias and convey to the students the standards against which they will be measured (Scott and Van der Merwe 2003).

The peer assessment rubric for the 3rd year group (see Table 2) was designed to reflect both the student’s contribution to the efficient functioning of the group and the extent of each student’s participation in the various tasks of the project. This approach combines the two different focuses of Goldfinch (1994) and Conway (1993). It consists of 10 criteria that must be rated, using the assignment of a numerical score for each item. Each member must rate themselves as well as the contribution of every other member in the group. The layout of the columns gives the student a clear picture of their ratings of each group member compared to themselves and the rest of the group. This was done to assist the student in making a fair and accurate comparative assessment of each group member. How the ratings are used to assess individual contributions to the group projects, will be explained in paragraph 3.2.3 below.

Table 2: Final peer assessment including self-assessment used in 3rd year course
The 4th year course uses a similar instrument, although the questions were posed at a different level (see Table 3). Group members were to rate their peers considering five areas, namely Ability to Work in a Group, Amount of Effort, Dependability, Intellectual Contribution and Overall Contribution to Project, on a scale of 1 to 5.

Table 3: Final peer assessment excluding self-assessment used in 4th year course

<table>
<thead>
<tr>
<th>INF414W Systems Development Project 2003</th>
<th>CONFIDENTIAL PEER ASSESSMENT SHEET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to Work with the Group</td>
<td>Amount of Effort</td>
</tr>
<tr>
<td>1 was disruptive of the group process</td>
<td>Minimal</td>
</tr>
<tr>
<td>2 participated, but wanted to go in a different direction than the group</td>
<td>Less than what was expected</td>
</tr>
<tr>
<td>3 ok</td>
<td>About what was expected</td>
</tr>
<tr>
<td>4 always participated, made sure everyone had a chance to participate</td>
<td>Above what was expected</td>
</tr>
<tr>
<td>5 helped get the group moving without dominating it</td>
<td>Did the whole thing (need to explain this)</td>
</tr>
</tbody>
</table>

According to Rust et al (2003), there is a need for transparency in the assessment process. For this reason, all assessment instruments used in the student group projects, for example the rubrics or mark sheets, are made available well in advance. This helps to create awareness of the assessment criteria and associated standards. Goldfinch (1994) recognises the fact that it is educationally unsound to withhold the assessment instrument from the students until they must use it, and recommends making it available in order for students to know exactly how they will be assessed.

3.2.3 Method

Group performance is based on group cohesion, group efficacy and team building (Bahli & Buyukkurt, 2003; Smith, 2004). During the development stages, groups are lead by a facilitator in the peer assessment of the group performance, which includes reviewing group dynamics, cohesion, effectiveness and communication. This is done openly in a reflective manner, with ample opportunity for feedback.

The final peer assessment is done secretly and includes a self-assessment component. Lejk & Wyvill (2002) concluded that secret peer assessment can be done more honestly and is therefore more accurate. There are strong arguments for and against the inclusion of self-assessment in the peer assessment of group work. While self-assessment is excluded from the final peer assessment in the 4th year course, it was decided to include it in the 3rd year course. For 3rd year students, this is their first encounter with peer assessment. Given their level of experience, it was felt that their inexperience may lead to them being less critical towards their peers and over-generous in their assessment. Goldfinch (1994) has found that students who are over-generous...
effectively penalise themselves if the rest of the group are more conservative. She therefore argues for the inclusion of self-assessment.

It is reasonable to accept that although students bring different strengths to a team, they are encouraged to participate in most areas, for example, all members of a group are expected to contribute towards the building phase (coding) of the project. A main focus in the design of the presentation and code evaluation instruments was to objectively reflect the quality of the product rather than just the effort. The authors are of the opinion that there is not necessarily a linear relationship between the quality of a product and the effort put into the delivering of it. For this reason it was decided to use the scores in the peer assessment process as penalties to adjust imbalances within certain groups rather than just increasing or decreasing the marks of individuals.

The process involved the following steps: forms were filled out individually, sealed to ensure confidentiality and handed in with the final shrink-wrapped product. For each group the ratings obtained were entered into a spreadsheet and averages were calculated. An average for each member was obtained, based on the self- and peer ratings. The 4th year calculations excluded the self-assessment ratings. From these averages an aggregate average across all the member’s ratings was calculated and was used to obtain differences by subtracting this aggregate average from the individual averages. From these differences a mark deduction table was developed in both courses to associate corresponding penalties to specific ranges for only those averages lower than the aggregate average.

A small committee of four academics was constituted to handle the review process for the 4th year students, whereas the course coordinator and the course administrator executed the process for the 3rd year students.

3.3 Results

Since 2002 peer assessment has been used effectively as an integral part of the 3rd year course in an attempt to enhance the real world experience of the course and reap the benefits of the process as discussed in section 2.3. It was incorporated for the first time into the 4th year course in 2003.

In the case of the 4th year course, a 5-point Likert scale was implemented. Ratings of 1 and 2 were seen as below average, whereas 4 and 5 were above the average. The differences for each team were analysed and penalties were only applied where a team member had rankings of 2’s and 1’s. Whilst 16 students with differences greater than 0.4 were identified the committee decided that 13 students would not be penalised as their marks were deemed acceptable as they were, on average, greater than 3. Of the 3 students who were penalised, one student received a 20 mark penalty, one 15 marks and one 5 marks from the final project mark. In 2002 and 2003 the peer assessment instrument for the 3rd year students excluded the last two categories as shown in table 2. Table 4 below summarises the peer assessment outcomes for 2002 and 2003. In 2002 10 groups were affected by the process. A severe problem, causing one group to split in two, occurred before the final evaluation process and was resolved by evaluating the project separately for the two sub-groups, and hence is not included in table 3. Three members in one group were each penalised by 5% as they contributed much less towards the project than the other two members. In another group two members were penalised by deducting 5% and 7.5% from their project score respectively. The case where 15% percent was deducted was again a severe case where the problem was already identified in the early stages of the project. In almost all the cases members admitted to not having contributed equally and the process was finalised and resolved with almost no conflict.

Table 4: Summary of peer assessment outcomes in 3rd year course

<table>
<thead>
<tr>
<th>Year</th>
<th>2002</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Students</td>
<td>204</td>
<td>183</td>
</tr>
<tr>
<td>Groups</td>
<td>41</td>
<td>38</td>
</tr>
<tr>
<td>Penalty deduction (%)</td>
<td>Students affected</td>
<td>Students affected</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>7.5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>12.5</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

During 2003 the technical skills transfer occurred via a series of workshops where students developed a pilot system prior to the build phase of their own projects. In 2002 this transfer occurred during a series of seminars where technical topics were addressed. Due to the additional focus on the enhancement of technical skills in 2003, all group members were expected to contribute substantially towards the building phase. Although problems were also experienced in 10 groups, they
seemed less severe than in 2002, this could be attributed to the increased focus on the enhancement of technical skills and the continuous feedback from regular peer assessment sessions. The differences of 3 of the 8 students, penalised by 5%, fell just inside the specific range for this penalty. In those groups some members indicated that they would prefer their team members not to be penalised for not contributing equally.

4. Conclusion

The authors found that the use of peer assessment during different stages of the course assisted to improve group cohesion. It also helped to identify areas within specific groups that needed attention. These might include aspects like an imbalance of workload and skills shortages. Mediation of project managers during early stages of the project aided to resolve some conflict and prepared students to effectively implement the peer assessment instruments provided. It also assisted to create an awareness of the diverse skills within each group and how these skills could facilitate learning and contribute positively to the final product. In addition this awareness prepared students for teamwork in industry.

Peer assessment as a component of the assessment strategy used for IS student group projects offers students the opportunity to become active participants in the assessment process. It also provides them with skills that are needed and can be applied in the real-world environment for which they are being prepared. As an assessment instrument it also benefits the lecturer or teacher, being an aid for a fairer and more accurate assessment of individual contributions to group projects.

As a future enhancement, informal peer assessments can be improved to include a more structured feedback process. This will benefit the subsequent occurrences of peer assessment in this course as well as other walks of life.

References


Elsje Scott, Nata van der Merwe and Derek Smith


view/prGuidelines.htm


Performance Evaluation of e-Business in Australia

Mohini Singh and John Byrne
School of Business Information Technology, RMIT University, Melbourne, Australia
mohini.singh@rmit.edu.au
john.byrne@rmit.edu.au

The Internet and related technologies have made a substantial impact on the way organisations conduct business in Australia and around the world. Australian organisations like their international counterparts have invested heavily to leverage the Internet and transform their traditional businesses into e-businesses in the last seven years. E-business investments are claiming a sizeable share of overall IT budgets in most organisations whether they are small, medium or large. However, managers are under constant pressure to justify e-business costs and to ensure that these investments keep paying off. Earlier research on e-business in Australia addressed issues of the rate of e-business uptake and the application of the Internet to certain business processes. Research discussed in this paper is one of the first attempts to evaluate the value of e-business. It is based on data collected, collated and analysed from the responses received from IT and e-business managers throughout Australia.

Research presented in this paper is based on a model developed in the USA (Barua et al, 2001) to identify the impact of e-business drivers on operational excellence of firms which influence financial improvements. It was initiated to quantify the success of e-business in Australia after huge losses from e-business projects were reported by a few large organisations. The paper includes a review of literature on e-business evaluation, research methodology, analysis techniques, a discussion of e-business performance in Australia and presents the impact of e-business on operational excellence and financial performance of the organisation.


1. Introduction

Australian organizations have invested heavily to leverage the Internet and transform their traditional businesses into e-businesses in the last seven years. E-businesses are defined as the use of internet based information and communication technologies (ICT) by organisations to conduct business transactions, share information and maintain relationships (Poon and Swatman, 1999). Two main trading models in e-business are B2C (business to consumer) and B2B (business to business). In Australia B2B e-business was worth 1.1% of the gross domestic product in 2001, and B2C e-business was worth 0.17% (NOIE, 2002). Business organisations doing business on the Internet with digitised business processes are expected to achieve business improvements from reduced operation costs, labour, time and paper. E-business investments are claiming a substantial share of overall IT budgets in most organisations based on anecdotal evidence that organisations achieve unprecedented benefits by leveraging the Internet as a medium of business. According to Kearney research report (2003) e-business budgets in Australia are about 27% of overall IT budgets. However, senior managers are increasingly under pressure to justify e-business costs. They are asked how these costs pay off and how can a company make sure they keep paying off? Managers in e-business organisations are striving to articulate where such benefits come from, and how to turn things around and cash in on e-business initiatives.

E-business research data in Australia to date is mostly on the number of business organizations trading electronically, type of e-business applications, potential benefits of e-business and the application of the Internet to certain business processes (NOIE, 2001 and 2002, Kearney, 2003). These reports indicate a substantial increase in the uptake of e-business and Internet applications by Australian organizations. Earlier research in Australia (Singh, 2000) highlighted the need for e-business metrics to evaluate benefits. In the year 2001 a media publication (The Age, 2001) reported huge losses incurred by large Australian organizations such as Fosters and the National Australia Bank from their e-business projects.

This research was initiated to evaluate the financial and operational performance of e-business and to quantify the success of e-business in Australia. It was accomplished via online and postal questionnaire surveys, and data was statistically analysed. This paper includes a review of literature on e-business evaluation, research methodology, analysis techniques, a discussion of e-business
performance in Australia and presents the impact of e-business on operational improvements and on the financial performance of the organisation.

2. Literature review

New technologies such as the Internet and the World Wide Web have made a profound impact on all businesses in Australia and around the world. E-business enables organisations to reduce costs, increase demand and create new business models. It has the potential to benefit all consumers through reduced prices and improved products and information flows (Dunt and Harper, 2002). Small and large firms alike can access the Internet and exploit near-zero marginal costs of distribution for their products (Dunt and Harper). Although e-business has been proved to be popular with large business enterprises, small and medium companies also create value by marketing and selling goods and services electronically (Dublisk, 2000). Each company is constrained by the amount of graphics and design capability that the Internet can deliver, so everyone starts from the same position with their Web sites.

Australian organisations like their international counterparts have increasingly resorted to e-business to capitalise on the opportunities of business efficiencies. These organisations adopted the B2C e-business model to increase market share, offer better customer service and to reach out to customers at greater geographic distances (Singh, 2000). Developments in B2B e-business in Australia has seen businesses and the government, both at the State and Federal levels adopting Web-based e-procurement to achieve volume purchase, dealing with a wider choice of buyers and suppliers, lower costs, better quality, improved delivery, and reduced paperwork and administrative costs (Singh and Thomson, 2002).

Benefits of e-business as outlined by Chaffey (2004), Singh (2002,a) and Turban et al (2004) are increased revenue from enhanced sales; reduced marketing costs with online advertising, reduced time in customer service and online sales; supply chain cost reductions from reduced inventory levels, increased competition from suppliers and shorter cycle time in ordering; and reduced administrative costs from automated routine business processes, order confirmation, accuracy of data and an improved competitive position. Other non quantifiable improvements achieved from e-business include a better corporate image, improved communication with customers and business partners via electronic channels, a faster product development lifecycle enabling quick response to market needs, improved customer service, better information and knowledge management, ability to incorporate positive feedback from customers to enhance sales, applications of intelligent software for data mining and forecasting trends and demands (Singh, 2000, Chaffey, 2004 and Turban et al, 2002 and Turban et al, 2004).

Amit and Zott (2001) advocate that business conducted over the Internet in the 21st century with its dynamic, rapidly growing and highly competitive characteristics promises new avenues for the creation of wealth. E-business models, methods and the volume of digitisation vary from industry to industry and from organisation to organisation depending on their size, nature of business, technology capability and in-house technical expertise. Although the value of adopting e-business has been recognised, actual achievements from it are not known. Most e-business reports (NOIE, 2003; ABS, 2003; Kearney, 2002) provide an understanding of the level of e-business adoption. Returns from e-business implementations have not been formally evaluated. Grey et al (2003) are of the opinion that much of the value associated with e-business comes not only from improvements in the technological infrastructure but from business and organisational transformations. They explain that a critical part of creating business value is identifying the processes to transform and selecting the right initiatives to enable the transformation. The IT infrastructure capability including speed, flexibility, capacity, efficiency, resilience, and security determine the type of applications that can be run and their performance. These applications affect the accuracy, speed and productivity of business processes in various functional areas of the organisation which have an impact on the overall business performance of the enterprise.

E-business payoffs are generally assessed as IT payoffs under the themes of metrics, environment, technology and processes (Kohli, Sherer and Baron, 2003). As suggested by Shi and Daniels (2003) success in e-business includes functionality, integration and scalability, and an evaluation of e-business applications is necessary for further improvements, management strategies and the deployment of technological developments. Devaraj and Kohli (2002) emphasise that it is
necessary to determine the strategic role of IT in the organisation as compared to other projects to get an overall economic picture. They also emphasise that IT projects have less apparent and longer payoff duration and IT metrics according to Devaraj and Kohli include profitability, productivity and customer value while e-commerce payoff measures address efficiency, effectiveness and innovation strategy measured along five dimensions of time, distance or geography, relationships, interactions, and product or service. IT evaluation is similar to benefits realisation as suggested by Ashburton and Doherty (2003) and it should be explicitly concerned with the on-going management and direction of the project, managing benefits as they are collected. They also emphasise that evaluation should be concerned with assessing the process of systems development as well as its product so that the systems development process can be improved over time, and that evaluation should be performed as an on-going process. Cronholm and Goldkuhl (2003) describe the strategies for Information Systems evaluation to be goal-based evaluation, goal-free evaluation and criteria-based evaluation. Goal based evaluation measure the IT system based on explicit goals from the organisational context. Goal free evaluation is an inductive and situational driven strategy, while criteria based evaluation means that some explicit general criteria are used as an evaluation yardstick. Barnes and Hinton (2004) emphasise that e-business performance measurement systems should include metrics on the performance of the website, business processes, customers and the link between e-business performance and business strategy.

Although all of the above it and e-business evaluation criteria discussed above address important issues, they are ideas or suggestions for evaluation. The only research that evaluated e-business performance exclusively at the time this project was undertaken was by Barua et al (2001) on which research discussed in this paper is based. They developed a conceptual model to evaluate e-business suggesting that e-business drivers contribute to operational improvements and operational excellence improves financial performance as shown in figure 1.

In the model e-business drivers come from the areas of IT applications, processes and readiness. As explained by Barua et al (2004), IT resources, processes and readiness are needed to create organisational information capability which is a resource essential for the exchange of strategic and tactical information with all stakeholders in an e-business. These drivers are actionable part of business value which have a direct impact on operational excellence and higher order impacts on financial performance of firms (Barua et al (2001). The operational excellence measures in this model are e-business specific as well as traditional. E-business specific measures are total business transacted online, existing customers conducting business online, new customers acquired online, MRO items and production goods procured online, and customer service provided online. Improvements in traditional measures refer to improvements in order delivery times, order fulfilment accuracy and better inventory management. Financial performance of an e-business firm in this model is similar to traditional financial measures of revenue per employee, gross profit margin and return on assets.

![Figure 1: Conceptual e-business value model](image-url)

### FINANCIAL PERFORMANCE INDICATOR
- Increase in revenue per employee
- Increase in gross profit margin due to e-Business
- Increase in Return on Assets due to e-Business
- Increase in Return on Investment due to e-Business

### OPERATIONAL EXCELLENCE MEASURES
- Online revenue (%)
- Online MRO procurement (%)
- Online production goods procurement (%)
- Online customer service (%)
- New customers acquired online (%)
- Existing customers online (%)

### ELECTRONIC BUSINESS DRIVERS
- Customer related processes
- Supplier related processes
- Customer orientation of IT applications
- Internal orientation of IT applications
- System integration
- Readiness of customers
- Readiness of suppliers

This research was accomplished by surveying e-business organisations using online and postal questionnaire surveys. Online surveys, were considered to be the apt method of investigating e-business organizations since it is technology based, quick, convenient, enables unlimited reach, seeks a response to all questions, responses are downloadable into a database and transportable to statistical

## 3. Research methodology

In this paper we sought to test this model in the Australian context.
packages for analysis. However, due to a disappointing response to online surveys, we resorted to a mail survey to complete the research project.

A set of questions to evaluate the performance of e-business in Australia was initially developed in MSWord. Some of these were adopted from the USA study (Barua et al, 2001). The questionnaire was divided into sections and included questions presented on Likert Scales, as ‘yes’ and ‘no’ answers and some fill in the blanks. Respondents were expected to provide answers with radio buttons, choosing an option from the drop down menus or filling in a word, phrase or numerical value in the space provided. The whole questionnaire was presented in six HTML pages for the online survey. At the end of page one a respondent was asked to select the ‘submit’ button and proceed to the next page. The questionnaire was designed so that a respondent could not proceed to the next page unless an answer to all questions on the current page was provided. On submission of first page, HTML codes were included to generate a tracking number enabling the respondent to complete the rest of the questionnaire at a later time or date if he or she wished to.

The questionnaire was sent to a random sample of 725 companies obtained from a database Business Who’s Who http://bww.dnb.com.au/default.asp. The database was sorted using different criteria to ensure that they were e-business organizations and included the top 500 companies. The questionnaire was disseminated via emails the addresses for which were obtained from the above database. It was addressed to the e-business manager based on the assumption that these organisations will have such a position, if not will be passed on to the person in charge. A short explanation of the objectives of the research and the URL for the survey was included in the email.

3.1 Response

At the end of one week 32 valid responses were received. The responses to the questionnaire were transferred to a database created in MySQL. Each page in the questionnaire was represented as a table in the database, and each response to a question was recorded as an element in the table. The database was designed to store both numeric and alphanumeric data. To elicit more responses, a hard copy of the questionnaire was sent by post to the same organisations. The package included a hard copy of the question and a covering letter explaining the purpose of the project and the URL for the online questionnaire. In the month following the mail out of the survey, online responses increased to 91, and valid hard copy responses received were 78. This research analysis is therefore based on a response rate of 23.3 %.

4. Findings and discussion

Findings of this research discussed in this paper reflect the development, application and achievements of e-businesses in Australia. Responses to part 1 of the questionnaire reflect e-business adoption and development in Australia. From this research it is apparent that in Australia e-business is adopted by organisations of all sizes small (32%), medium (34%) and large (34%). Titles of individuals who responded to the surveys were IT managers (22%), e-business managers (14%), managing directors (11%), marketing managers (8%) and other middle managers (45%). It is important to note that many of the respondents held postgraduate (30.8%) and graduate (37.2%) level qualifications. All categories of Australian industries have some form of e-business in their organisations although it is most widely adopted by the manufacturing (24%) and the service industry (22%) sectors. Others were grouped as transport/utility (18%), retail/wholesale (15%) and other (21%).

The most popular e-business model adopted was business to business (47%), followed by business to consumer (18.2%). Some adopted more than two types of online trading models. Most e-business development in Australia took place in the year 2000 (22.4%). Other development percentages indicate an increase up to the year 2000 (1998 (16.4%) and 1999 (18.4%)), and a decline in 2001 (11.2%) and 2002 (6.6%). The downward trend in e-business adoption since 2000 is commensurate with the concept that the dot.com crash slowed or discouraged e-business development. E-business like most new initiatives requires substantial resources in terms of technology, finance, people and time. Most Australian organisations made sizeable investment in e-business resources. Findings indicate that 71% assigned all development planning duties to a dedicated group in the organisation, 65% indicated that they allocated large financial resources to e-business projects and 68% had allocated dedicated personnel to manage and implement e-business projects.
4.1 e-Business drivers and operational success

e-business drivers included in this research were system integration, customer orientation of IT, supplier orientation of IT, informational (quality, supply continuity, and relationship management) and transactional; internal operation of IT, customer related processes, supplier related processes, customer e-business readiness and supplier e-business readiness. These are key factors that e-business organisations invest in and commit resources to in order to achieve improved operational performance. Development of e-business drivers in the organisations surveyed are discussed in a separate paper (Singh and Byrne, 2004), highlighting a moderate development in all organisations.

Table 1: Correlation coefficients between e-business drivers and operational excellence measures (** = p<.05, * = p<.1).

<table>
<thead>
<tr>
<th>E-Bus Drivers</th>
<th>Online Revenue</th>
<th>MRO Procurement</th>
<th>Prod Goods Procurement</th>
<th>Customer Service</th>
<th>New Customers</th>
<th>Existing Customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer Processes</td>
<td>.064</td>
<td>.154</td>
<td>.294 **</td>
<td>-.076</td>
<td>.198</td>
<td>.117</td>
</tr>
<tr>
<td>Supplier Processes</td>
<td>.007</td>
<td>.137</td>
<td>.247 **</td>
<td>.002</td>
<td>.147</td>
<td>.139</td>
</tr>
<tr>
<td>Customer IT application</td>
<td>.227 *</td>
<td>.110</td>
<td>.205 *</td>
<td>.085</td>
<td>.331 ***</td>
<td>.045</td>
</tr>
<tr>
<td>Internal IT application</td>
<td>.383 ***</td>
<td>.146</td>
<td>.255 **</td>
<td>.107</td>
<td>.309 **</td>
<td>.046</td>
</tr>
<tr>
<td>System Integration</td>
<td>.424 ***</td>
<td>.197</td>
<td>.263 **</td>
<td>.118</td>
<td>.307 **</td>
<td>.064</td>
</tr>
<tr>
<td>Customer readiness</td>
<td>.193</td>
<td>.262 **</td>
<td>.324 ***</td>
<td>.095</td>
<td>.250 **</td>
<td>.232 *</td>
</tr>
<tr>
<td>Supplier readiness</td>
<td>.205 *</td>
<td>.300 **</td>
<td>.330 ***</td>
<td>.077</td>
<td>.366 ***</td>
<td>.211 *</td>
</tr>
</tbody>
</table>

From Table 1 it is apparent that in Australia customer and supplier readiness had a significant impact on the procurement of production goods and some impact on the procurement of MRO goods. Buyer (customer) and supplier interaction in an e-marketplace is based on how ready they are to trade online either through net markets (intermediaries) or by net exchanges in-house. E-procurement is a major application of B2B e-business in Australia (Singh, 2004), and this finding highlights the importance of customer and supplier e-readiness to engage in e-procurement. Supplier readiness, however, also had a positive impact on acquisition of new customers, a small increase in online revenue and managed to get some existing customers to trade online. Higher levels of electronic business readiness of customers and suppliers are positively associated with customer and supplier side informational capabilities reducing uncertainty through better information sharing and coordination and managing demand, inventory and capacity information (Barua et al, 2004).

In the following section we discuss the impact of e-business drivers on operational improvements and financial success. In order to determine the impact of e-business drivers on operational excellence measures, we calculated the correlation coefficient between each driver and each operational excellence measure. We display our results in Table 1. Next to the value of the correlation coefficient, we indicate whether the relationship is significant, using three asterisks (***), when the correlation is significant at the 0.01 level, two asterisks (**) when the correlation is significant at the 0.05 level and one asterisk (*) when the relationship is significant at the 0.10 level.

Internal applications of IT prepare employees to be responsive to customer needs and readily access internal information through easy-to-use interfaces. Data analysis presented in Table 1 indicates that Australian organisations with internal e-business initiatives achieved an increase in online revenue, supported e-procurement and acquired some new customers. The impacts of intranets and internal communication, project management, internal process improvements and internal orientation of electronic business processes are pivotal for e-business success (Singh and Byrne, 2004).

Customer orientation of IT applications helps customers access product related information, use FAQ’s for quick answers, find post product information, customise orders and receive online customer service (Barua, et al, 2001). From Table 1 it is seen that customer orientation of e-business led to the acquisition of new customers by firms, and had a small impact on e-procurement of production goods and on online revenue. Supplier related processes are important to reduce approval steps in online purchases, paper work and exception handling. However, from Table 1 it is clear that in Australia this only slightly impacted e-procurement as did customer related processes of improved online services.

### 4.2 Explanation of results

In Table 2 we summarise the results of the impact of e-business drivers on operational improvements. For each of the e-business drivers, we identify those that were reported to have a significant relationship with the identified operational improvement measures. For example, the e-business driver customer related process has a significant impact on e-procurement only.

#### Table 2: Significant relationships between e-business drivers and operational measures.

<table>
<thead>
<tr>
<th>E-business driver</th>
<th>Relationship with operational measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer related processes</td>
<td>None except Online goods procurement</td>
</tr>
</tbody>
</table>

From Table 2 it is apparent that in Australia the impact of e-business drivers on operational improvements is small. This leads to the conclusion that a lot more effort in developing e-business drivers is essential for achieving operational improvements from e-business.

### 4.3 e-Business operational excellence and financial performance

Financial performance indicators used in this research were also adopted from Barua et al, (2001), and these included percentage increases in revenue per employee, gross profit margin, return on assets and return on invested capital attained from e-business initiatives. In considering the relationship between operational improvements from e-business and their impact on financial performance indicators, we divided responses into whether there was an increase in the financial performance indicator or not. We then calculated the change in the operational excellence measure for those firms showing an increase in the financial performance indicator and the change in the operational excellence measure for those firms not showing an increase in the financial performance indicator. Thus those reporting an increase in revenue (the financial performance indicator) collected on average 11.3% of their revenue online (operational excellence measure). Whereas those who did not report an increase in revenue collected on average only 5.0% of their revenue online. The results are shown in Table 3.

#### Table 3: Impact of operational improvements on financial performance

<table>
<thead>
<tr>
<th>Operational Improvement Factors</th>
<th>Increase in Revenue</th>
<th>Increase in profit margin</th>
<th>Increase in ROA</th>
<th>Increase in ROI</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Online revenue</td>
<td>5.0%</td>
<td>11.3%</td>
<td>5.4%</td>
<td>9.1%</td>
</tr>
<tr>
<td>MRO procurement</td>
<td>5.0%</td>
<td>3.0%</td>
<td>5.0%</td>
<td>3.0%</td>
</tr>
</tbody>
</table>
From the data presented in Table 3 it is clear that although insignificant, Australian organisations managed to achieve some financial improvements from e-business in the short time they adopted and implemented online trading at firm level. E-business is thus a useful venture for Australian businesses.

When we consider the relationship between operational excellence measure and financial performance indicator, there does not seem to be any significant differences in operational excellence measure between those firms reporting increases in financial performance indicators and those firms reporting no increase in financial performance indicators. This Australian result, however, is very different from the result reported by Barua et al (2001). They reported a significantly higher value of the operational excellence measure for those firms reporting an increase in the financial performance indicator in every case. A possible explanation for this phenomenon may be the much higher rate of online revenue collected by the US firms compared to Australian firms. For example, for those firms reporting an increase in revenue per employee, the US firms reported that they collected on average 40.4% of their revenue online, whereas the Australian firms collected 11.3% of their revenue online. In Table 4, we summarise the percentage of revenue collected online by the US and Australian firms for those reporting an increase in each of the financial performance indicators.

Table 4: Revenue collected online by US and Australian firms reporting an increase in the financial performance indicator

<table>
<thead>
<tr>
<th>Financial Performance Measure</th>
<th>US firms</th>
<th>Australian firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue per employee</td>
<td>40.4%</td>
<td>11.3%</td>
</tr>
<tr>
<td>Gross profit margin</td>
<td>42.2%</td>
<td>9.1%</td>
</tr>
<tr>
<td>Return on assets</td>
<td>44.8%</td>
<td>8.9%</td>
</tr>
<tr>
<td>Return on Investment</td>
<td>45.0%</td>
<td>9.3%</td>
</tr>
</tbody>
</table>

Although business value of e-business in the USA is not a benchmark for Australian organisations, the figures in Table 4 are included in this paper to suggest the importance of further research needed to address the impact of the political and legal environment, the economic and social environment, and the technological environment in the two countries as well as other parts of the world as determinants of e-business success. However, from the findings presented in Table 4 it can be said that in Australia a greater emphasis on e-business strategies, initiatives and support is required for organisational information capability development in order to achieve business value from net enabled business transformation.

5. Conclusion

The findings presented in this paper clearly indicate that e-business is the new way of doing business and adopted widely by Australian organisations irrespective of their size. From this it is inferred that the Internet is increasingly transforming traditional businesses to e-businesses. Although most industry sectors have embraced e-business, manufacturing and service industries are capitalising on the opportunities more than others. Although it is not new that larger organisations are in a better position to allocate specific resources for new projects, e-business implementation in Australia shows that the medium and small organisations have also invested substantially in e-business. This indicates that the value of e-business has been realised by all Australian organisations although real benefits of e-business have not been achieved.

Testing the conceptual model developed in the USA for e-business evaluation in Australia reveals that substantial resources have been allocated to implement and maintain e-business drivers. These include customer and supplier related processes, buyer and supplier acceptance and readiness of e-business, internal orientation of e-business and integrated systems. Analysis of data indicates a need for a greater emphasis on the development of e-business drivers in order to achieve operational improvements and financial success.

This research is one of the first in Australia that has attempted to evaluate the performance of
e-business. Although e-business benefits are similar to IT benefits which are long term and some are non quantifiable, slow achievement of improvements from e-business can also be attributed to the fact that it was new, Australia had a lack of expertise for implementation and management, and a cultural change in shopping, negotiating and dealing with customers needs customer acceptance of the new medium.

From the findings discussed in this paper it can thus be concluded that in Australia e-business requires more attention in terms of technology applications, business and management strategies and user acceptance.

References
Dubish, S., (2000), Retailing and the Internet, proceedings of the First World Congress on the Management of Electronic Commerce, McMaster University, Hamilton, Canada (CD ROM).
Singh , M., 2002 (b), 'E-Services and their role in B2C E-commerce', Journal of
Mohini Singh and John Byrne

Management Service Quality, vol. 12, no. 6, November, pp 434 - 446.
The Age, March, 2001