A Method For Multiple Criteria Evaluation Of DSS In A Multiple-Constituency Environment

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

David Arnott, Frada Burstein
School of Information Management and Systems
Monash University, Melbourne, Australia
Email   david.arnott@sims.monash.edu.au
frada.burstein@sims.monash.edu.au

Abstract

Many studies discuss approaches for the evaluation of DSS. Few, however, provide a method to allow the context of the evaluation to be explicitly considered. This paper proposes a DSS evaluation method based on multiple criteria and incorporating a multiple constituency perspective. The approach proposed can be used for any group - constituencies, involved in DSS project at any stage. With multiple-constituency DSS evaluation many criteria may, at any time, be valuable to a constituency group, whilst others may be unimportant, or inconsequential. This paper introduces the concept of multiple-constituencies and discusses some methods for multiple criteria decision making and evaluation. A prototype tool implementing the approach is also described.

1. Introduction

The evaluation of a decision support system (DSS) is usually based on the opinions of a single reference group, or constituency, rather than on all relevant groups involved with the DSS project. Consequently, many methods of evaluation focus on a single group, usually the users or decision-makers, depending on the terminology used. Including each relevant constituency in the evaluation process may produce a more context-sensitive evaluation that could be used to improve the DSS from each group's perspective (Maynard, Arnott and Burstein, 1995). An example of this type of approach is the multiple-constituency approach (Connolly, Conlon and Deutsch 1980).

The evaluation process must measure the success of the system from each constituency's perspective to be considered useful. To accomplish this, a number of evaluation criteria will be required, with many being valid criteria for one, or many constituency groups. Any method used in such an evaluation environment must adequately utilise those criteria that are valid for the particular constituency evaluating the DSS. The context of the evaluation may also strongly influence the evaluation results. Thus, this is important to have a flexible tool for DSS evaluation that takes into consideration all these issues.

Maynard and Arnott (1994) adopt a multiple-constituency approach for DSS evaluation. This approach allows evaluation to be performed from multiple, and some times very different, perspectives of the DSS project. The method was extended with multiple-criteria evaluation (Maynard; Arnott and Burstein, 1995). A comprehensive set of criteria useful for measuring a success of DSS project was identified. The criteria were presented as a hierarchy classified by the different perspectives they measure DSS from; for example effectiveness, efficiency, satisfaction and use, are the classes of the upper level of this hierarchy. The approach allows various groups of people concerned with the DSS to identify which criteria are relevant to their views and how important they are in the current context of evaluation.

Before a useful multiple-criteria, multiple-constituency evaluation of DSS can be undertaken, a method of measuring the often volatile, criteria for each constituency must be developed. An aggregation procedure for the multiple criteria evaluation provided by the constituencies has been proposed. We were unable to directly apply any of the multiple criteria methods that exist for hierarchies of criteria. None of these approaches could adequately cater for the dynamic nature of multiple-constituency evaluation: many using a static number and structure of criteria which is not useful where multiple constituencies are concerned. With multiple-constituency DSS evaluation many criteria may, at any time, be valuable to a constituency group, whilst others may be unimportant, or inconsequential. The proposed procedure is capable of considering the relative importance of the criteria provided by the constituency, as well as individual scores the particular DSS received for these criteria. This is used to show the relative success of the DSS overall and how it performed with respect to any individual group at any level of the criteria hierarchy.

Whilst there are some studies focusing on multiple criteria DSS evaluation (Adelman et al. 1985 and Goicoechea et al. 1992), no studies have approached the evaluation process from a multiple group perspective in a DSS context. As a result there is little research on the use of criteria in which those criteria have failed in the evaluation process, especially where multiple groups are concerned. The authors proposed an approach that attempts to reduce the likelihood of criteria failing in the evaluation process through explicitly defining the criteria ensuring no misinterpretation can take place across constituency groups (Maynard et al 1995).

The aim of this paper is to describe a method that allows for the measurement of those criteria that a constituency sees as relevant at any time to enable the evaluation of a DSS. First, the concept of multiple-criteria, multiple-constituency evaluation of DSS will be introduced. Several methods for dealing with multiple-criteria, both in evaluation and in decision making will be discussed and their advantages and disadvantages highlighted. A method for the measurement of criteria in multiple-criteria, multiple-constituency evaluation will then be presented. A computerised tool implementing the approach is also described.

2. The Multiple-Constituency Approach to Evaluation of DSS

The Multiple-constituency approach was proposed as a way to introduce multiple perspectives on the evaluation of the effectiveness of the organisation (Connolly et al, 1980). The assumption of this approach is that evaluation process needs to be flexible enough to accommodate subjective perceptions about organisational performance to avoid unnecessary biases and distortions from using just one view and measure. The result of such evaluation does not present a common view and may not be generalisable beyond the time the evaluation was done. However, in the situation where there is a need to capture such 'contextual' factors in order to, perhaps, compare the results obtained from the different constituencies or evaluations done over a period of time, the approach gives definite advantage.

The important aspect of the multiple-constituency approach includes identification of all relevant constituencies to be involved in the evaluation process. As part of this research we have identified five major groups of people, constituencies, involved and related to DSS projects at the different stages (Maynard et al, 1995). These constituencies are: *DSS Developer*, *User*, *Decision-maker*, *Management* and *Decision-consumer*. The first four being identified from the past DSS studies, the last, *Decision-consumer* was introduced to fill the gap in the existing literature (Maynard, 1997). The *Decision-consumer* group comprises those people who are directly and significantly influenced by the decision made through using the DSS. The opinion of these people is often different from any other group involved with the DSS and it is very important not overlook them in the evaluation process. These five constituencies represent distinct roles in the DSS project, however, they may be performed by the same people.

3. A Multiple-Criteria, Multiple-Constituency Approach to Evaluation of DSS

After identifying the perspective the evaluator is going to have depending on the role they play in DSS project at the evaluation time, there is a need to formulate some suitable measurements - criteria for evaluation. As it was mentioned above there is a need for dynamic set of such criteria as some of them play more important role for some constituencies but not for the others, and may become irrelevant if the context of evaluation has been changed. As a result of an extensive literature review, a generic set of DSS evaluation criteria was identified. This set is presented to the evaluators in the evaluation as a source from which they identify the subset of criteria suitable for the particular perspective or constituency role that they adopt for the evaluation process. For the logical consistency this set was classified in a hierarchy.

At the top level, the measurement of system success can accomplished through the assessment of DSS from four different perspectives, or domains. These domains are *effectiveness*, *efficiency*, *use* and *satisfaction* (Maynard et al. 1995). Within each of these domains a number of criteria exist that may be important for one or more constituencies. Usually, criteria are referred to using various terminology and are rarely defined. In a multiple-constituency approach, however, constituency groups may not have the same understanding of meaning for criteria when they are not formally defined. Thus, it becomes critical for criteria to be defined. Maynard et al (1995) explicitly define each criterion to avoid confusion amongst constituencies as to its meaning.

In the multiple-constituency evaluation process described later, criteria are 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 criteria "perceived usefulness of the system". Measurement may occur along the bi-polar scale "not useful - very useful". Within the evaluation process, members of constituency's will also weight the importance of each criterion to their work, and will then rate the DSS with respect to the each criterion at the bottom levels of the hierarchy. These ratings and weights then will be aggregated throughout the hierarchy and each node in the hierarchy will be given an evaluation "score". Scores equate to the success rating of the DSS based on the perspective of each constituency for that criteria set.

Using the hierarchies of criteria for evaluation process allow each constituency to view their evaluation outcomes within any level of detail within the hierarchy. This may enhance the interaction and understanding that constituencies have of the evaluation.

In the next section we present a method of aggregating the results of multiple-criteria, multiple-constituency evaluation of DSS. This method has been developed and implemented in a tool that automates the proposed approach.

3.1. Methods for multiple criteria evaluation

Several formal methods exist for decision making. In general, they use a number of criteria, and use a process of weighting and scoring each criterion based on several possible outcomes, to select, or suggest the "best" outcome for a given situation. These multiple criteria methods are termed multi-criteria decision-making methods. Multiple-criteria methods can be split in to two major areas: decision making and decision aiding. Each area, in general, uses a hierarchical structure to aid in selecting the best alternative from several alternatives - decision making approaches, or in aiding the decision-maker to select an alternative - decision aiding approaches (Saaty 1990; Burstein, et al. 1983). Whilst in evaluation we are measuring the success of a DSS rather than selecting an alternative, some methods used in multiple-criteria decision making may be adapted to evaluating DSS.

Probably the most widely used aggregation rule and scoring procedure is a weighted linear average (Edwards 1977), similar to Equation 1 below and to the methods described by Adelman et al. (1985) and Goicoechea et al. (1992).

Equation 1: A Weighted Average

$V_j = \sum_j W_j V_{ij}$, $\sum_j W_j = 1.0$, Where V_j is the overall value

V_{ij} is the value of any element.

W_j is the weight associated with that element.

Edwards (1977) states: “the weighted linear average is a reliable method that produces values that are extremely close to very much more complicated non linear and interactive ‘true’ utility functions”. The weighted linear average method is complimented by its ease of use and ease of understanding when compared to other methods (see Wilks (1938), Dawes and Corrigan (1974), Einhorn and Hogarth (1975), Ridel and Pitz (1986)). To use the linear weighted average in a hierarchical structure, the weighted average is initially applied at the bottom level of the hierarchy. Then, each successive level of the hierarchy has the weighted averages calculated. This continues until the top level of the hierarchy is calculated. This may be useful in the evaluation process.

Huber (1980), from a multiple-criteria decision making perspective, describes a method that uses weights to determine the importance of criteria. His model has essentially four phases:

1) Elicit weights for each criterion

   Rank each criterion in terms of their relative importance. Assign 100 to the most important criterion and assign other values between 0 and 100 to the other criteria that reflect their importance from the least to the most important criterion.

2) Normalise weights for each criterion

   Convert the weights of each criterion into proportional weights by taking the weight of the criterion and dividing it by the total weight of all criteria. As such, the proportional weights of the criteria add to 1.

3) Calculate weights for all attributes

   For the attributes of a criterion, rank the attributes in terms of their importance, assign a value of 1 to the most important or most relevant attribute. Assign values between 0 and 1 to the other attributes. Divide each attribute's value by the sum of them all and determine the proportional weights for the attributes by multiplying each attribute’s weight by the criterion’s proportional weight.

4) Calculate each alternatives score

For each alternative, score the alternative between 0 and 100 for each of the attributes. The utility of the alternative becomes the sum of its score multiplied by the proportional weight for each attribute.

Keeney and Raiffa (1976) and Edwards (1977) discuss similar approaches. Huber's (1980) method differentiates between several alternative courses of action, for instance, the selection of whether to locate a service station at X, Y or Z. The determination of proportional weights maintains the internal consistency of this approach. For any level in the criteria attribute hierarchy, the sum of all weights on that level is 1. So, the total weight for all attributes is 1, as is the total for all the weights of criteria. This approach may be useful in an evaluation process.

From an evaluation perspective, Adelman et al. (1985) suggest a method for evaluating DSS similar to those discussed so far. In this method a hierarchy of attributes is created with each attribute having the same weight. The method first scores each of the bottom level attributes. 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 use this method in DSS evaluation (Adelman and Donnell 1986, Hopple 1989, Andriole 1989, Adelman 1989).

Adelman (1989) lists several caveats to this approach. The comparison of attributes in the manner he describes requires that a common scale must be used, otherwise comparisons would be meaningless. In his work, since only attitudes are measured, the common scale used was a utility (or value) scale that has scores ranging from 0 to 10. Adelman (1989) also states that it is possible to weight the attributes and then use the weights to determine the utility of attributes throughout the hierarchy. But when Adelman (in each of the papers which discuss this method) uses the approach he ignores weights so that each criteria has the same weight, because, as he explains, it is "inappropriate at the moment to weight criteria" (Adelman et al. 1985, Adelman and Donnell 1986, Adelman 1989).

Goicoechea, Stakhiv and Li (1992) build on Adelman et al.'s (1985) approach. They average the responses for questions relating to a criterion and then use weighted averages to progress up the hierarchy. This method is distinct from Adelman et al.'s (1985) in that weights of each family are used to determine the score for the parent node. In Adelman et al.'s (1985) method the weights were ignored altogether. Yet, the method still uses averages to get the score of the bottom level nodes. This may be detrimental. For example, if alternatives A, B, and C were scored as 0.1, 0.6, and 0.9 respectively a score of 0.5333 would be produced for the parent (Equation 2).

Equation 2: An averaging method of scoring

$$Score_{Parent} = \frac{Score_A + Score_B + Score_C}{N}$$, Where N equals the number of alternatives.

$$Score_{Parent} = \frac{0.1 + 0.6 + 0.9}{3}$$

$$Score_{Parent} = 0.5333$$

But if the alternatives were weighted to indicate importance then this score would not be the same. For instance, the importance weights for alternatives A, B and C may have been 0.9, 0.04, and 0.06 respectively. This would make the score of the parent (using a weighted average) 0.056 (Equation 3), a vast difference from 0.5333.

Equation 3: A weighted average method of scoring

$$Score_{Parent} = \frac{(Score_A * Weight_A) + (Score_B * Weight_B) + (Score_C * Weight_C)}{N}$$

Where N equals the number of alternatives.

$$Score_{Parent} = \frac{(0.1 * 0.9) + (0.6 * 0.04) + (0.9 * 0.06)}{3}$$

$$Score_{Parent} = 0.056$$

Given this, the importance of weighting the bottom level criteria is clear. This approach, as it stands, may not be appropriate for the evaluation process.

Another approach is the analytical hierarchy process, AHP (Saaty 1980). This is a general approach to measurement that derives ratio scales from both discrete and continuous paired comparisons which are used to produce an overall rank of alternatives (Saaty 1994, Saaty 1990). To rank the alternatives the AHP uses pair-wise comparisons along a ratio scale (usually language expressions such as *equal importance*, *strong importance* and *extreme importance*, which equate numerically to 1, 5 and 9 respectively). To determine which alternative is best several steps are carried out.

a) Determine which criterion is most important

b) Determine which alternative is best for each criterion

c) Determine the best alternative to select

At the completion of these three steps, the decision-maker is presented with the best alternative given the weighting and scoring of the criteria.

3.2. Advantages and Disadvantages of these Methods

The advantages and disadvantages of the above methods will influence selection of an appropriate method for weighing and scoring criteria in the evaluation process. Saaty (1994) criticises the approach Huber (1980) and others used because of differences in scale. He states that the most objectionable approach is to assign a set of numbers to judgements on alternatives under a particular criterion and to then normalise these numbers (so they add to 1 by multiplying the weight by the reciprocal of their sum). Problems 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. For example, if one scale set was the range 1-5 and the other was 12-19 each set would still be normalised to the interval 0-1. 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 (1980) approach allows the calculation of meaningful scores for parents, as weights of the whole tree are applied at the bottom level. It is then a simple case to add weighted scores up the hierarchy to get scores for successive parents. But 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 meaning. For example, a three level hierarchy branch with weights of 0.1 for each level would yield a criterion weight of $0.1 \bullet 0.1 \bullet 0.1 = 0.001$. If this was factored down another level this would become $0.1 \bullet 0.1 \bullet 0.1 \bullet 0.1 = 0.0001$. Thus, for unevenly levelled trees the method Huber discusses becomes less practical.

In the 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 $\frac{(N^2 - N)}{2}$. This compares poorly to the approach Huber suggests where the number of questions required to determine weights of criteria equals the number of criteria. The intrusiveness and time required conducting the AHP in the DSS evaluation process will be detrimental and may result in the evaluation process not being conducted, as the number of criteria would be usually large. In addition, the AHP attempts to compare several alternative actions, but the evaluation process does not require a comparison. The evaluation, rather, is attempting to measure the attitudes of each constituency towards the DSS, rather than comparing their attitudes for one DSS with another.

The approaches based on weighted averages (Adelman et al. 1985, Goicoechea et al. 1992) assume that the criteria are linearly related and thus use addition to get the score for the upper levels of the hierarchy. This assumption can only be made when additivity conditions are met (Keeney and Raffia 1976). This, however, is generally reasonable assumption. Edwards (1977) 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”. Dawes and Corrigan (1974) also concur with Edwards’ views stating that linear models offer good approximations to hierarchical and multivariate models in many decision-making situations. Additive linear approaches have some advantages: they are simple to use, do not add much time to the evaluation process, and make it easy to explain to those evaluating the DSS (if necessary) how the values throughout the hierarchy are obtained, giving further meaning to the evaluation. This type of approach, defined by Goicoechea et al. (1992) and Adelman et al. (1985), requires further definition before it can be used directly for an evaluation. For instance, these approaches neglect to specify how weights were obtained throughout the hierarchy (where they were obtained, rather than being ignored completely, as in Adelman’s case) and whether they have been normalised, and if so, how.

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 number of levels (all the bottom level criteria are at level three for instance). However, once varying depths of the hierarchy 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 (1980) approach). Depending on the purpose of the hierarchy this may be suitable, but for an evaluation procedure where differing levels may be the norm, this is not the case.

Thus, the review presented above justifies that for the purpose of the DSS evaluation process, none of the approaches discussed seemed adequate. The approach we adopt is based on the combination of these approaches and addresses the problems identified. The next section describes this approach.

4. A multiple-constituency, multiple-criteria, hierarchical based approach to the evaluation of DSS

This section describes an approach for weighting and scoring the criteria in evaluating a DSS. It uses several approaches described previously, drawing specifically from those Huber (1980), Adelman et al. (1985) and Goicoechea et al. (1992) discuss to form a method of evaluating DSS in a multiple-constituency, multiple-criteria, hierarchical based environment. The following describes this combined approach.

The evaluation of a DSS essentially aims to measure its success from the perspective of those with a significant stake in it, using the criteria each the stake-holder believes is important. The problem with traditional approaches is that they lack the concept measuring the DSS based on distinctly separate stakeholders and often produce a single aggregate evaluation result as a consequence.

The approach used to weight and score each criterion here assumes that the person doing the evaluation builds a hierarchy of criteria. In our case, the generic hierarchy exists and specific criterion may be added to it. The approach then uses these criteria to determine the success of the DSS. Rather than producing a single outcome for the success of the DSS for criteria for all stakeholders, the proposed method produces an outcome for each relevant criterion in the hierarchy for each of the stakeholders. This produces a comprehensive, cohesive evaluation outcome. The steps for accomplishing this are outlined below.

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

For each of the nodes in the hierarchy, the person evaluating the DSS rates how important each criterion is to them. This can be done on a purely subjective basis and responses transformed to numerical values (Saaty 1990). In the proposed approach, the person evaluating the DSS is not aware of the numerical values used. They only see a bi-polar scale that is labelled at each end. Figure 4.2 shows an example of such a scale. The method weights the each node of the hierarchy on a scale of 1 to 100 with 0 reflecting that the criterion is not applicable in this instance.

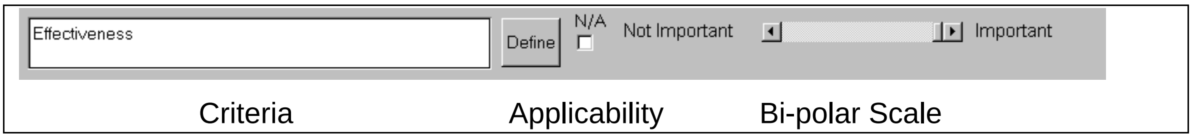


Figure 4.2: Scale for weighting criteria

Practically, the scale could take on any reasonable values as the person evaluating the DSS has no knowledge of it. The only impact on evaluation from the scale would be on the size of the scores. If the scale was kept constant for the evaluation of a particular DSS over its life this would not be an issue. But one cannot change the scales used in the evaluation process and expect to compare scores generated with the new scale with those using the old scale, as differences of scale would be apparent. This means that throughout the life of the DSS the rating scale must remain the same if any comparison of evaluation scores is made over time. A change to the scale at any time after the initial evaluation will make any comparison over time worthless.

Step 2: Convert the weights into normalised weights (add to 1 for each level of the hierarchy)

For each element of a level determine the normalised weight working from the bottom to the top of the hierarchy.

In a hierarchy with N levels ($i = \overline{0, N}$) where N is the maximum depth in the hierarchy.

M_i = number of objects in level i

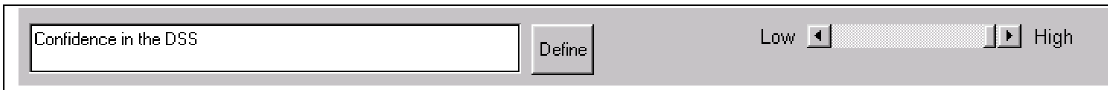
$M_0 = 1$

$$\overline{W}_{ij} = \frac{W_{ij}}{\sum\limits_{i \le j \le M_i}^{m_i} W_{ij}}$$

Where $\overline{W}_{ij}$ is the normalised weight for criteria.

Step 3: Get the scores for each criteria (bottom levels only) from the evaluator

For each of the bottom level elements get the evaluator to rate how well the system matches the criteria. These answers will be subjective and be converted to numerical values. In a similar manner to the weighting of the criteria, the evaluator will rate the system on a scale. The evaluator is unaware of the scale used and as long as the scale is used consistently throughout the evaluation it can take on any reasonable range of values. Like the weighting of criteria the scale used is between 1 and 100. A bi-polar scale is used similar to the one presented in Figure 4.3.



Criteria Bi-polar Scale

Figure 4.3: Scale for scoring criteria

$S_{N_j} = scores$

$S_{mi}(1 \le m \le M_i)$ - scores for the objects which are on the last level of the hierarchy in their branch

Step 4 Use the 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.

$\overline{S}_{Mi} = \overline{W}_{iM} \bullet S_{Mi}$ calculate weighted score

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

Start from the level N - the deepest level

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, meaning the dependant objects from the level immediately below in the hierarchy.

$$\tilde{S}_{mi} = \overline{W}_{mi} \bullet \sum_{i=1}^{n} \overline{S}_{mi-1}$$

Where i is a number of level (*0* to N), m is the number of objects in the level i $(1 \le m \le M_i)$

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 additive weighted sum method can be used. But, as Edwards (1977) 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.

5. MaltiVal - A tool for DSS Evaluation

The method described above was implemented in a prototype DSS evaluation system called MaltiVal (Maynard, 1997). It acts as an aid to the evaluation process, rather than controlling it in a similar manner to a DSS helping the decision-maker rather than making the decision. It has been implementing using Microsoft Excel™, which allows it to be used in both Windows™ and Macintosh® operating system environments. This makes the tool even more flexible as it cam be portable and can be used in various software platforms.

The system fully supports the entire evaluation process of the DSS project as well as storing the results for further comparison. It begins by registering a particular user in a particular role of the constituency group they represent. Then the user is identifies a sub-set of criteria relevant to the situation out of the full set provided. The system provides definitions of all the terms and criteria used to prevent misunderstanding and to make sure the criteria measure what the user intends.

The evaluation module of the system collects the weights and the scores from the user and aggregates them according to the procedure described above. The results are presented to the user in any necessary level of details, eg. by level, by group of criteria in the hierarchy, by the area, as well as how this compares to the maximum possible result of the perfect DSS (based on their importance weights). This shows the user how well the current system performed in different areas.

Different users of the systems have various level of access to the aid. For example, there is a possibility of adding some criteria, if the user deems it necessary.

6. Summary And Conclusions

In this paper we presented a method of DSS evaluation which allows the consideration of an interest groups opinion in a DSS project. The approach lets the person evaluating the system give weights to the criteria used based on their perceived perception of the importance of those criteria for the particular DSS. Such an approach provides a dynamic environment to capture different perspectives as well as monitor changes that are necessary to be managed in order to achieve an overall success of the project.

The method proposed to evaluate a DSS helps to reveal the differences in perceptions about the project and its outcomes, as well as capture and reflect the context of the evaluation explicitly. Using the MaltiVal tool, the evaluation can be preformed quickly and efficiently at any point to give material for discussion of the further improvements to the DSS under consideration before the project is complete. It can also be used to provide feedback after the system is complete and is operational.

The MaltiVal tool was applied in an illustrative example that showed a definite usefulness of this approach. The tool was used by the user and developer of a DSS prototype and the results of these two evaluations were different enough to give good indication of the different expectations these people had. The results have also indicated what were the areas of improvements that the system could have.

There are some issues still to be researched if such an approach is adopted as a real world application. They are mainly concerned with the logistics of the process and some managerial implications that the introduction of such a process can have. For example, the privacy of the results and the ownership of the information are some issues to consider. However, these questions can only be addressed in the field study stage of this research that is now under consideration.

References:

Adelman, L. ‘Integrating Evaluation Methods into the DSS Development’. *Information and Decision Technologies*, 15:4, 1989, 227-241.

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.

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).

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

Burstein, F.V.; Biniashvili, N.M. and Muskhelishvili, K.G. ‘Development of procedure for criteria importance values estimation on hierarchical structure for decision making tasks.’ *Theoretical Cybernetics*-2, Tbilisi: “Metsniereba”, 1983, 27-34.

Connolly, T., Conlon, E.J. and Deutsch, S.J. 'Organisational Effectiveness: A Multiple Constituency Approach', *Academy of Management Review*, 5:2, 1980, 211-217.

Dawes, R.M. and Corrigan, B. ‘Linear models in Decision Making’, *Psychological Bulletin*, v81, 1974, pp97-106.

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.

Einhorn, H.J. and Hogarth, R.M. ‘Unit Weighting Schemes of Decision Making’ *Organisational Behaviour and Human Performance*, v13, 1975, pp171-192.

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.

Hopple, G.W. ‘Decision Support Systems: Software Evaluation Criteria and Methodologies’. *Large Scale Systems*, 12:3, 1987, 285-300.

Huber, G.P. *Managerial Decision Making*. Scott, Foresman and Company. Glenview, Illinois, 1980.

Keeney, R.L.; and Raiffa, H. *Decisions With Multiple Objectives: Preferences and Value Tradeoffs*. John Wiley and Sons Inc. New York. 1976.

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

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.

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.

Ridel, S.L. and Pitz, G.F. 'Utilisation - Orientated Evaluation of DSS'. *IEEE Transactions on Systems, Man and Cybernetics*, SMC 16, 1986, 980-996.

Saaty, T.L. *The Analytic Hierarchy Process*, McGraw-Hill, New York, 1980.

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

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.

Wilks, S.S. 'Weighting Systems for Linear Functions of Correlated Variables Where There is no Dependant Variable' *Psycometrika*, v3, 1938, 23-40.