A Multi-faceted Decision Support System Evaluation Approach

Sean Maynard* Frada Burstein and David Arnott****

** Department of Information Systems,*
University of Melbourne, Australia
Email: s.maynard@dis.unimelb.edu.au

*** School of Information Management and Systems*
Monash University, Melbourne, Australia
Email:{ frada.burstein; david.arnott}@sims.monash.edu.au

ABSTRACT: *The evaluation of the performance of any information system is important for the further improvement of that system. This paper proposes a Decision Support System (DSS) evaluation method based on multiple-criteria techniques within a multiple-constituency perspective. With multiple-constituency DSS evaluation, many criteria may be valuable to a particular group of stakeholders, whilst others may be unimportant, or inconsequential. The proposed approach can be used by any DSS project stakeholders, and at any stage of the project. A system architecture for computer-based DSS evaluation is developed. A prototype tool, MultiVal, that has been developed based on this architecture is described.*

RÉSUMÉ:

KEY WORDS: *decision support systems, evaluation approach, multiple-constituency, multi-criteria tool*

MOTS-CLÉS:

1. Introduction

The evaluation of information systems development and use is integral in ensuring that quality systems are built for organisations. The evaluation of decision support system (DSS) can be defined as the process, which measures the success of a DSS. In general, evaluation usually reflects the viewpoint of the person who performs or is responsible for the evaluation. It is useful in developing and evaluating systems as critical as DSS to be as inclusive of as many people and reference groups as possible. Including each relevant reference group may produce a more balanced evaluation that could be used to improve the DSS from each stakeholder's perspective. One way of capturing multiple perspectives in DSS evaluation is to apply the multiple-constituency approach [CON 80], [MAY 94]. The multiple-constituency approach treats products, information systems, or organisations as systems generating different opinions in relation to the effectiveness of their functions [CON 80]. Thus, unlike other evaluation approaches the result of multiple-constituency evaluation does not aim to produce a single measure, rather it allows the tailoring of the evaluation environment to reflect the needs and opinions of the constituencies involved and to produce a final result which reflects these differential factors.

Six generic constituencies are appropriate to the evaluation of a DSS: the developer, the user, the decision-maker, the chauffeured decision-maker, the decision-consumer and management constituencies [MAY 95]. Each of these constituencies may apply different criteria when evaluating the success of the DSS project. Given that a number of constituencies are involved with a DSS project, an adequate evaluation process should consider the success of the system from each constituency's perspective. To accomplish this a number of sets of evaluation criteria will be required: one generalised set for each constituency and any specific criteria relevant for that constituency.

There are studies that focus on multiple-criteria DSS evaluation (for example, [ADE 85] and [GOI 92]. None have approached the evaluation process from a multiple group perspective in a DSS context. This paper presents an approach that allows various groups of people concerned with a DSS to identify which criteria are relevant to them and to indicate their importance within the current evaluation context. We have identified a comprehensive set of criteria useful for measuring the success of DSS projects [MAY 95]. The criteria are presented as a hierarchy classified by the different perspectives from which DSS is measured. The major groups of criteria considered at the upper level of this hierarchy are effectiveness, efficiency, satisfaction, and use. To reduce the likelihood of misinterpretation of criteria each of them is explicitly defined the criteria to ensure that no can take place across constituency groups [MAY 95], [MAY 99].

In this paper we describe a method and a tool for multifaceted evaluation of a DSS. The next section presents the conceptual framework of multiple-constituency multiple-criteria evaluation of DSS. Then, a computerised tool, MultiVal, which implements the framework, is described.

2. The multiple-constituency approach to evaluation of DSS

The multiple-constituency approach was proposed as a way to introduce a variety of perspectives to the evaluation of the effectiveness of an organization [CON 80]. The fundamental assumption of this approach is that the evaluation process needs to be flexible enough to accommodate subjective perceptions about organizational performance in order to avoid unnecessary biases and distortions that may arise from using just one view and measure. The result of such an evaluation may not present a common or consensus view and may not be generalisable. However, in the situation where there is a need to capture contextual factors in order to compare the results obtained from the different constituencies or evaluations performed over a period of time, the approach provides a definite advantage.

An important aspect of the multiple-constituency approach is the identification of all relevant groups or constituencies that are involved with the subject of evaluation. We have identified five major groups of people that are involved in DSS projects to some extent namely: DSS Developer, User, Decision-maker, Management, and Decision-consumer. The first four were identified from past DSS studies, the last, Decision-consumer was introduced to fill a gap in the existing literature [MAY 97]. The Decision-consumer group comprises those people who are directly and significantly influenced by the decision using the DSS. The opinion of these people is often different from other groups involved with the DSS

and it could be important to involve them in the evaluation process. These five constituencies represent distinct roles in DSS project and individuals may perform a number of roles in one project.

It should be noted that these groups contribute to different stages of the DSS implementation lifecycle. The expectations of the constituency regarding DSS success may vary quite significantly depending on the stage of the project. The factors that they will take into consideration while evaluating the system should also reflect their position and perspective. The aim of the proposed approach is to provide a dynamic evaluation environment that will be capable of capturing these differences while assuring some consistency in the process. The next section presents a discussion of the criteria that can be used in DSS evaluation to capture the multiple perspectives of the evaluator. This generic set can also be tailored to reflect the role of the evaluator and a stage of the DSS project the evaluation is performed at.

3. Multiple Perspectives of DSS Evaluation

The roles within a DSS project will each have differing perspectives of the DSS being evaluated. Consequently, suitable measurements or criteria for evaluation need to be formulated. A dynamic set of criteria is required as some criteria are more important for some constituencies than others, and some may become irrelevant if the context of evaluation is changed. This section introduces a comprehensive set of generic criteria relevant to DSS evaluation. This set is based on extensive literature analysis and represents a generic set of DSS evaluation criteria [MAY 95]. This set is useful as a starting point for the evaluation process. It can be presented to the evaluators as a source from which they identify the subset of criteria suitable for the particular perspective or constituency role.

3.1 DSS evaluation Criteria

Evaluation is the process that measures the success of a DSS. The measurement of DSS success can be accomplished through the assessment of four different domains. These domains are effectiveness, efficiency, use and satisfaction [MAY 95]. Within each of these domains a number of criteria exist that may be important concerns for one, or a number of constituencies. In past research, criteria have been referred to using various terminologies. Often they are identified as important, and then are directly used in questionnaire analysis to determine their relevance in an evaluation process. Some of the criteria are common in many studies and tend to have inherent meaning for a particular reference group. In a multiple-constituency approach, constituency groups may not have the same understanding of meaning for each criteria when criteria are not formally defined. Thus, it becomes critical for each criteria to be explicitly defined. This section identifies criteria within each of the mentioned domains and defines these criteria to enable an improved and shared understanding from each constituency's perspective.

3.1.1 Effectiveness

The effectiveness of a DSS is essentially the level to which the goals of the DSS project are fulfilled. Table 1 summarises criteria within the effectiveness domain and presents definitions and references to these criteria. In addition, criteria numbers are included which are referred to in the hierarchies presented later.

The ability of the system to adjust to changing requirements and to provide simulation type capabilities is termed the “flexibility of the system”. Similar terms used elsewhere to describe these criteria include the “ability of the DSS to carry out ad hoc analysis” and the “ability of the system to produce alternative solutions”. These terms are considered to be synonyms for the “flexibility of the system” as they measure a similar concept of effectiveness.

“Individual differences” including the gender, attitudes and experience of a constituency, may influence that constituency’s perception of the effectiveness of the system. Individual differences are thought to be important in determining how DSS should be developed for each constituency group.

Table 1: Effectiveness criteria

Criteria Number	Criteria	Definition and reference sources
Effec1	Accuracy of information	The correctness and exactness of the information provided by

		the DSS. [BAI 83], [EVA 89], [UDOD 92], [IVA 83]
Effec2	Adequacy of information provided	Whether the information provided to the DSS is sufficient for the decision task. [HAM 81]
Effec3	Cognitive style	The habitual ways that individuals process and utilise information. [ALA 92], [RAM 92], [BAR 88]
Effec4	Completeness of data files	How complete the data files are with regard to the decision task. [ADE 85]
Effec5	Completeness of information	Output that contains all the information required for the decision. [BAI 83], [IVA 83], [BAR 88], [UDOD 92]
Effec6	Data summarisation	The ability of the system to summarise output. [MAH 89]
Effec7	Effect of DSS on organisation	The organisation's view on how the DSS affects its make-up. [ADE 85], [HOP 87]
Effec8	Effect of DSS on people's position in the organisation	The organisation's view on how the DSS alters its job structure. [ADE 85]
Effec9	Effect on information flow	The effect of the DSS on how information flows throughout the organisation. [ADE 85]
Effec10	Effect on organisational effectiveness	The effect of the DSS on the fulfilment of the goals of the organisation. [FIN 93, [SAN 84]
Effec11	Flexibility of system	The ability of the system to adjust to changing requirements and provide simulation type capabilities. [HOP 87], [UDOD 92], [MAH 89], [BAI 83], [IVE 83], [KEE 81], [ADE 85], [UDO 92]
Effec12	Increase in analytical tools used	Any increase in the amount of analytical tools used as a result of the DSS. [SAN 85]
Effec13	Individual ability improvement	An improvement in the ability for you to carry out your job resulting from the DSS. [MAH 89]
Effec14	Individual differences (gender, attitude, experience)	Those demographic variables that influence the make-up of the individual. [ALA 92], [RAM 92], [BAK 90], [ADE 85], [ZMU 79]
Effec15	Individuals willingness to change	The extent to which you are willing to undergo changes that the DSS may cause to your job [BAK 90].
Effec16	Level of task independence	Independent tasks are those that do not require assistance from others for completion. [SAN 85]
Effec17	Market growth	The growth of the organisation in the market as a result of the DSS. [FIN 93]
Effec18	Match between technical approach and task requirements	The compatibility of the problem and the DSS. [ADE 85], [HOP 87], [SAN 84]
Effec19	Operational needs	The needs that you have to complete your work. [ADE 85]
Effec20	Personality	The cognitive and affective structures maintained by individuals to facilitate adjustment to events, people, and situations. [ALA 92], [RAM 92]
Effec21	Political acceptability	The acceptability of the system by senior management from a political viewpoint. [ADE 85]
Effec22	Provision of training	The adequacy of the training conducted. [ADE 85], [HOP 87], [BAR 88] , [HAM 81] , [IVE 83], [BAI 83], [UDOD 92], [FIN 93], [EIN 78]
Criteria Number	**Criteria**	**Definition and reference sources**
Effec23	Quality of the decision making process	The affect of the DSS on the quality of the decision making process. [SAN 84], [EIN 78], [FIN 93], [KEE 81], [SAN 85] , [ADE 85], [UDOD 92], [UDO 92], [HOP 87], [BAR 88]
Effec24	Range of objectives	The objectives of the DSS project. [ADE 85], [UDOD 92]
Effec25	Role of I.S.	The role that information technology has within the organisation. [RAM 92]
Effec26	Structurability of task	How well the task can be structured for the DSS. [ADE 85], [BAI 83], [IVE 83] , [FIN 93], [HOP 87], [RAM 92], [CHAN 82]
Effec27	Support from the organisation	The support of the organisation for the DSS and in your use of the DSS. [RAM 92], [BAI 83], [IVE 83] , [KEE 81], [UDOD 92], [BAK 90] , [SAN 85] , [SAN 84]
Effec28	Time-frame of task	The time required completing the decision task. [ADE 85], [RAM 92]
Effec29	Timeliness of information	Information that is current and available in a time frame allowing the decision to be made within the decision time frame. [BAI 83], [IVE 83], [EVA 89], [UDOD 92]

Table 1: Effectiveness criteria (cont.)

For the decision task to be supported successfully by DSS the system must be appropriately designed for the decision environment. This is measured using the criteria "match between technical approach and task requirements". Synonyms used for these criteria include the "match between the decision task and the system", and the "application of the DSS to the business problem". These terms focus on the measurement of the compatibility of the decision task with DSS. For instance, the DSS should process information in similar stages to those required by the decision task.

Training is an issue of importance in determining the effectiveness of DSS. It is referred to as "training quality" and sometimes the "amount of training provided to users". The adequacy of training provided to relative constituencies is measured by these terms. The criteria "provision of training" is introduced here to reflect this characteristic.

A major focus of DSS is improving the decision making process. The criteria "quality of the decision making process" is used to measure how DSS effects the decision making process. It is referred to by the terms "quality of decisions", "better decision making", "decision accuracy", "decision process quality" and "increase in alternatives considered". The similarity between these terms and the "quality of the decision making process" criteria is their measurement of how effectively DSS supports decision making quality. For instance, the quality of the decision making process is thought to be improved through the use of DSS as it may allow more alternative solutions for the decision problem to be considered.

The support of DSS from an organisational perspective is thought to be critical to the effectiveness of the DSS. This is measured using the "support from the organisation" criteria, which are also referred to as "top management support". The focus of each of these is the measurement of the backing of organisation for the DSS project. Another critical aspect of effectiveness is having the information required for decision-making available when it is needed. The criteria "timeliness of information" is used to determine if the system provides current information within a time period that allows the decision to be made. This criteria has also been referred to as the "currency of output", reflecting that output which is current and on time can be used for decision-making.

The other criteria within the effectiveness domain have been consistently termed throughout the literature. The definitions of each of these criteria, where they have been given are also similar and have been used here. Where criteria are not formally defined definitions have been created based on our own experiences and understanding.

3.1.2 Efficiency

The efficiency domain focuses on the degree of performance of a DSS project. This includes how well each constituency performs, in addition to the performance of a DSS. Criteria, other than those discussed below, are used uniformly in other studies. Table 2 summarises these criteria listing their definitions and references.

The productivity of each relevant constituency may be affected by a DSS project. Ideally, as a result of the DSS, the productivity of a constituency will be improved. This has been referred to by the terms "increased productivity", "decision-maker efficiency", "efficiency improvement", "user decision process utilisation", and "user performance". From a multiple-constituency perspective it is useful to measure the effect of the DSS on each constituency's efficiency. A group of criteria have been introduced to measure this from the basis of each constituency.

The benefit of a DSS from a cost viewpoint may be determined by its profitability. The criteria "cost effectiveness" measures the effect of a DSS on the profit of the organisation. The terms "cost-benefit", "cost savings", "cost-profit" and "profitability" are also used in previous research as synonyms for this criteria. All of these factors measure the trade off between the cost of a DSS and the benefits saved and gained through its use.

Table 2: Efficiency criteria

Criteria Number	Criteria	Definition and reference sources
Effic1	Chauffeured decision-maker productivity	Effect of the DSS on chauffeured decision-maker efficiency. [FIN 93], [SAN 84], [MAH 89], [EVA 89]
Effic2	Cost effectiveness	Effect of the system on the profit of the organisation. [UDOD 92], [EVA 89], [FIN 93] , [KEE 81], [ALA 92], [UDO 92], [EIN 78], [SAN 84]

Effic3	Decision-consumer productivity	Effect of the DSS on decision-consumer efficiency. [MAY 97]
Criteria Number	**Criteria**	**Definition and reference sources**
Effic4	Decision-maker productivity	Effect of the DSS on decision-maker efficiency. [EIN 78], [SAN 84], [MAH 89], [EVA 89]
Effic5	DSS developer productivity	Effect of the DSS on DSS developer efficiency. [MAY 97]
Effic6	Management productivity	Effect of the DSS on organisational efficiency. [EIN 78], [SAN 84]
Effic7	Reliability of system	Dependability and reliability of the DSS. [BAR 88], [RAM 92], [HOP 87] [CHAN 82], [ADE 85], [BAI 83], [IVE 83], [EVA 89]
Effic8	Response/turnaroun d time	The time required for the DSS to process queries. [ADE 85], [IVE 83]
Effic9	Throughput	The sufficiency of the amount of material flowing though the DSS. [CHAN 82], [EVA 89] , [VAN 90]
Effic10	Time taken for task accomplishment	Effect of the DSS on the time required for completing a decision making task. [ALA 92], [SAN 85], [UDOD 92], [EVA 89], [FIN 93], [UDA 92], [HOP 87], [KEE 81], [VAN 90]
Effic11	User productivity	Effect of the DSS on user efficiency. [MAH 89], [HAM 81], [HOP 87], [UDOD 92], [UDO 92]
Effic12	Utilisation of assigned resources	The allocation of staff, machines, materials and money within the DSS project. [BAI 83], [IVA 83], [EVA 89], [CHAN 82], [HAM 81]

Table 2: Efficiency criteria (cont.)

For a DSS to be successful from an efficiency perspective it must be available when required and should be reliable. The criteria “reliability of system” measures the dependability and reliability of DSS. Synonyms that are used to measure this factor include “output reliability” and the “availability” of DSS.

The efficiency of DSS from a decision standpoint can be determined through measuring any time savings within the decision making process. “Decision speed”, “faster decision making”, “time savings” and “time taken for decision making” are terms that have been used to measure the time required for the decision process. If DSS is to be considered efficient the system should reduce, or at worst cause no change to the decision making time. This is measured through the criteria “time required for task accomplishment”.

3.1.3 Satisfaction

The satisfaction domain deals essentially with how adequate the DSS is with respect to the views of each constituency. A summary of satisfaction criteria; their definitions and references are shown in Table 3. Criteria not mentioned in the discussion below have been consistently described in other research. Where criteria have not been formally defined, definitions have been provided based on our interpretation of the criteria.

Participation of constituencies in the development of DSS is considered a useful indicator of the satisfaction they exhibit towards the system. Terms used to reflect development involvement include “top management involvement”, “ensuring user involvement” and “ensuring decision-maker involvement”.

Table 3: Satisfaction criteria

Criteria Number	**Criteria**	**Definition and reference sources**
Satis1	Accurate model construction	The model constructed for the DSS accurately details the decision. [MAH 89], [RAM 92]
Satis2	Adaptiveness	How well the system can change with respect to the decision situation. [MAH 89]
Satis3	Alternative discovery	The ability for the DSS to allow for judgements to be made based on the output of the system and to then be incorporated within the system. [MAH 89], [KEE 81]
Satis4	Attitude of constituency towards the DSS	How you feel towards the DSS. [BAI 83] , [IVE 83], [BAR 88], [ZMU 79] [BAK 90], [Udo 92], [FIN 93], [HOP 87], [UDOD 92]
Satis5	Chauffeured decision-maker participation in	Chauffeured decision-maker involvement with the DSS project during development. [BAR 88], [ZMU 79]

	development	
Satis6	Communication between groups in the DSS project	The manner and methods of information exchange between DSS project groups. [MAH 89], [BAI 83], [IVE 83], [BAR 88], [KEE 81], [UDO 92], [UDOD 92]
Satis7	Confidence in the DSS	Feeling certain that the system will perform correctly. [BAI 83], [IVE 83], [ADE 85], [HOP 87], [MAH 89]
Satis8	Convenience of access	The ease of physical access to the DSS. [BAI 83], [IVE 83],
Satis9	Critical question answering	The ability for the DSS to support questions critical to the business. [MAH 89]
Satis10	Decision complexity decrease	Affect of the DSS on the complexity of the problem from your perspective. [MAH 89] [RAM 92], [ADE 85]
Satis11	Decision-consumer participation in development	Decision-consumer involvement with the DSS project during development. [MAY 97]
Satis12	Decision-maker participation in development	Decision-maker involvement with the DSS project during development. [BAR 88], [ZMU 79]
Satis13	Decision process extension	Extension in the depth of the decision process. [MAH 89]
Satis14	Documentation	The notes provided on how the DSS operates. [BAI 83], [IVE 83]
Satis15	DSS interface applicability	How well the DSS interface matches how you work. [ADE 85], [HOP 87], [UDOD 92]
Satis16	Expectations of computer support	Your anticipated benefits of DSS support are met. [BAI 83], [IVE 83], [UDOD 92]
Satis17	Extensive retrieval functions	The access of the system to a wide range of information sources. [MAH 89]
Satis18	Learning facilities provided	The DSS enables learning about the decision task. [MAH 89]
Satis19	Management participation in development	Management involvement with the DSS project during development. [BAK 90], [BAI 83], [IVE 83], [SAN 85]
Criteria Number	**Criteria**	**Definition and reference sources**
Satis20	Perceived information quality	The quality of information from your perspective. FIN 93], [VAN 90], [ZMU 79], [IVE 83], [EVA 89], [RAM 92] , [HOP 87]
Satis21	Perceived usefulness of system	How useful you think the DSS is to your work. [ALA 92], [BAI 83] [SAN 85]
Satis22	Planning horizon broadening	The DSS helps to extend the planing horizon. [MAH 89]
Satis23	Program modification facilities	Facilities for user defined procedures or functions. [MAH 89] , [ADE 85]
Satis24	Relevance of information	Output that relates directly to the decision situation. [BAI 83], [IVE 83], [UDOD 92], [BAR 88],
Satis25	Reliance on DSS	Your dependence on the DSS for your work. [SAN 85] , [MAH 89]
Satis26	Security of data	How well the DSS is protected through security measures. [BAI 83], [HOP 87]
Satis27	Technical competence	The competence of staff in the development of the DSS. [BAI 83], [IVE 83],
Satis28	Understandability of system	How well the DSS can be comprehended. [BAI 83], [IVE 83], [BAR 88], [MAH 89], [ADE 85]
Satis29	User participation in development	User involvement with the DSS project during development. [BAI 83], [HAM 81], [ZMU 79]

Table 3: Satisfaction criteria (cont.)

From a multiple-constituency perspective, criteria used to measure the involvement of constituencies in development are “decision-consumer participation in development”, “decision-maker participation in development”, “management participation in development” and “user participation in development”.Each of these criteria measure the level of involvement of each constituency in the development of the DSS. This may affect the level of satisfaction of each constituency with a DSS.

The development and use of a DSS may provide opportunities for constituencies to learn about the decision or to discover new facets of the decision task. Constituency satisfaction with this is measured using the criteria “alternative discovery”, which is also referred to as “new insights and learning”. The focus of each of these is the measurement of whether the DSS project has enabled constituencies to learn about the decision task.

The satisfaction that constituencies have for a DSS may be affected by their attitudes towards computers, the problem and IT staff. Terms used to measure this elsewhere include “attitude of EDP staff”, “attitude of IS staff”, “attitude towards computers”, “attitude with IS.” and “attitude towards the problems addressed”. Each of these terms focuses on an aspect of the criteria “attitude of constituency towards the DSS”.

The satisfaction for a DSS may be affected by the ability of constituencies involved in the DSS project to communicate. This may occur from two major perspectives: the communication dealing with the decision task, and the communication of development issues. Terms used to describe this include “communication facilities”, “communication with IT staff” and “improved communication with IT staff”. “Communication facilities” refers to the facilities incorporated into the DSS that enable communication during the decision making process (particularly important for complex decisions that may have multiple decision-makers and users). The other two terms refer to the communication between the DSS developer constituency and the user, decision-maker, and chauffeured decision-maker constituencies. The criteria “communication between constituencies in the DSS project” is introduced here to cater for these issues.

A factor thought to effect the satisfaction of constituencies with a DSS is their confidence in the system. The terms used to refer to this factor include “user confidence” and “decision-maker confidence” which enable measurement from the perspective of only these two constituencies. The criteria used to measure the confidence that a constituency has for the DSS in the multiple-constituency approach is called “confidence in the DSS”. A positive outcome of a DSS project is thought to be the apparent reduction of decision complexity from the perspective of relevant constituencies. The terms “complexity of task” and “understandability of task” were used to measure the effect of a DSS on the complexity of the decision task. These terms are considered to be synonyms for the criteria “decision complexity decrease”.

The criteria “extensive retrieval functions” focuses on the measurement of the ability of the DSS to retrieve information for the decision task when it is required. This criteria is also referred to by the term “external database access” which measures a particular source of information. The functions provided by the system should be able to access information from a wide range of sources and present it in a number of formats.

The satisfaction of each constituency will be affected by their perceptions of the usefulness of a DSS. The criteria “perceived usefulness of the system” focuses on the measurement of how useful the DSS is for the work of a constituency. This essentially measures the utility of DSS. The “perceived utility” of the system and “DSS usefulness” have also been used as terms to describe this factor.

A useful outcome of a DSS project for some constituencies is the ability to extend their planning horizon. The satisfaction that each relevant constituency has with a DSS may be affected by this factor. “Long range planning” and “more effective strategic management” are terms that are used as synonyms for the criteria “planning horizon broadening”. These terms concern the effect that a DSS has on the time frame of strategic plans within the organisation.

An indicator of the satisfaction of constituencies with respect to a DSS may be how much they rely on the DSS to complete their work. The terms “DSS dependence” and “dependability on DSS” have been used as synonyms for the criteria “reliance on a DSS”. The focus of each of these is on measuring the dependence of relevant constituencies on DSS to complete the decision task. The criteria “understandability of system” measures how easy the system and its output are to understand from each relevant constituency’s perspective. System understanding is referred to elsewhere as the “understanding of text” which measures a subset of this criteria.

3.1.4 Use

The use domain focuses on the direct utilisation of the DSS for the tasks that were intended. Table 4 summarises criteria within the use domain, presenting definitions and references to each criterion. The criterion “ease of use”, “voluntary use” and “widespread use” are widely used in research to refer to how easy the system is to use, whether use is forced upon constituencies and how widely a DSS is used. How often the DSS is used is measured by the criterion “frequency of use”. This criterion is termed elsewhere as “repeat use” and “utilisation”. If the frequency of use is high (relative to the type of decision task) then it is likely that the system would be considered successful, although in some cases a very successful system may only be used once.

Table 4: *Use* criteria

Criteria Number	Criteria	Definition and reference sources
Use1	Ease of use	How easy you find the system to use. [ADE 85], [VAN 90]
Use2	Frequency of use	How often the DSS is used. [BAK 90], [UDOD 92], [UDO 92], [LUC 78], [FIN 93]
Use3	Voluntary use	The use of the DSS without compulsion. [BAK 90], [LUC 78]
Use4	Widespread use	How widely the DSS is used. [EIN 78], [SAN 84]

3.2 Hierarchies of criteria

It is not sufficient to list criteria within each domain from a multiple-constituency evaluation perspective, as some criteria will not be relevant to all constituencies. One means of presenting the criteria important to each constituency within a multiple-constituency evaluation process is through hierarchical structures, one for each constituency. A hierarchy is defined as a structure consisting of nodes, which occur at different levels, one below the other within a tree root like structure. At the top of the structure there is a single node. As we progress down the hierarchy the number of nodes on each level increases as higher level nodes are broken down into their components. The use of a hierarchical structure of criteria has several advantages.

First, a hierarchical structure allows criteria to be grouped within similar areas. In addition, sub-criteria can be introduced as subordinates to criteria if further depth of evaluation is required. Secondly, a hierarchical structure allows criteria and sub-criteria to be added and removed from the structure as necessary without distorting the evaluation of a DSS. Criteria that are not important for a particular system may be removed from the hierarchy and disregarded during the evaluation process. In addition, if specific criteria are identified by a constituency for a particular DSS these may be included in the hierarchy at any level. Using a hierarchy of criteria in the evaluation process will enable the evaluation of the DSS at any level. That is, a constituency may view the evaluation outcome at any level within the hierarchical structure. For example, management may view the evaluation from the perspectives of the effectiveness, efficiency, satisfaction, and use domains to determine how the DSS is performing. If any of these domains indicate that a problem exists with the DSS then management can “drill down” on that domain to determine what criteria are contributing to the problem. Finally, the presentation of the hierarchy in a graphical form enhances the evaluation process, as each constituency is able to easily visualise the relationships between criteria. This will enable each constituency to view, in as much detail as they deem necessary, how they have evaluated the system. This will in turn allow problem areas that each constituency have with the system to be quickly identified.

As an illustration the next two sections present the hierarchy of criteria relevant to the decision-makers and management groups in a DSS project. Criteria are categorised within the domains on the second level, with success forming the top level. They are further grouped, where necessary, on the third level within criteria subgroups dealing with specific areas of interest within the project. At subsequent levels the criteria and sub-criteria, where necessary, are present. These criteria are presented in the hierarchies using criteria numbers from Tables 1, 2, 3 and 4. A number of third level categories are present which are used to group criteria with a similar focus. These categories include issues dealing with individual personnel, information quality, the decision task, the system, the organisation, the IS group, time, flexibility, and costs. Not all of these may be present in each constituency’s evaluation hierarchy

3.2.1 A Hierarchy of DSS Evaluation Criteria for Decision-Makers

A hierarchy of criteria important in measuring the success of DSS from a decision-makers perspective is presented in Figure 1. The decision-maker is one of two constituencies whose job it is to analyse and carry out the decision task. As such, major criteria for the decision-maker focus on the support of a DSS for the decision task, the quality of the decision making process, and the effect of the system on the decision-maker’s ability to make decisions. In addition, as the decision-maker normally uses the system, a focus on criteria dealing with how they work is important. These are captured within the “individual personnel issues” criteria group. These criteria are common for all constituencies who directly interact with the system and its output. Thus, the overall concerns for the decision-maker constituency encompass the needs of the decision task (time, quality, and flexibility) and the ability of the DSS to support the decision task.

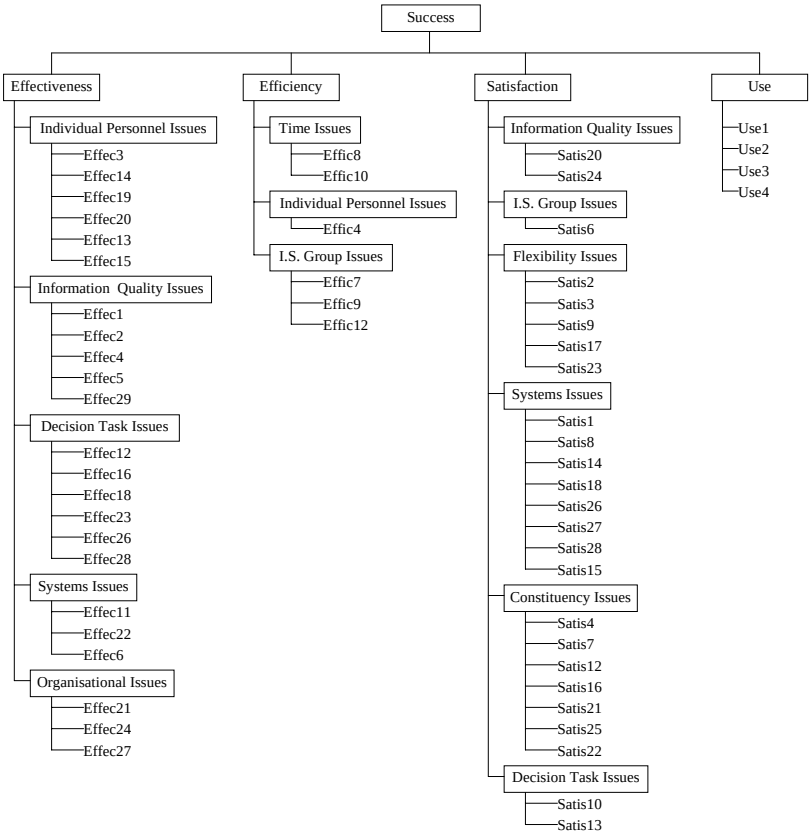

Figure 1: A hierarchy of decision-maker DSS evaluation criteria

3.2.2 A Hierarchy of DSS Evaluation Criteria for Management

The management constituency concerns deal with how the DSS has affected the business. A hierarchy of criteria important in the measurement of success from the management perspective are shown in Figure 6. The role of the management constituency in a DSS project is to determine the benefits that the system will provide to the organisation and to then support the project if the benefits are worthwhile. A focus on the productivity of constituencies, the quality of decisions, and the effect of the DSS on organisational effectiveness will enable the management constituency to rate their perspective of the success of a DSS.

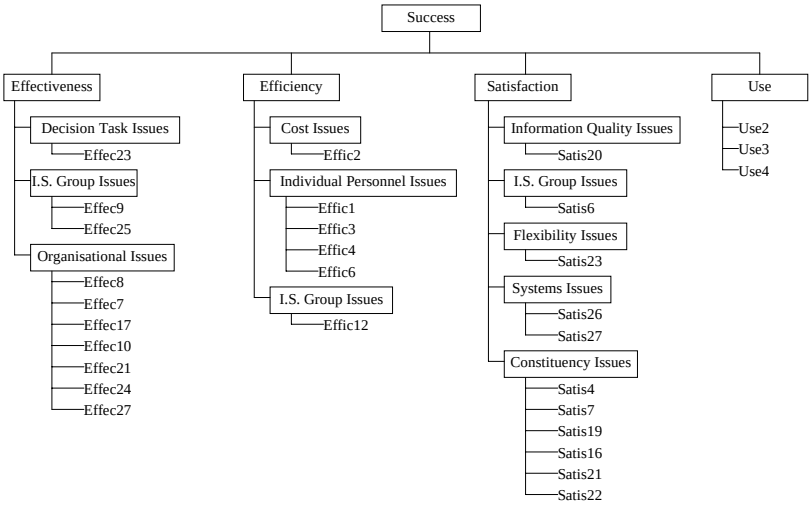

Figure 2: A hierarchy of management DSS evaluation criteria

4. The measurement of criteria

Criteria can be measured using a bi-polar attitudinal scale. For example, a constituency may be asked to rate how well the system performs with respect to the "perceived usefulness of the system". Measurement may occur along the bi-polar scale from not useful to very useful. Within the evaluation process, members of constituencies can weight the importance of each criteria to their work, and can then rate the DSS with respect to the criteria at the bottom levels of the hierarchy. These ratings and weights can be aggregated throughout the hierarchy and each node in the hierarchy will be given an evaluation "score". Scores are the success rating of the DSS based on the perspective of each constituency for that criterion [MAY 99].

Several formal methods exist for decision-making using a multiple-criteria approach [HUB 80], [BAN 90], [VIN 92]. In general these use a process of weighting and scoring each criterion based on several possible outcomes and to suggest the "best" outcome for a given situation.

Within these formal methods a number of techniques are available for the aggregation of utility. Keeney and Raffia [KER 76] describe three main utility functions: the multilinear utility function, the multiplicative utility function and the additive utility function. The most widely used method for producing an utility function is the additive utility function which is operationalised as a linear weighted average [EDW 77], [BAL 98]. The weighted linear average method is favoured for its ease of use and ease of understanding when compared to other methods. It requires the criteria to be independent by preferences, which is not a serious constraint in the case of hierarchically structured set of criteria. One of the reasons for splitting criteria into groups and levels is to achieve an unbiased and uninfluenced evaluation so that the evaluator assigns values within similar categories [MAY 99].

The use of a hierarchical structure of criteria has several advantages: In general, a hierarchical structure is used to either select the best alternative in decision-making approaches, or to aid the decision-maker to select an alternative in decision aiding approaches [SAA 90].

- A hierarchical structure allows criteria to be grouped in similar areas. In addition, sub-criteria can be introduced as subordinates to criteria if further depth of evaluation is required. Finlay [FIN 94] states that a hierarchy '… enables the user to disaggregate highly complex and generic criteria into their measurable components. Expert judgement and existing data are likely to be more effectively incorporated when using the more concrete, lower-level criteria'
- A hierarchy allows criteria and sub-criteria to be added and removed without distorting the evaluation of a DSS. Criteria that are not important for a particular system may be removed from the hierarchy and disregarded during the evaluation process. In addition, criteria specific to the system evaluated may be included in the hierarchy at any level without adversely affecting the evaluation.
- Using a hierarchy of criteria in the evaluation process will enable evaluation of the DSS at any level. A constituency may view the evaluation outcome at any level in the hierarchical structure, and can 'drill down' to lower levels of the hierarchy at will.
- Presenting the hierarchy in a graphic form enhances the evaluation process as each constituency can easily visualise the relationships between criteria and thus view, in as much detail as necessary, how they have evaluated the system. This will, in turn, allow problem areas in the system to be identified.

From an evaluation perspective, Adelman et al. [ADE 85] suggest a multiple-critera method for evaluating DSS. In this method, a hierarchy of attributes (shown as nodes in Figure 1) is created with each attribute having the same weight. The method scores each of the bottom level attributes (shown in bold in Figure 1), then for each hierarchy branch, the scores of each lower level attribute are averaged to give a score for that parent attribute. This progresses through the entire hierarchy until each attribute is scored. Others have used this method in DSS evaluation ([ADE 89], [HOP89], [AND 89], [ADE 89], [GOI 92]). There are, however, several caveats to this approach. The comparison of attributes requires that a common scale must be used otherwise comparisons are meaningless. Adelman [ADE 89] states that it is possible to weight the attributes and then use the weights to determine the utility of attributes throughout the hierarchy – but in practice he ignores weights because, it is "inappropriate at the moment to weight criteria" .

Goicoechea, Stakhiv and Li [GOI 92] built on Adelman et al.'s [ADE 85] approach by averaging the responses for questions relating to a criterion and then using weighted averages to progress up the hierarchy. This method is distinct as the weights of each family (see Figure 1) are used to determine the score for the parent node. Yet, the method still uses averages to obtain the score of the bottom level nodes. This may be detrimental as no importance is assigned to the bottom level criteria.

The various advantages and disadvantages of these methods have influenced the selection of an appropriate method for weighing and scoring criteria in the DSS evaluation process. Saaty [SAA94] criticises the approach Huber [HUB 80] and others used because of differences in scale. Problems can occur when different sets of numbers are used to scale the judgements for the alternatives under different criteria. When the numbers are normalised, all sets would lie in the interval [0-1] no matter from which scale set they originated. Thus, values of scale are lost with this approach. This argument, however, becomes irrelevant in attitudinal research where an attitude towards an object (in our case the DSS) is being measured using the same scale for each criterion.

The calculation of weights for attributes in Huber's approach allows the calculation of meaningful scores for the parent nodes in the hierarchy, as weights of the whole tree are applied at the bottom level. It is then a simple case of adding the weighted scores up the hierarchy to get scores for successive parents. Problems occur when there are differing numbers of levels for each branch in the hierarchy. For each successive level of the tree, if weights are between 0 and 1, a factor of about one tenth is applied to the weights in that level. This may not be appropriate, as weights at lower levels would have less impact on the resulting score. Thus, for unevenly levelled trees, which are likely in DSS evaluation, the method Huber discusses becomes less practical.

The normalisation of weights in a hierarchy is usually completed through two main approaches. Weights on any one level of the hierarchy are normalised so that they total to 1 (or 100%), or weights in a family are normalised so that they total to 1. These approaches work effectively when each branch of the hierarchy has the same depth. However, once varying depths are encountered the second approach tends to bias the weighting procedure so that weights at the bottom level have less meaning (as in Huber's approach). Depending on the purpose of the hierarchy this may be suitable, but for an evaluation procedure where differing branch depth may be the norm, this is not the case.

In AHP the process of calculating priority vectors and conducting pair-wise comparisons for a large number of attributes becomes time consuming when compared to other approaches. For instance, if N equals the number of criteria, then the number of comparisons to be conducted equals $\frac{\left(N^2 - N\right)}{2}$. For a large N the number of comparisons becomes too impractical to deal with. This compares poorly with other approaches where the number of questions required equals the number of criteria. This may dramatically increase the intrusiveness and time required for the DSS evaluation process possibly resulting in the evaluation process not being conducted.

The approaches based on weighted averages assume that the criteria are linearly related and then use addition to obtain the score for the upper levels of the hierarchy. This assumption can only be made when additivity conditions are met [KEE 76. This, however, is generally a reasonable assumption. Edwards [EDW 77] states "quite substantial deviations from value independence will make little difference to the ultimate utility and even less to the ranked order of weights of criteria".

Additive linear approaches have some advantages for this context, as they are simple to use, are time efficient, and are easy to explain to those evaluating the DSS. This type of approach, defined by Goicoechea et al. [GOI 92] and Adelman et al. [ADE 89], requires further research before it can be used directly for an operational evaluation. For instance, these approaches neglect to specify how weights were obtained throughout the hierarchy and whether they have been normalised, and if so, how. From the review presented above, it is clear that for the process of DSS evaluation no single existing approach is adequate.

5. A Multiple-constituency approach to evaluation of DSS based on multiple-criteria hierarchy

This section describes an approach for multifaceted DSS evaluation. The method proposed for evaluating DSS in a multiple-constituency environment, in respect of a hierarchical structure of multiple-criteria is based on approaches described above. The approach we adopted is based on combination of the approaches reviewed in the previous section and addresses most of the identified problems.

Our approach to weighting and scoring each criterion assumes that the person conducting the evaluation builds a hierarchy of criteria. A generic hierarchy exists as a starting point for the evaluator who may delete or add to the hierarchy any specific criterion that they feel is relevant. The approach then uses these criteria to evaluate the success of a DSS. Rather than producing a single outcome as a measure of the success of a DSS for all stakeholders, the proposed method produces an outcome for each relevant criteria structure for each of the stakeholders. This produces a more comprehensive, cohesive evaluation. The steps for accomplishing this are outlined below.

Step 1: Obtain the weights of all nodes from the evaluator

For each of the nodes in the hierarchy, the person evaluating the DSS indicates how important each criterion is to them. This can be done on a purely subjective basis on a sliding scale with responses later transformed to numerical values [SAA90]. In the proposed approach, the person evaluating the DSS is not concerned with the numerical values. Evaluators only see a bi-polar scale that is labelled at each end. Figure 3 shows an example of such a scale. The method weights each node of the hierarchy on a scale of 1 to 100 with 0 indicating that the criterion is not applicable in this instance.

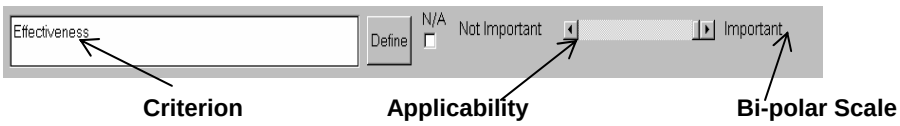

Figure 3: Scale for weighting criteria

Step 2: Convert the weights into normalised weights

For each element of a level determine the normalised weight working from the bottom to the top of the hierarchy. This will convert all the weights of each level of a tree to be in the range 0 to 1 and will mean that the sum of the weights in the hierarchy will be 1.

Step 3: Obtain the scores for each bottom level criterion from the evaluator

For each of the bottom level elements get the evaluator to indicate how well the system matches the criteria. These answers will be subjective and will be represented in a similar manner to the weighting of the criteria. The evaluator will rate the system on a scale from the top to the bottom values and the position on the sliding bar will be converted to numerical values. As in step 1, the evaluator is unaware of the numerical scale used and as long as the scale is used consistently throughout the evaluation, it can take on any reasonable range of values. A sample scale is shown in Figure 4.

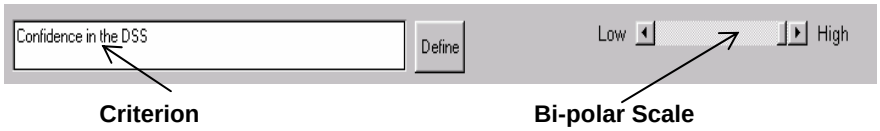

Figure 4: Scale for scoring criteria

Step 4 Use bottom level weights to convert the score to a weighted score

Multiply each of the scores from Step 3 by the normalised weights to get scores for each criterion at the lowest level for each hierarchy branch.

Step 5: Use the weighted scores to calculate scores for the hierarchy

The weighted score for the object is derived as a product of a normalised weight of the object and a sum of the weighted scores of its children (the dependant objects from the level immediately below in the hierarchy).

The method of calculating the weights and scores assumes that a hierarchy of evaluation factors has been created. A value independent relationship between the criteria should be apparent so that an linear weighted sum method can be used. But, as [EDW 77] points out, this is not critical for the success of the method, as little difference will be made to the utility of attributes in the hierarchy.

6. An architecture for a DSS evaluation tool

This section describes a system architecture for a tool based upon the multi-faceted DSS evaluation theory described in the previous section. There are two overriding concerns with the functionality of the multiple-consistency DSS evaluation tool. These are the adequacy of the tool in catering for each different constituency and the focus on the evaluation criteria, process, and outcomes. The architecture should:

- support the multiple-constituency DSS evaluation framework,
- allow generic and specific evaluation criteria,
- display and store the results of the evaluation process,
- provide a secure environment in which to conduct the evaluation,
- anable the seamless integration of all tasks.

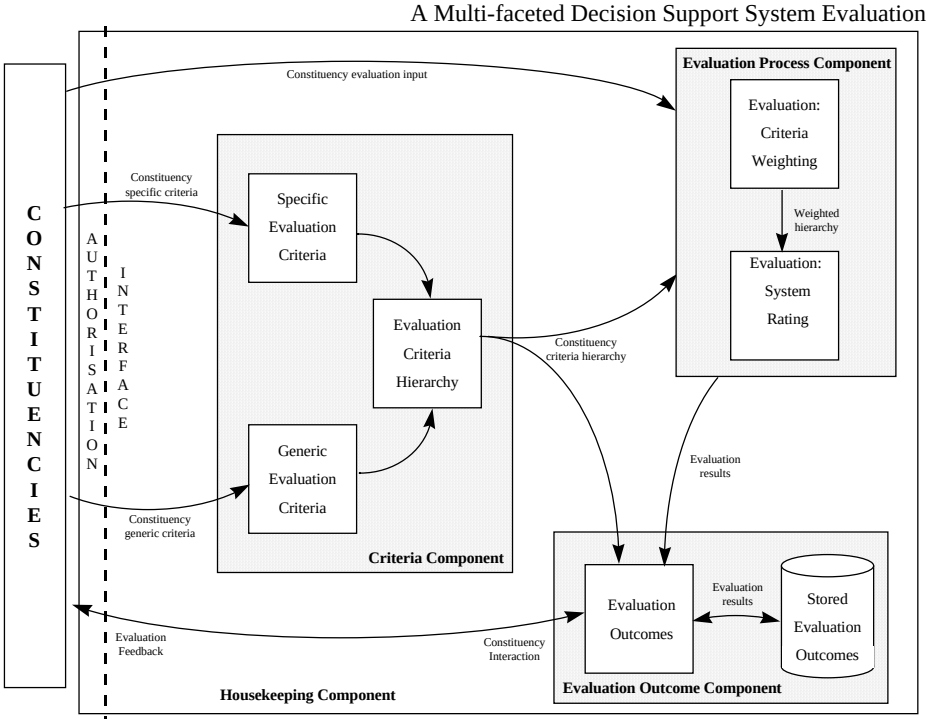

Figure 5. A Multiple Constituency DSS Evaluation Architecture

The architecture, shown in Figure 5 consists of four major components: housekeeping; criteria database; evaluation process, and the evaluation outcome.

6.1 The Housekeeping Component

The housekeeping component is responsible for integrating the evaluation process, evaluation outcome, and criteria database components. It monitors the required functionality for all components of the architecture and the interaction between them.

The success of the evaluation process may depend on the ability of the interface to provide the DSS evaluator with an appropriate information about the tool and its underlying architecture. An important function, which the interface provides, is access control. The authorisation dialogue identifies the person who wishes to use the evaluation tool. It determines if this person has authorisation to use the tool and ascertains the perspectives from which they are allowed to use it. This limits use of the tool to those constituencies in which the person evaluating the DSS is a member of and provides security for the evaluation outcomes.

The housekeeping component provides management of several other important aspects of the architecture. Management of the specific criteria in the criteria database component is an integral part of the architecture. The housekeeping component allows these criteria to be added to, modified, and deleted as required by each constituency. Likewise, management of evaluation outcomes is required. The housekeeping component provides a way to add or retrieve an evaluation outcome. Also, it allows deletion of all outcomes.

The housekeeping component also handles the linkages between other architecture components and the constituencies. Partially, this is accomplished by the interface between the constituencies and the underlying architecture of the evaluation tool. It is also accomplished by moving information through the architecture. For instance (as shown in Figure 5), a constituency may add specific criteria in the criteria database component and these are placed into that constituency's evaluation criteria hierarchy. The housekeeping component passes the criteria hierarchy to the evaluation process component and provides an interface to enable the constituency to evaluate the DSS based on their hierarchy of criteria. Housekeeping then passes the results of the evaluation to the evaluation outcome component. They are then shown through the interface and stored accordingly.

6.2 The Criteria Database Component

The criteria database component of the architecture supports generation of the generic and specific criteria and provides a mechanism for producing the hierarchy of criteria, both generic and specific for each constituency (see Figure 5). While generating all of the criteria required for the evaluation process by each constituency would be useful, the time required for such an exercise precludes its practical application. As such, it is necessary to use a generalised set of criteria complimented by a set of criteria specific to an instance of the evaluation. The architecture allows criteria specific to the current evaluation to be elicited. The specific criteria area of the criteria database component captures these criteria.

The final part of the criteria database component generates a hierarchy of criteria for the evaluation to take place. Specific criteria defined by the administrator, combined with the hierarchy of criteria specified for each constituency , form a comprehensive hierarchy of criteria for the evaluation of the selected DSS.

6.3 The Evaluation Process Component

The evaluation process component comprises two major parts: measuring weights for criteria and scoring criteria for the specific DSS. It uses the hierarchy of criteria provided by the criteria database component and requires input from the evaluator and implements the process described above in Section 4.

6.4 The Evaluation Outcome Component

The purpose of the evaluation outcome component is to present the results of the evaluation process to each constituency. This can show the results of a current evaluation or those of previous evaluations. One major function that the evaluation outcome component must provide is flexibility. The understanding of each constituency may be improved if the evaluation results can be viewed at various levels of detail. The component focuses on the delivering of evaluation outcomes to the person evaluating the system in a manner that compliments the hierarchical criteria structure and allows them to view the hierarchy in a meaningful way, at any level. It uses the results of the evaluation process, the criteria hierarchy for the constituency carrying out the evaluation and requires some interaction with this constituency.

In addition, the evaluation outcome component must allow the outcome of the evaluation to be stored for future use. The comparison between two evaluations of the same system may prove valuable to each constituency, as it may point out where the system has improved or degraded over time.

7. MultiVal - a tool for DSS evaluation

The evaluation architecture described in Section 6 has been implemented in a prototype DSS evaluation system called MultiVal. The first prototype was implemented using Microsoft Excel™. Further development, using Java, is currently taking place (see prototype of MultiVal'2001 at http://km-svr.sims.monash.edu.au/ MultiVal2001/).

Whilst there are many products available that focus on multiple-criteria decision-making (see HIVIEW, VISA and Resolve*Ballot for example), there seems that focus on DSS evaluation. The aim of multiple-criteria decision-making packages is on comparing alternate strategies or options in solving a particular problem. Rather than focusing on alternates MultiVal seeks to identify the factors that are crucial for the success of a particular artefact, based on multiple perspectives of multiple criteria. In essence MultiVal extends the concepts of these other products and changes the focus from a decision-making perspective, to that of evaluating an artefact. It allows criteria to be added, removed or to be marked as inactive where necessary. It is important to allow for different groups to have different criteria.

MultiVal is based on the architecture described above. As such, there are four major components: housekeeping, criteria database, evaluation process, and evaluation outcome. Figure 6 illustrates the

structure of the evaluation tool. Each of the steps is numbered and these numbers are referred to in the discussion.

All components of the tool accept for the Administration component, can be used by any type of users as long as they are taking part in the DSS project and have a particular role in it. The Administration Component of the MultiVal performs three distinct functions: managing specific criteria, managing people using the tool and the evaluation outcome database. Specific criteria options include adding, modifying and deleting specific criteria and deleting all specific criteria. Adding, deleting and modifying an evaluator's personal details, and modifying evaluator privileges are contained in the evaluator's area. The evaluation outcome database is also linked to this component.

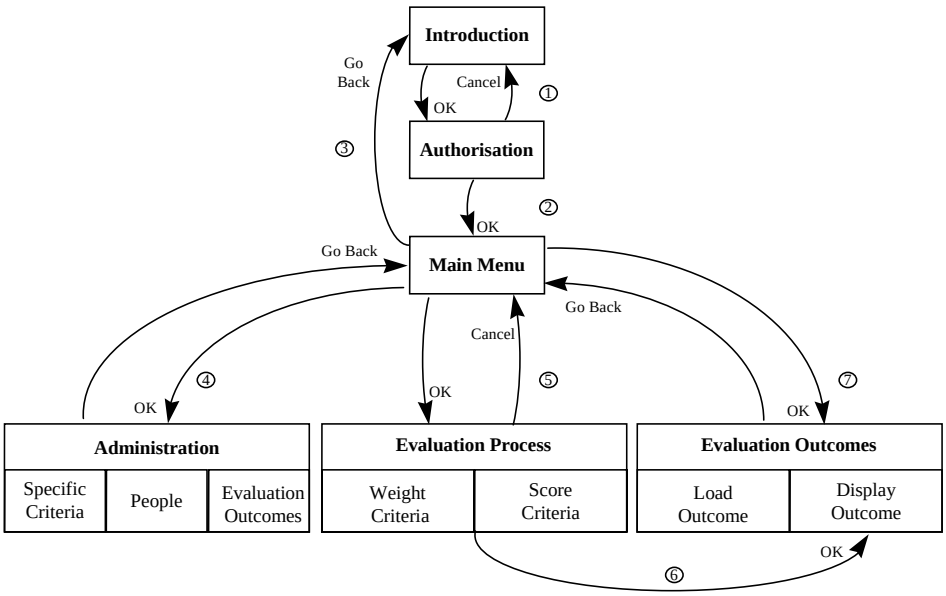

Figure 5: **The Structure of MultiVal.**

When MultiVal is started, the introduction dialogue is shown (see Figure 6). This explains the purpose of the evaluation aid and gives a brief overview of the system.

When the person evaluating the DSS presses OK ① from the Introduction Dialogue, the authorisation dialogue is invoked (Figure 7). The role the person using the aid wishes to play is selected. In Figure 7 the person evaluating the DSS has selected the User constituency role.

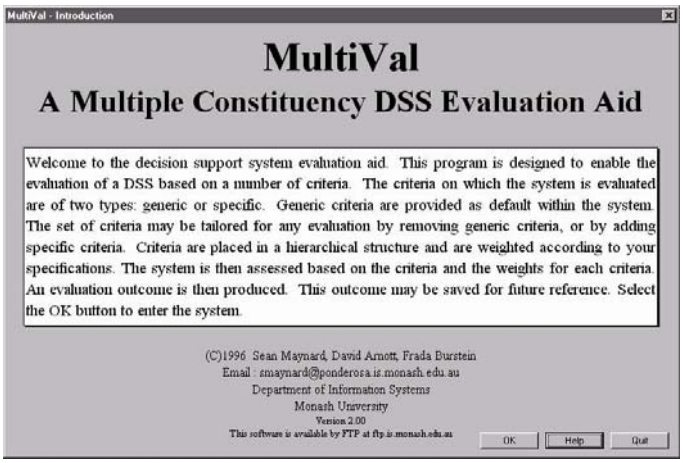

Figure 6: Introduction Screen

In the evaluation that follows, the criteria hierarchies the tool produces will be for the user constituency alone. If the person using the evaluation aid selects a constituency role in which he or she does not belong, or has not been set up to use, then access to the tool will be denied.

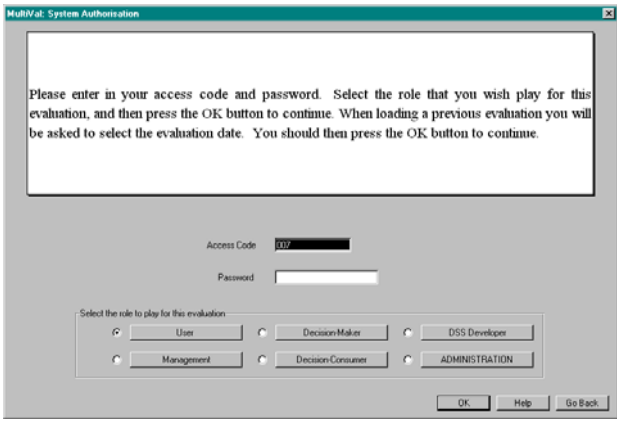

Figure 7: **Authorisation Dialogue Screen**

The user of the evaluation aid can see the definitions of each constituency by clicking on the relevant button. Once the user has completed the authorisation dialogue, they gain access to administration ④, evaluation ⑤ or outcomes ⑦.

Selecting the Evaluate the DSS ⑤ option initiates the evaluation process. The tool retrieves the generic criteria and the specific criteria for the selected constituency role and places them in the criteria hierarchy. MultiVal then displays a dialogue (Figure 8), which asks the evaluee to rate how important each of the criteria is to them (without a reference to the particular system in question).

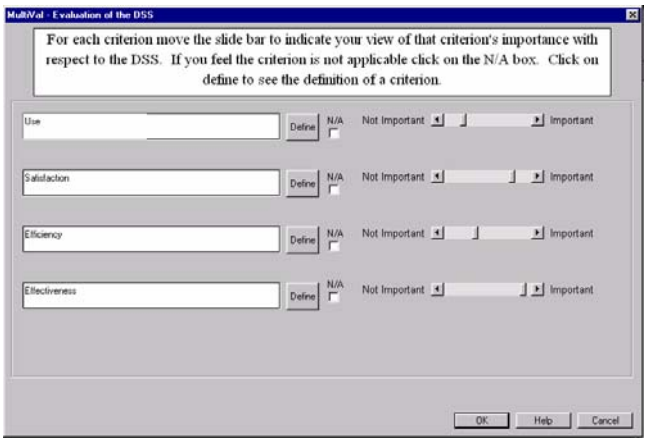

Figure 8: Criteria Weighting Screen

MultiVal then displays the second level of the hierarchy and asks the person evaluating the DSS to indicate how important each of the criteria on this level is to their work. The evaluation process assumes that the lop level of the hierarchy ie. *success* is weighted with a value of 1. MultiVal then progressively asks for the same information about the criteria at each level of the hierarchy. At any level in the hierarchy the person evaluating the DSS may indicate that any particular criterion is not applicable in (relevant to) the current situation. In such an instance, all criteria below this criterion in the hierarchy are also considered inapplicable and therefore ignored in the evaluation process. The definition of each criterion can be accessed by pressing the define button immediately to the right of the criterion's name. The criteria are assigned weights reflecting the position of the indicator in the slide bar along the bi-polar axis.

Table 5 shows a sample implementation of the generic criteria structure available for all constituencies in the MultiVal tool while Table 6 shows an implementation of the specific criteria structure.

Table 5: Implementation of the Generic Criteria Structure

Level 0	Level 1	Level 2	Level 3	Definition	User	DM	DSS	MGN	DC	Adjective Pairs (Weighting)	Adjective Pairs (Scoring)
Success				The attainment of DSS project objectives.	YES	YES	YES	YES	YES	Not Important/ Important	
	Satisfaction			The adequacy of the DSS.	YES	YES	YES	YES	YES	Not Important/ Important	
		Systems satisfaction issues		Issues dealing with the satisfaction that you have with the DSS	YES	YES	YES	YES	YES	Not Important/ Important	
			Accurate model construction	Whether the model constructed for the DSS accurately details the decision.		YES	YES			Not Important/ Important	Not Accurate/ Accurate
			Convenience of access	The ease of physical access to the DSS.	YES					Not Important/ Important	Low/High
			Documentation	The notes provided on how the DSS operates.	YES		YES			Not Important/ Important	Bad/Good
			DSS interface applicability	How well the DSS interface matches how you work.	YES					Not Important/ Important	Not Adequate/ Adequate
			Learning facilities provided	Whether the DSS enables learning about the decision task.	YES		YES			Not Important/ Important	Not Adequate/ Adequate
			Security of data	The adequacy of the security provided in the DSS.	YES	YES	YES	YES	YES	Not Important/ Important	Not Adequate/ Adequate
			Technical competence	The competence of staff in the development of the DSS.	YES	YES	YES	YES		Not Important/ Important	Low/High
			Understandability of system	How easy the DSS is to comprehend.	YES					Not Important/ Important	Low/High

Once the relevant criteria in the hierarchy are weighted, the bottom level criteria are presented to the evaluator so that they can rate the system with respect to each criterion (Figure 9).

Table 6 Implementation of the Specific Criteria Structure

Level 0	Level 1	Level 2	Level 3	Definition	User	DM	DSSD	MGNT	DC	Adjective Pairs (Weighting)	Adjective Pairs (Scoring)
Success0					YES	YES	YES	YES	YES	Not Important / Important	
	Satisfaction				YES	YES	YES	YES	YES	Not Important / Important	
		Constituency satisfaction issues			YES	YES	YES	YES	YES	Not Important / Important	
			Suitability of Interface for DSS Support	How suitable the DSS interface is for the individual involved	YES		YES			Not Important / Important	Low/High

The scoring of each criterion is completed through positioning an indicator on a bi-polar axis. Unlike the weighting of criteria, the rating of criteria does not allow a criterion to be marked as inapplicable.

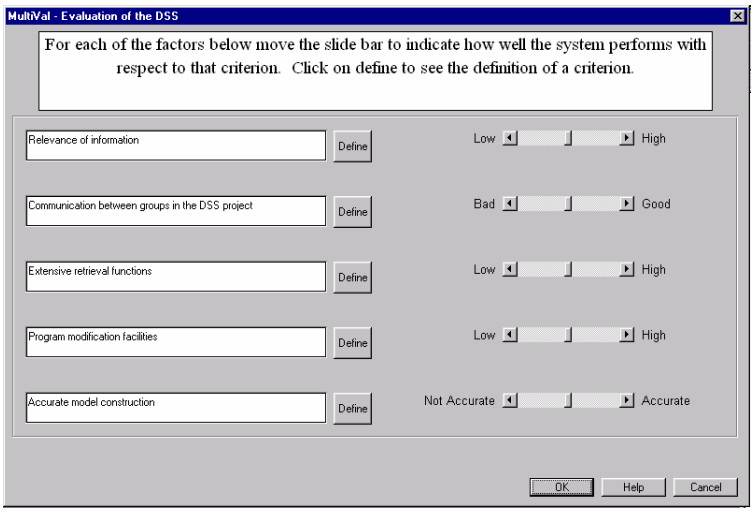

Figure 9: The criteria scoring screen

Once the criteria are rated, scores are computed and the outcome is shown ⑥ (Figure 11). This completes the evaluation process described in Section 5. If at any stage during the evaluation process the cancel button is selected the control will be returned to the main menu and the partial evaluation will be discarded.

The person using the evaluation aid can also select *View Evaluation Outcomes* ⑦ from the main menu in order to access past evaluation results. To clarify which one to retreave they will be asked to select the particular evaluation to view (Figure 10). The system determines the valid evaluations that this particular person may view based on the access code and on the constituency role selected in the authorisation screen.

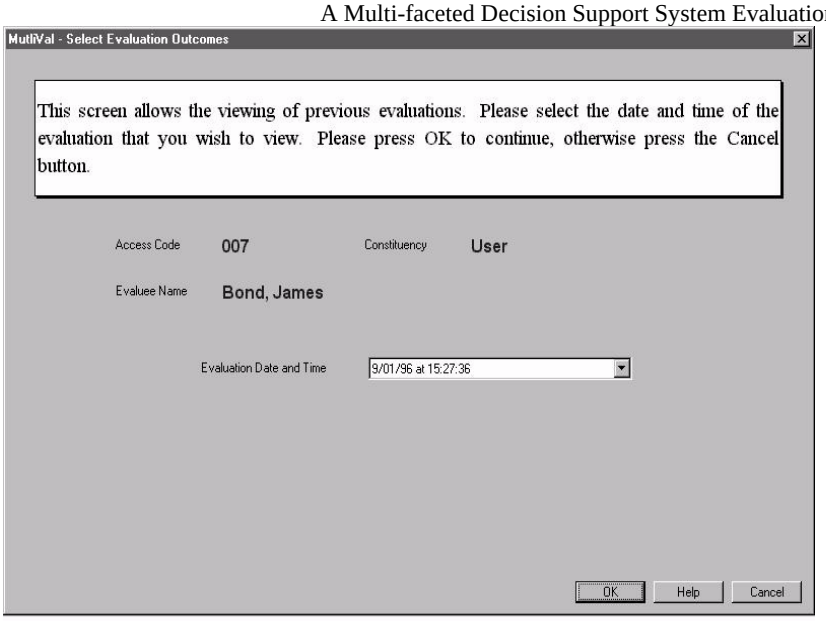

Figure 10: Select Evaluation Outcome Screen

The person using the evaluation aid selects the date and time of the evaluation that they wish to view. This will display the evaluation results for the evaluation undertaken by that person, in the constituency selected on the date and time selected (Figure 11).

The top of the screen in the Figure 11 displays the current evaluation constituency along with the date and time of the evaluation. Directly below this on the left-hand side of the dialogue the parent level criterion is shown (in this case it is Success). Next to this two buttons appear that allow the user of the tool to select whether to

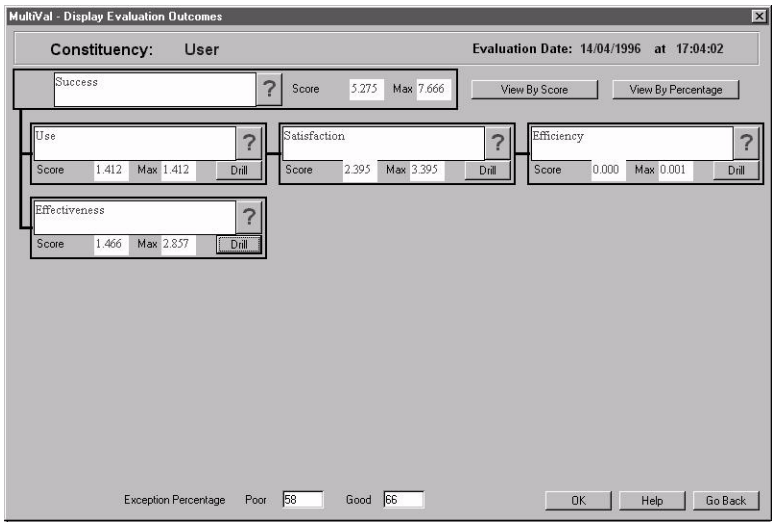

Figure 11: Evaluation Outcomes Screen

view the evaluation results as a score (as shown) or as a percentage. Below this, each of the applicable criteria in the hierarchy in the level below the parent level criterion is shown. For each criterion shown in this dialogue a number of areas are visible (see Figure 12).

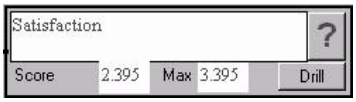

Figure 12. Criterion Outcome Box

The name of the criterion is displayed in the main part of the criteria box. Directly under the name, the score and maximum value is shown. In this same area the percent would be shown if the evaluator was viewing it from that perspective. Next to the scores is the drill button that allows the criteria at the next level down in the hierarchy to be shown. The final area in the outcome box is the '?' button that allows information to be retrieved about the meaning of this particular criterion.

In the evaluation outcomes screen the exception percentage for poor and good results is shown in the lower third of the screen. The user can place limits on what is considered to be good or poor evaluation results. These figures are used to determine the colour to display each criterion. For instance, if the person evaluating the system thought that any criterion rating below 50% was poor and any criterion rating above 90% was good, criteria receiving a rating of less than 50% would be displayed in red, all values between 50% and 90% displayed in black, and all receiving a rating greater than 90% would be displayed in green. This is similar to the drill down navigation and exception reporting in EIS systems.

8. Concluding comments

In order to be effective the evaluation process should measure the success of the system from the perspective of all stakeholders in a DSS project. In this paper we presented a method of DSS evaluation based on research from a number of fields that allows the consideration of all interest groups' opinion in a DSS project. The approach allows the people evaluating the system give weights to the criteria based on their perception of the importance of those criteria for the particular DSS case. This method of DSS evaluation helps to reveal differences in perceptions about the project and its outcomes, as well as capture and reflect the context of the evaluation. Such an approach provides a dynamic environment to monitor changes that need to be made to achieve overall project success. We then presented a system architecture based on the evaluation method.

The MultiVal tool shows how this architecture can be operationalised. Using the tool allows the evaluation to be preformed quickly and efficiently at any point of DSS development and operation. This is achieved by using generic evaluation hierarchy as a template for each evaluation. It provides a rigorous basis for improvements to the DSS before the project is complete. It can also be used to provide feedback to stakeholders after the system is complete and is operational.

The next stage of the project is to test MultiVal in the field. This stage will use an action research approach whwere the tool is used in real project evaluation, its performance analyses, changes made, and the application/evaluation cycle repeated.

References:

[ADE 85] Adelman, L.; Rook, F.W. and Lehner, P.E. 'User and R&D Specialist Evaluation of Decision Support Systems'. IEEE Transactions on Systems, Man, and Cybernetics, 15:3, 334-342. (May-Jun 1985).

[ADE 86] Adelman, L. and Donnell, M.L. 'Evaluating Decision Support Systems: A General Framework and Case Study' in Microcomputer Decision Support Systems, (Ed.) Andriole, S. J., Wellesley, MA. QED Inform. Sci. Ch. 12., 1986.

[ADE 89] Adelman, L. 'Integrating Evaluation Methods into the DSS Development'. Information and Decision Technologies, 15:4, 1989, 227-241.

[ALA 92] Alavi, M. and Joachimsthaler, E.A. 'Revisiting DSS Implementation Research: A Meta Analysis of the Literature and Suggestions for Researchers'. MIS Quarterly, 16:1, 95-116. (March 1992).

[AMA 93] Amoako-Gyampah, K. and White, K.B. 'User Involvement and User Satisfaction'. Information & Management, 1993. 25, 1-10.

[AND 89] Andriole, S.J. Handbook for the Design, Evaluation, and Application of Interactive Military Decision Support Systems. Prinston, NJ. Pertrocelli Books Inc., 1989.

[BAI 83] Bailey, J.E. and Pearson, S.W. 'Development of a Tool for Measuring and Analysing Computer User Satisfaction'. Management Science, 29:5, 530-544. (May 1983).

[BAK 90] Barki, H. and Huff, S.L. (). 'Implementing Decision Support Systems: Correlates of User Satisfaction and System Usage'. INFOR (Canada), 28:2, 89-101. May 1990.

[BAL 98] Ballestero, E. and Romero C, Multiple criteria decision making and its applications to economic problems, Boston, Mass.: Kluwer Academic Publishers, 1998.

[BAN 90] Bana e'Costa C. (ed.). Readings in multiple criteria decision aid, Berlin ; New York : Springer-Verlag, 1990.

[BAR 88] Baroudi, J.J. and Orlikowski, W.J. 'A Short-Form Measure of User Information Satisfaction: A Psychometric Evaluation and Notes on Use'. Journal of Management Information Systems, 4:4, 44-59. (spring 1988).

[CHAN 82] Chandler, J.S. 'A Multiple Criteria Approach for Evaluating Information Systems'. MIS Quarterly, 61-74. (March 1982).

[EDW 77] Edwards, W. 'Use of Multi-attribute Utility Measurement for Social Decision Making' in Bell, D.E., Keeney, R.L. and Raiffa, H. (1977) Conflicting Objectives in Decisions, John Wiley and Sons, Chichester, 1977.

[EIN 78] Ein-Dor, P. and Segev, E. 'Organisational Context and the Success of Management Information Systems'. Management Science, 24:10, 1064-1077. (June1978).

[EVA 89] Evans, G.E. and Riha, J.R. 'Assessing DSS Effectiveness Using Evaluation Research Methods'. Information & Management, 16:4, 197-206. (April 1989).

[FIN 93] Finlay, P.N. 'Measurement of Success for Lone User MIS'. Journal of Management Information Systems, 3, 1993 47-67.

[GOI 92] Goicoechea, A.; Stakhiv, E.Z.; and Li, F. 'A Framework for Qualitative Experimental Evaluation of Multiple Criteria Decision Support Systems' in Proceedings of the Ninth International Conference on Multiple Criteria Decision Making: Theory and Application in Business, Industry and Government, (Ed.) Goicoechea, A.; Zionts, S.; Duckstein, L., New York. Springer-Verlag, 1992, 1-17.

[HAM 81] Hamilton, J.S. and Chervany, N.L. 'Evaluating Information Systems Effectiveness'. MIS Quarterly, 5:3, 55-59. (September1981).

[HIG2001] HIGHVIEW 'HIGHVIEW for Windows', http://www.enterprise-lse.co.uk/hiview.htm. Accessed 22/03/2001.

[HOP 87] Hopple, G.W. 'Decision Support Systems: Software Evaluation Criteria and Methodologies'. Large Scale Systems, 12:3, 1987, 285-300.

[HUB 80] Huber, G.P. Managerial Decision Making. Scott, Foresman and Company. Glenview, Illinois, 1980.

[IVE 83] Ives, B.; Olson, M. and Baroudi, J., 'The Measurement of User Information Satisfaction'. Communications of the ACM, 1983 26:10, 785-793.

[KEE 81] Keen, P.G.W. 'Value Analysis - Justifying Decision Support Systems'. MIS Quarterly, 5:1, 1-14. (March1981).

[KER 76] Huber, G.P. Managerial Decision Making. Scott, Foresman and Company. Glenview, Illinois, 1980.

[LUC 78] Lucas, H.C. 'Unsuccessful Implementation: The Case of a Computerised Order Entry System'. Decision Sciences, 9, 68-79, 1978.

[MAH 89] Mahmood, M.A. and Sniezek, J.A. 'Defining Decision Support Systems: An Empirical Assessment of End-User Satisfaction'. INFOR (Canada), 27:3, 253-271. Aug 1989.

[MAY 94] Maynard, S. and Arnott D.R. 'A Multiple Constituency Approach to Decision Support Systems Evaluation', Proceedings of the Fifth Australasian Conference on Information Systems. 1994.

[MAY 95] Maynard, S., Arnott, D.R., and Burstein, F. 'DSS Evaluation Criteria: A Multiple Constituency Approach', Proceedings of the 6th Australasian Conference on Information Systems. 1995.

[MAY 97] Maynard, S. 'A Multiple Constituency Approach for the Evaluation of Decision Support Systems', Masters Thesis, Department of Information Systems, Monash University, 1997.

[MAY 99] Maynard, S., Arnott, D.R., and Burstein, F. "A Method For Multiple Criteria Evaluation of DSS In A Multiple-Constituency Environment", (ed) F.Burstein, Proceedings of the Fifth International Conference of the International Society for Decision Support Systems, Melbourne, 1999 (CD ROM).

[RAM 92] Ramamurthy, K.; King, W.R. and Premkumar, G. 'User Characteristics - DSS Effectiveness Linkage: An Empirical Assessment'. International Journal of Man Machine Studies, 36:3, 469-505. (March 1992).

[RES2001] RESOLVE*Ballot 'The Resolver Ballot Features', http://www.barrettsaunders.com/Products/ prodBallot_Key.htm. Accessed 22/03/2001

[SAA 80] Saaty, T.L. The Analytic Hierarchy Process, McGraw-Hill, New York, 1980.

[SAA 90] Saaty, T.L. 'How to make a decision: The Analytic Hierarchy Process'. European Journal of Operational Research, 9:26, 1990, 9-26.

[SAA 94] Saaty, T.L. 'Highlights and Critical Points in the Theory and Application of the Analytic Hierarchy Process', European Journal of Operational Research, 74, 1994, 426-447.

[SAN 84] Sanders, G. 'MIS/DSS Success, Measures'. Systems, Objectives, Solutions, 4, 1984, 29-34.

[SAN 85] Sanders, G.L. and Courtney, J.F. 'A Field Study of Organisational Factors Influencing DSS Success'. MIS Quarterly, 9:1, 77-93. (March 1985).

[UDOD 92] Udo, G.J. and Davis, J.S. 'Factors Affecting DSS Benefits'. Information and Management, 23:6, 359-371. (December 1992).

[VAN 90] Van Tran, H. 'Successful DSS Development with Traditional Tools and Techniques'. Journal of Information Systems Management, 7:3, 46-55. (Summer 1990).

[VIN 92] Vincke, P. Multicriteria decision-aid. Chichester [England] New York : Wiley, 1992.

[VISA 2001] 'V.I.S.A Multiple Criteria Decision Analysis' http://www.simul8.com/visa.htm. Accessed 22/03/2001.

[ZMU 79] Zmud, R.W. 'Individual Differences and MIS Success: A Review of the Empirical Literature.'. Management Science, 25, 1979, 966-979.