

A NEW MODELLING FRAMEWORK FOR ORGANISATIONAL SELF-ASSESSMENT: DEVELOPMENT AND APPLICATION¹

C. H. R. Siow, J. B. Yang and B. G. Dale, Manchester School of Management, UMIST, UK

ABSTRACT

This paper reports the development of a multiple attribute decision making (MADM) modelling framework and methodology for application to the scoring process against the criteria of the European Foundation for Quality Management (EFQM) Model for Excellence. The MADM model is based on the evidential reasoning approach with the focus of the initial development and application on the people management enabler. The model built has been tested with real data from two award application documents from an electricity distribution utility and the scores achieved using the methodology have proven the validity and reliability of the Model.

KEY WORDS: EFQM Excellence Model, Self-Assessment Process, Multiple Attribute Decision Making

INTRODUCTION

Ahire and Rana (1995) and Hwang and Yoon (1981) outline how a Multiple Attribute Decision Making (MADM) model allows the decision-maker to evaluate, along common multiple criteria, various competing alternative courses of action to achieve a pre-specified goal. These and other writers describe how MADM methods are technical decision aids for evaluating alternatives that are characterised by multiple qualitative (or fuzzy) and quantitative attributes. It is also pointed out by Eom (1989) that multiple attribute decision making problems with these types of attributes are commonplace in business practice. Self-assessment against the European Foundation for Quality Management (EFQM) Excellence Model as typically outlined by Hakes (2000) can be considered to be part of the general class of organisational evaluation problem. Hwang and Yoon (1981) argue that all problems share the following characteristics:

- . Each problem has multiple objectives/attributes and the decision-maker must generate relevant attributes for problem setting.
- . Multiple criteria usually conflict with each other.
- . Each attribute may have a different unit of measurement.

These characteristics are found in the self-assessment process against the criteria of the EFQM Model and therefore it is possible to apply the MADM modelling logic to develop a methodology for assisting the scoring process against the criteria of the EFQM Model. Ahire and Rana (1995) have identified the following three elements that a MADM model application to a decision problem should possess:

- . identifying the decision hierarchy in terms of criteria, sub-criteria, sub-sub-criteria;
- . determining the relative importance weights of these elements of hierarchy; and

¹ Quality Management Journal, Vol.8, No.4, pp.34-47, 2001

- . assessing and ranking various alternatives using these criteria.

The first step in applying the MADM logic to a problem is to evaluate and quantify the state of an attribute. This approach may be conceptually clear and easy to understand, however, Yang and Singh (1994) and Yang and Sen (1994) have demonstrated the type of difficulties that can be encountered in this process. The main difficulty is the requirement for attributes to be potentially independent and quantitative. Measurements of qualitative attributes are highly subjective and imprecise and, as such, it is more effective to evaluate them using subjective judgements with uncertainty. This is typical of the problem under consideration in this paper - self-assessment against the criteria of the EFQM Excellence Model - in which the process of scoring is influenced by: the background and prejudices of the assessors; how they interpret an organisation's approach; deployment, and assessment and review; their level of understanding of the TQM/excellence/continuous improvement philosophy and the sub-criteria under consideration; and how they interpret the detail in an organisation's submission document. As a consequence there is some degree of subjectivity and uncertainty in the scoring activity and, as a result, the process is not scientifically robust with respect to numerical evaluation. The evidential reasoning approach (Yang and Singh (1994), Yang and Sen (1994) and Yang (2001)) provides an alternative way to deal with such uncertain synthesis problems by means of evidence combination for multiple attributes.

The application of MADM methods to the self-assessment process against the criteria of a Quality/Excellence Award Model to date have been given no attention by academic researchers. This may be due to the recent rise in popularity of self-assessment and the technicality of the MADM methods that tend to favour scientific decisions and with which few quality management researchers are familiar. A review of MADM methods carried out by the researchers has shown their potential for application to the self-assessment environment; this also established that the evidential reasoning approach was the most appropriate. This paper reports the development of a MADM modelling framework and methodology to assist in the scoring of the sub-criteria of the EFQM Model. The focus of the initial development and application has been the people management enabler of the Model.

EVALUATION FRAMEWORK

In a MADM model, criteria can be described on the basis of attributes. These attributes are broken down into their simplest form in such a manner that they can be readily and easily evaluated by means of either numerical values or subjective judgements with a level of uncertainty. By using the evidential reasoning approach, the evaluations of these basic attributes can then be aggregated into an overall evaluation of the people management criteria, the subject of the examination with respect to the 1998 version of the EFQM Model. Yang and Sen (1994) have developed a new general multi-level and distributed evaluation process based on an evidential reasoning framework where uncertain subjective judgements for the evaluation of qualitative attributes can be accommodated. With this new approach, qualitative attributes may be evaluated by uncertain subjective judgements through multiple levels of factors and each of the judgements may be assigned by single or multiple assessors in a rational way within the evidential reasoning framework. This approach has recently been enhanced so that both qualitative and quantitative attributes can be dealt with under the same framework (Yang (2001)) and has been integrated in a newly developed window-based software package called Intelligent Decision System (IDS) (Yang and Xu (2000)). The process of constructing the MADM Model and methodology and performing the evaluation will be based on this platform.

Undertaking a self-assessment against the EFQM Excellence Model is based on the transformation of original data into equivalent subjective statements with uncertainty. As precise numerical values are not usually available for the enablers (i.e. leadership, policy and strategy, resources, people management, and processes) of the Model, it is then more natural to articulate subjective judgements with uncertainty as original evaluation data, as detailed by Yang and Sen (1997). Furthermore the scoring of such qualitative issues will generate high uncertainty during evaluation (Van der Wiele et al (1995)) and therefore the evidential reasoning approach is the ideal tool to handle this situation.

With the MADM approach an overall set of evaluation grades is first defined, and each of the base level attributes is judged relative to these grades, which may range from excellent to very poor, or from fundamental to minor. According to Yang and Sen (1997) the MADM approach works on the assumption that 'if the evaluation of a basic attribute is, to a certain extent, judged to be good, then the evaluation of its associated upper-level attribute will, to some degree, also be good'. The evidential reasoning approach provides a systematic way of synthesising such uncertain evaluations of basic attributes to produce an evaluation for an associated upper level attribute.

The following are some of the points on which the construction of the evaluation framework has been based:

1. Attributes on the same level should be of similar importance. If they are not, then too much credence can be given to any one of them.
2. A decision must be made as to when to stop the breakdown of attributes. The attribute can be broken down to an infinite number of levels but there is a stage where further breakdown will provide little further accuracy. On the other hand, too little breakdown can also result in an evaluation that is too subjective and inaccurate.
3. The framework must be an accurate representative of the people management criteria being modelled.

The EFQM Excellence Model (1998) consists of five enablers and four results criteria broken down into 32 sub-criterion parts (e.g. the people management criterion has 6 sub-criteria). This structure has been employed as the guideline for the construction of the evaluation framework. The people management criteria and its 6 sub-criterion parts are referred to as level 0 and level 1 attributes respectively. The level 1 attributes can be further broken down into the level 2 attributes of areas to address which are detailed in the evidence for each sub-criteria. An example of the detailed breakdown of each attribute of the people management criteria is given in Table 1 and their respective descriptions can be found in EFQM (1998). A key decision is whether or not to carry out a more detailed breakdown of these attributes, requiring a clear understanding of the attributes and their underlying factors. If the attributes are broken down beyond level 2, researchers can produce a different framework for the evaluation of the same enabler and as a result the benefit of the generic MADM Model is lost.

The sub-criterion 3(a) - How people resources are planned and improved - is used as an example

to illustrate how the framework has been formulated. Under this level 1 attribute, the EFQM Model breaks this down to the following 5 attributes (level 2); in self-assessment terms these are the areas which need to be addressed in collecting evidence in support of the 3(a) sub-criteria.

1. How the organisation aligns the human resources plan with policy and strategy.
2. How the organisation develops and uses employee satisfaction surveys.
3. How the organisation ensures fairness in terms of employment.
4. How the organisation aligns its remuneration, redeployment, redundancy and other terms of employment with policy and strategy.
5. How the organisation uses innovative work organisation strategies and methods to improve the way of working.

Each of these level 2 attribute is further broken down into two level 3 attributes - approach and deployment. These are defined in the EFQM (1998) as:

“The enabler criteria are concerned with how the organisation approaches each of the criterion parts.

Information is required on the excellence of the approach used and the extent of deployment of the approach - vertically through all levels of the organisation and horizontally to all areas and activities.”

EVALUATION GRADES AND DEFINITIONS

The MADM model requires a generalised set of evaluation grades before an assessment can be undertaken. Central to decision theory and to the evaluation of grades in the MADM model is the concept of ‘value’. This is the measure of what is good or desirable about a design or its attributes (Siddall (1972)) and demonstrates that the desirability of each outcome can be decided by attaching to it some value. For example, when assessing a quantitative attribute such as: the ‘number of team briefings conducted annually’ there is little chance of evaluating whether the reported result (e.g. 24 times) per year is good or bad. This indicates the need for a range of values to evaluate the performance of an attribute. According to Siddall (1972) utility value can be described as that relating to the function or usefulness of an attribute. Using the example of team briefing assessments, here is a range of values from 0-24, with 24 having a utility value of 1 (the best), 0 as having a value of 0 (the worst), and other values having utilities between 1 and 0.

Whilst the same guiding principles apply to qualitative judgements, there are some differences. The judgements given by assessors must also relate to unity. The best judgement possible is related to the utility value of one and the worst to zero. For example, if a 5 point evaluation scale is chosen and calibrated starting with the end points, then 1 point may be given to the maximum value that is practically or physically realisable and 0 point to the minimum value. The midpoint (0.5) will also be a basis for calibration, since it will be the breakpoint between values that are favourable (or better than average) and values that are unfavourable (or worse than average).

With reference to the examples of qualitative and quantitative judgements, the conventional procedures that derive these numerical values use addition and multiplication across the attributes. This type of scaling assumes that a value of 0.75 is three times as favourable as a value of 0.25. The combination of values across attributes implies that the difference between any two specific values is the same for each attribute. It is obvious that the assignment of such values is arbitrary. Many scales are possible and it is important to provide some consistency checks or definitions in these scales. These checks/definitions are desirable but they make the scaling procedure a tedious activity, posing many questions to the assessor.

The evaluation grades and definitions have been formed in three stages, (i) the assessment of each level 2 attributes according to a group of definitions or checklist items, (ii) the establishment of the scale for the evaluation grades; five grades were selected - 'world-class', 'award winners', 'improvers', 'drifters' and 'uncommitted' and (iii) each checklist item is then defined in terms of the five grades, using the scoring matrix approach (EFQM (1998), Porter and Tanner (1998)).

Checklists for each qualitative attribute

To provide some consistency and definition in the evaluation scales, a checklist of items is used for evaluation of each of the level 2 attributes in the 'people management' enabler. These items have been formed from the work by Godfrey et al (1998) and various checklists acquired from management consultancies. An example for the criteria "How the organisation aligns the human resources with policy and strategy" is shown in Table 2. These lists of items can be further developed from the documents produced by Award winning companies (e.g. BT United Kingdom (1997) and TNT United Kingdom (1998)).

The checklist enables the assessors to have a clear idea of what to look for in the evaluation of each attribute. It also serves as a reference to enable the assessors to make reliable and accurate judgements and thereby minimise the subjectivity of the assessment process. The checklists can be modified to suit an organisation's operating environment, however, changes should only be done as a last resort when there are too many 'un-assessable' items. It is recommended that during assessment, 'un-assessable' items should not be graded and taken into consideration during the final assessment.

Evaluation grades scale

Dale and Lascelles (1997) have identified six different levels of characteristics and behaviours - *world class*, *award winners*, *improvers*, *tool pushers*, *drifters* and *uncommitted* - in relation to the adoption of Total Quality Management (TQM). These give definitions to the evaluations grades and also show hierarchical differences between the levels. This is required as there is a hierarchical difference in the utility value of each evaluation grade in the MADM model, the best judgement possible is related to the numerical value of one and the worst to zero. To prevent the evaluation scale becoming too complex and difficult to differentiate between the levels, the six levels of TQM adoption have been reduced to 5 levels; based on the advice of Dale and Smith (1998) the '*tool pushers*' and '*drifters*' levels have been combined. The scale of '*world class*' to '*uncommitted*' is used to describe the evaluation grades, as detailed below:

The general scale of evaluation grades, H , is defined as

$$H = \{H_1 \quad H_2 \quad H_3 \quad H_4 \quad H_5\}$$

$$p\{H\} = \{1 \quad 0.75 \quad 0.50 \quad 0.25 \quad 0\}$$

{ *World-class, Award winners, Improvers, Drifters, Uncommitted* }

The description of each grade level is taken from Dale and Lascelles (1997) and Dale and Smith (1998).

Evaluation grades definitions

During the evaluation of complex qualitative attributes such as people management, there are no mathematical formulas and any attribute can be assigned to the general set of grades without validity. To think of an attribute merely in terms of ‘*world class*’ to ‘*uncommitted*’, places an undue burden on the assessor to translate judgements directly into single word evaluations. A better approach is to define each evaluation grade with more precision of what a ‘*world class*’ attribute can achieve and, on the other hand, what an ‘*uncommitted*’ organisation cannot achieve.

In this MADM model, all the level 2 attributes have been defined in terms of the checklist of items with each item in turn defined in evaluation grades from H_1 to H_5 . For example, considering the attribute - how the organisation aligns the human resources with policy and strategy, a *world class* grade definition for each of the checklist items is as follows:

- | | |
|--------|--|
| Item 1 | Comprehensive evidence of the total integration of Human Resources Management (HRM) and quality management as part of the same initiative and moving towards a common goal of excellence. |
| Item 2 | Comprehensive evidence of the Human Resources (HR) function having an influence in the development of business policy and strategy. |
| Item 3 | Comprehensive evidence of an HR strategy incorporating plans for selection, training and succession. Clear evidence of regular review of these plans leading to improved business effectiveness. |

On the other hand, the ‘*uncommitted*’ evaluation definition of this attribute is the reversal of ‘*world class*’.

- | | |
|--------|--|
| Item 1 | No evidence of the integration of HRM and quality management. |
| Item 2 | No evidence of the HR function having an influence or input in the development of business policy and strategy. |
| Item 3 | No evidence of an HR strategy incorporating plans or selection, training and succession. There is also no evidence of assessment and review. |

The definition of each grade takes into account the guidance specified by the EFQM (1998) in terms of approach and deployment.

Based on EFQM (1998) and Porter and Tanner (1998), Table 3 shows the scoring matrix that is used to evaluate the approach, differentiating between the different evaluation grades.

Table 3 - Approach Scoring Matrix

Score dimension	0	25	50	75	100
Sound prevention-based system	No evidence	Little evidence	Evidence	Extensive evidence	Comprehensive evidence
Review and refinement	No evidence	Occasional review	Regular review	Regular review leading to refinement	Regular review leading to improved business effectiveness
Integration into normal operations	No evidence	Some	Well established	Good	Total
Evaluation grades	Uncommitted	Drifters	Improvers	Award winners	World class
Utility value	0	0.25	0.5	0.75	1

The ‘deployment’ attribute, on the other hand, is concerned with the extent to which the approach has been implemented taking into account its full potential and the appropriateness and effective application of the approach. The following are used to define the different evaluation grades of $H_1 - H_5$:

World Class	Complete and full integration into normal operations
Award Winners	Well integrated but not totally integrated into normal operations
Improvers	Average levels of integration
Drifters	Some areas of integration
Uncommitted	No integration

Weightings

To design an accurate MADM model the weightings of each attribute have to be considered as they allow the alteration of the composition of attributes evaluation in favour of certain sub-attributes. Without weightings, all attributes are considered equal in making up the overall evaluation of the higher level attributes. All the attributes in this evaluation framework are given the same weightings as the EFQM Model (i.e. in the people management criteria there are 6 sub-criteria). The 9% or 90 points weighting for the people management enabler in the EFQM Model will only be considered when all the enablers are combined.

There is a need to have a clear understanding of what is important when assessing people

management to effectively determine the weightings for each item on the checklist. The total weightings given to all the checklist items for one parent attribute should be no more than 100%. The weightings given are subjective since they are based on presumption, understanding and experience, hence the relative importance of each checklist item will vary, depending on the researcher and organisation. To reflect their own culture, operations and processes it is possible for an organisation to change the weightings for the more important items.

COLLECTION AND EXTRACTION OF TEST DATA FOR EVIDENCE

The data used for testing this model has been taken from the 1997 and 1998 EFQM Model application documents of an electricity distribution utility. Unlike other MADM projects such as design selection where the information obtained is in no particular format, hence making the extraction of data a difficult task, this MADM model has been designed in line with the 1998 EFQM Excellence Model. The evidence for the assessment of a particular criterion part will be in the same format and the assessor will just have to classify this evidence with respect to the defined grades. There are no set rules on the usage of data and it is possible that data from one criterion part in the application document can be used as evidence for assessing a different criterion part.

The framework has been tested by evaluating the data to the set of evaluation grades. Whilst it was easy to extract the relevant data from the two application documents, this was not the case in mapping the extracted data to the set of evaluation grades. For example, when assessing the attribute - evidence of a systematic approach to succession plans - the following is the data presented in the 1997 Application Document.

‘The appraisal process helps to identify people for succession planning. By assessing appraisal notes and their own personal knowledge the utility identified potential senior managers and specific training plans were developed to enhance their careers. With the introduction of the new organisation in October 1996 this planning enabled the utility to appoint staff of a suitable calibre into all new posts. As a result of SVS the age profile of the company changed and so a new project is underway to create a comprehensive succession planning process’.

The wordings and terms used in the document does not always coincide with the definition and meaning of the evaluation grades as given below:

World class - Comprehensive evidence of a systematic approach to succession plans. Clear evidence that these plans are being regularly reviewed, leading to improve business effectiveness.

Award winners - Extensive evidence of a systematic approach to succession plans. Clear evidence that these plans are being regularly reviewed and refined.

Improvers - Evidence of a systematic approach to succession plans. Clear evidence that these plans are being regularly reviewed.

Drifter - Little evidence of a systematic approach to succession plans. Clear evidence that these plans are being occasionally reviewed.

Uncommitted - No evidence of a systematic approach to succession plans and any plans are not reviewed.

The assessor needs to interpret this extracted evidence and categorise it into one or more of the five grades. For such qualitative evaluation as people management, precise numerical values are not available and in this process of evaluation an item's grade can lie between 'world class' and 'award winner', and using subjective judgement the assessor can assign degrees of confidence or belief to these grades. If it is felt that the item lies closer to the 'world class' definition, then the evaluation may become 'world class' (0.8), 'award winner' (0.2), indicating a degree of confidence/belief in the evaluation. The total degree of belief for one grade could be less than 100% but should not be more than 100% and this is one of the main advantages of the evidential reasoning approach. The degrees of belief are multiplied by the weighting to produce an assessment for each of the level 2 attributes. These assessments can be aggregated and the final assessment for the people management enabler can be generated using the IDS package.

An assessment can be characterised using either a distribution in terms of degrees of belief or expected utilities (scores). There are three possible utilities - 'worst possible', 'average' and 'best possible' - and these are obtained by taking into consideration the missing evidence. As a result of missing evidence, some degree of belief may be left 'un-graded' and could be assigned to any of the grades. The IDS software uses the theory of evidence to analyse these uncertainties by taking account of two extreme scenarios. One scenario is that the missing evidence could support the worst grade 'uncommitted' and the other is that the missing evidence could support the best grade 'world class'. This leads to the generation of two utilities 'worst possible' and 'best possible' utilities between which a utility interval is comprised. This allows the organisation to assess where they stand in terms of the worst and best possible scenario, without having total certainty of the assessment. The average result is obtained by giving no preference to any of the grades and this is done by assigning equally 'ungraded' degrees of belief to the missing evidence.

During an evaluation there are some checklist items that cannot be fully evaluated and assessed with confidence without a site visit. A site visit is an opportunity to talk with people in the organisation and to confirm the validity of the data and to clarify aspects that are unclear. If in the evaluation of an attribute the evidence justifies a degree of belief of 50%, but this evidence is questionable because it is not clear how widely the attribute is deployed within the organisation, then as a result, the remaining 50% degree of belief is left 'un-graded', resulting in a total degree of belief for that attribute of less than 1. If a site visit can make clear the scope of deployment, the remaining degree of belief may be assigned to individual grades. However, there may be a case when there is no wish to conduct a site visit to verify the evidence. In this situation it is possible for the assessor to implant some personal judgements into the evaluation process. Here, the assessor's perception in terms of inference and deduction for the missing evidence is based on his/her belief of how well the organisation can possibly perform in that attribute. This method of assessing missing evidence is subjective but useful if a site visit is not possible.

TESTING OF THE MODEL

The model has been tested using the utilities two application documents. Analysis of the 1997 document reveals that the average overall utility value obtained for the assessment is 0.366 or

36.6% and is midway between the drifters and improvers levels of TQM adoption (Siow (1999)). Figure 1 provides a graphical representation of this type of analysis.

The highest score for a level 1 sub-attribute is - (3f) How people are cared for - with an average score of 45.8%, see Figure 2. This score is obtained from its child level 2 attributes scores which has a highest average score of 62% and a lowest average score of 31% (Yang et al (2001)). The lowest score of the level 1 attribute is - (3a) How people resources are planned and improved - with a score of 27.7%, which is obtained from considering all its level 2 attributes (Siow (1999)). The highest score among these level 2 attributes is 45.4% and the lowest score is 14%. As explained by Yang et al (2001) the overall score of 36.6% is within the range of the lowest and highest score of each lower level attribute, indicating that the mathematical process for this model is accurate and that the score of 36.6% is valid.

The MADM model can also help with identifying strengths and areas for improvement (AFIs) by analysis of the appropriate attributes. This can be done in two ways:

- . By examining the lowest or highest score of each level's attribute through the framework to the base level. (Hierarchical framework method)
- . By identifying all the sub-attributes that score higher (strengths) or lower (AFIs) than the overall score. (Relative score method)

The hierarchical framework method is less accurate than the relative score method. With the former method only those attributes that are in most need of improvement and those in which the organisation has done well are identified. By following the highest score of each attribute through the framework to the base level attribute, the assessor can identify the checklist items that have scored high, along with the evidence and comments made (see Figure 3). The AFIs are identified in the same manner by following the lowest score of each attribute through the framework to the base level attribute to identify which checklist item received the lowest score.

The hierarchical framework method has two weaknesses. Firstly, the assessor is only able to identify areas of attention through its parent attribute. That means that if the parent attribute is scored low or high then more investigation is needed on the respective child attributes. However, there may be cases where the parent attribute has been scored high but one of the child attributes scored low. This means the low scoring sub-attribute will be 'disguised' amongst a group of high scoring sub-attributes. Secondly, this method, by focussing on the level 1 attribute with the lowest or highest score, allows the remainder of the level 1 attributes to be left out of the identification process and, as a result, some areas which need attention may not be identified. Having made these points it must be said that the more focussed hierarchical framework method allows a quick assessment to be made.

This problem can be minimised by modifying the method so that it starts from a level 1 rather than a level 0 attribute. This allows an examination of every level 1 attribute that gives the assessor a wider view of the scored document. This method can be rather tedious and is only recommended if a very detailed feedback report is required.

The relative score method uses the relative score (overall score) of the level 0 attribute to identify

strengths and AFIs. It allows all the attributes of the same level (usually the lower or base level attribute) to be compared to a relative score with those attributes with a higher or lower score identified for further investigation. To identify the strengths from the scoring of the 1997 application document all the level 2 attributes with scores higher than 36.6% are considered to be a strength. It is also possible to examine in this way all the level 3 attributes (i.e. approach and deployment), however it is felt that the findings could be too detailed. AFIs are located by identifying all attributes that have scored lower than 36.6%.

It is clear that the hierarchical framework and relative score methods have advantages and disadvantages. The former method is able to establish a list of items linked to a parent attribute, emphasising the weakest or strongest parent attribute, however it suffers from being too narrow in its scope of investigation. On the other hand, the relative scoring method is more elaborate, taking into account all criterion parts and is also more flexible as any strengths and weaknesses above or below a specified score are considered. This will lead to a large number of areas identified which, dependent on the assessor's objectives, can be minimised by concentrating on a smaller number. Building on their respective advantages the two methods can be used together to produce a feedback report.

COMPARATIVE ANALYSIS OF THE AWARD SIMULATION DOCUMENTS

The IDS software employed in the analysis can be used by an organisation to keep track of its progress over a period of time, offering a logical and systematic way for comparative analysis. With the relevant data inputted into a database, comparison can be done between alternative courses of action and benchmarking undertaken. This form of comparative analysis is given in Figure 4. It can be seen that the 1998 submission has improved from a score of 36.6% to 43.3%. The score of 36.6% is somewhere between that of drifters and improvers, whilst that of 43.3 is close to improvers.

With the hierarchical framework and the IDS software, the AFIs from the previous years assessment can easily be identified from the difference in the scores. The methodology can also be used to simulate different improvement strategies, based on the number of AFIs identified from the self-assessment process. Each checklist item of attributes relating to a sub-criterion can be categorised into short or long-term improvements. The items identified are then assessed to an evaluation grade that the assessor believes can be achieved within the chosen time frame.

IMPLEMENTATION CHALLENGES AND LIMITATIONS

The MADM modelling framework developed on the basis of the evidential reasoning approach allows the assessors to map evidence extracted from self-assessment documents and site visits to the defined evaluation grades. This mapping process provides a natural way of interpreting a wide range of evidence, most of which is of a qualitative nature. In this process, the assessors are not forced to pre-aggregate qualitative evidence into a single numerical score. However, the mapping process does require the assessor to provide his (or her) assessments in a different way from conventional scoring and it is not straightforward to map the extracted evidence to the set of evaluation grades. Appropriate training is therefore needed to implement the MADM modelling framework. It should be noted that the ER approach can only help to reduce but cannot completely eliminate subjectivity in self-assessment. This is because the mapping process still requires the subjective interpretation of extracted evidence in terms of the evaluation grades by individual assessors. It should also be noted that in the current ER modelling framework evaluation grades must represent distinctive standards in assessment. In other words, they are

required to be mutually exclusive. This is satisfied in the EFQM model as the five evaluation grades are related to percentage scores from 0% to 100%. Research is currently being conducted to allow the use of dependent (or non-exclusive) evaluation grades in the framework.

As discussed previously, the use of the checklists (or guidelines) can help assessors to improve the accuracy and consistency in their self-assessment. It has been recognised that it requires in-depth research and takes time and resources to identify commonly used guidelines for assessment criteria and areas to address. As part of such effort, several research projects are being conducted by the authors to investigate and organise self-assessment guidelines for constructing a knowledge base and decision models. Obviously, there is a need to stimulate more research in this area to develop widely accepted knowledge bases to support the self-assessment process. It should be noted that for comparative analysis within an organisation or among a number of organisations, it is important to keep the modelling framework consistent and intact, which means that the same checklists and weighting systems should be used to assess all organisation in the analysis.

Finally, the criteria aggregating process of the ER approach is not as simple as the additive weighting method as used in the conventional scoring method. The implementation of the proposed modelling framework and the methodology relies on the availability of easy-to-use computer software packages. The intelligent decision system mentioned in this paper provides a window-based package to support self-assessment. However, it takes time to get the awareness and acceptance of this and similar software packages. With the rapid development of computer technology, however, it will only be a matter of time for practitioners to employ computer software packages like IDS for self-assessment.

CONCLUSIONS

The process of defining the evaluation grades was done as part of a three stage process, first by defining the attributes with a precise list of checklist items, secondly by deciding on the evaluation grades and lastly by defining each checklist item in terms of the different grades. Subjectivity is an inevitable problem due to the qualitative nature of the people management enabler which was selected for investigating the feasibility of the application of MADM. By designing a MADM model with precise definition for each qualitative attribute and using a mathematical method via the evidential reasoning approach, it has the potential to reduce the subjectivity and biases of the scoring activity in undertaking a self-assessment against the criteria of the EFQM model and will improve the validity and reliability of the scoring process. The MADM model is also able to identify areas of attention for feedback. This is done using either or a combination of the hierarchical framework method and the relative score method.

A comparative analysis of the host utilities two award simulation documents showed that the MADM model coupled with the use of the IDS software provides a more logical and systematic way for comparative analysis of the results. This is extremely useful as it enables an organisation to keep track of its progress over time.

The assessor, using the MADM model, is required to allocate an evaluation grade which they believe to be appropriate and in this way the model is able to effectively simulate the results of a real life self-assessment. By simulating different improvement strategies it allows an organisation to plan effectively which improvement and changes will give the best results in terms of achieving excellence. Two different scenarios had been designed into the MADM

model in terms of the time taken for improvement to be implemented (i.e. short term and long term).

ACKNOWLEDGEMENT

The constructive comments on this paper from the anonymous referees are greatly appreciated. This work is partly supported by the UK Engineering and Physical Science Research Council (EPSRC) under the Grant No: GR/N27453/01 and the Grant No: GR/N65615/01.

REFERENCES

Ahire S. L. and Rana D. S., (1995), Selection of TQM Pilot Projects Using an MCDM Approach, *International Journal of Quality and Reliability Management*, 12 (1).

BT United Kingdom, (1997), *Quality Award 97*, BT, Belfast.

Dale B. G. and Lascelles D. M., (1997), Total Quality Management Adoption: Revisiting the Levels, *The TQM Magazine*, 9 (6), 418-428.

Dale B. G. and Smith M., (1997), Spectrum of Quality Management Implementation Grid: Development and Use, *Managing Services Quality*, 7 (6), 307-311.

Eom H., (1989), The Current State of Multiple Criteria Decision Support Systems, *Human Systems Management*, 8 (2), 113-119.

European Foundation for Quality Management, (1998), *Self-Assessment: Guidelines for Companies*, EFQM, Brussels.

Godfrey G., Dale B. G., Wilkinson A. and Marchington M., (1997), *An Audit Tool for Self-Assessment of Human Resources Policies Within a Total Quality Management Environment*, Manchester School of Management, UMIST.

Hakes C., (2000), *The Business Excellence Handbook (Fifth Edition)*, European Quality Publications, London.

Hwang H. and Yoon K., (1981), *Multiple Attribute Decision Making - Methods and Applications*, Springer-Verlag, New York.

Porter L. and Tanner J., (1998), *Assessing Business Excellence*, Butterworth-Heinmann, London.

Siddall J. N., (1972), *Analytical Decision Making in Engineering Design*, Prentice Hall, Englewood Cliffs.

Siow C. H. R., (1999), *Development of a Distributed Evaluation Model for Self-Assessment Based on the EFQM Model for Business Excellence*, M.Sc. Dissertation, Manchester School of Management, UMIST.

TNT United Kingdom, (1998), *Submission for the 1998 European Quality Award*, TNT, Warwickshire.

van der Wiele T., Williams A. R. T., Kolb F. and Dale B. G., (1995), *Assessor Training for the European Quality Award*, *Quality World Technical Supplement*, March, 12-18.

Yang J. B. and Singh M. H., (1994), An Evidential Reasoning Approach for Multiple Attribute Decision Making with Certainty, IEEE Transactions on Systems, Manufacturing, and Cybernetics, 24 (1), 1-17.

Yang J. B. and Sen P., "A general multi-level evaluation process for hybrid MADM with uncertainty", IEEE Transactions on Systems, Man, and Cybernetics, Vol.24, No.10, 1994, 1458-1473.

Yang J. B. and Sen P., (1997), Multiple Attribute Design Evaluation of Complex Engineering Products Using the Evidential Reasoning Approach, Journal of Engineering Design 8 (3), 211-230.

Yang J. B. and Xu D-L., (2000), Intelligent Decision System Via Evidential Reasoning, Version 1, IDSL, Cheshire, UK.

Yang J. B., (2001), Rule and Utility Based Evidential Reasoning Approach for Multiple Attribute Decision Analysis Under Uncertainties, European Journal of Operational Research (in press, EJOR Paper #98287).

Yang J. B., Dale B. G. and Siow C. H. R., (2001), Self-Assessment of Excellence: an Application of the Evidential Reasoning Approach, International Journal of Production Research (under review).

Table 1 - Evaluation Framework**3a. How people resources are planned and improved**

1. aligns the human resources plan with policy and strategy
2. develops and uses employee satisfaction surveys
3. ensures fairness in terms of employment
4. aligns its remuneration, redeployment, redundancy and other terms of employment with policy and strategy
5. uses innovative work organisation strategies and methods to improve the way of working

3b. How people capabilities are sustained and developed

1. identifies, classifies and matches people's competences with its needs
2. manages recruitment and career development
3. establishes and implements training plans
4. reviews the effectiveness of training
5. develops people through work experience
6. develops team skills
7. promotes continuous learning

3c. How people agree targets and continuously review performance

1. aligns individual and team objectives with its targets
2. reviews and updates individual and team objectives
3. appraises and helps people improve their performance

3d. How people are involved, empowered and recognised

1. encourages and supports individuals and teams participation in improvement
2. encourages people's involvement through in-house conferences and ceremonies
3. empowers people to take action and evaluates effectiveness
4. designs the recognition system to sustain involvement and empowerment

3e. How people and the organisation have an effective dialogue

1. identifies communication needs
2. shares information and has a dialogue with its people
3. evaluates and improves communication effectiveness
4. structures top down, bottom up and lateral communication

3f. How people are cared for

1. promotes awareness and involvement in health, safety and environmental issues
2. sets the level of benefits (such as pension plan, health care, child care, etc.)
3. promotes social and cultural activities
4. provides facilities and services (flexible hours, transport, etc.).

Source: EFQM (1998)

Table 2 - Definitions for Evaluation Grades in Terms of Checklist Items**3a. How people resources are planned and improved****3a1. How the organisation aligns the human resources plan with policy and strategy*****World class*****Approach**

- . Comprehensive evidence of total integration of Human Resource (HR) management and quality management as a part of the same initiative and moving towards a common goal of business excellence (35%).
- . Comprehensive evidence of HR function having influence in the development of business policy and strategy (35%).
- . Comprehensive evidence of HR strategy plan incorporating selection, training, and succession plans. Clear evidence of regular review of plans leading to improved business effectiveness (30%).

Deployment

- . Comprehensive evidence of skills of employees mapped to all levels of organisational needs (50%).
- . Comprehensive evidence of HR strategy plan incorporating selection, training, and succession plans is used throughout the organisation (50%).

Award winners**Approach**

- . Clear evidence of good integration of HR management and quality management as a part of the same initiative and moving towards a common goal of business excellence (35%).
- . Extensive evidence of HR function having influence in the development of business policy and strategy (35%).
- . Evidence of HR strategy plan incorporating two out of the three plans - selection, training and succession. Clear evidence of regular review of plans leading to refinement (30%).

Deployment

- . Evidence of skills of employees mapped to most of the organisation needs (50%).
- . Evidence of HR strategy plan incorporating selection, training and succession plans is used in most of the organisation (50%).

Improvers**Approach**

- . Evidence of integration of HR management and quality management as a part of the same initiative and moving towards a common goal of business excellence (35%).
- . Evidence of HR function having influence in the development of business policy and strategy (35%).
- . Evidence of HR strategy plans incorporating only one of the three plans - selection, training, and succession. Clear evidence of regular review of plans (30%).

Deployment

- . Evidence of skills of employees are mapped to half of the organisation needs (50%).
- . Clear evidence of HR strategy plan incorporating selection, training, and succession plans is used in half of the organisation (50%).

Drifters

Approach

- . Evidence of some area of integration of HR management and quality management as a part of the same initiative and moving towards a common goal of business excellence (35%).
- . Little evidence of HR function having influence or input in the development of business policy and strategy (35%).
- . Little evidence of HR strategy plans incorporating only one of the three plans - selection, training, and succession plans. Evidence of occasional review of plan (30%).

Deployment

- . Evidence of skills of employees mapped to a small area of the organisation needs (50%).
- . Evidence of HR strategy plan incorporating selection, training, and succession plans is used in a small part of the organisation (50%).

Uncommitted

Approach

- . No evidence of integration of HR management and quality management (35%).
- . No evidence of HR function having influence or input in the development of business policy and strategy (35%).
- . No evidence of HR strategy plan incorporating any selection, training, and succession plans. No evidence of review of plan (30%).

Deployment

- . No evidence of skills of employees mapped to organisation needs (40%).
- . No evidence of HR strategy plan incorporating selection, training, and succession plans is used in the organisation (50%).

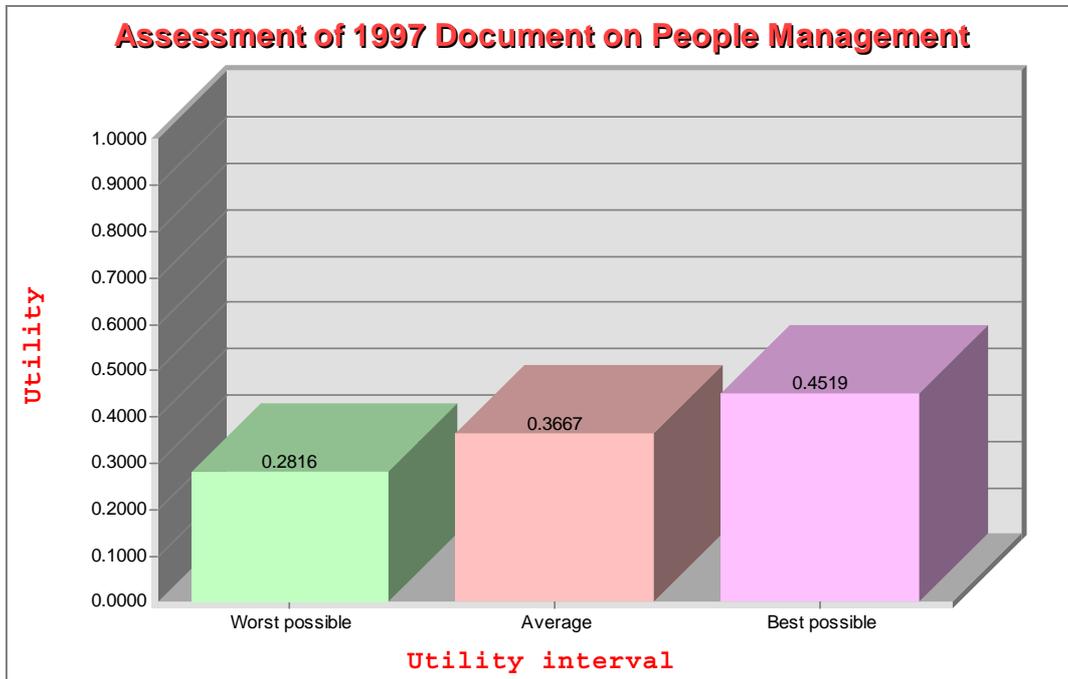


Figure 1 Self-Assessment Results: 1997 Submission Document

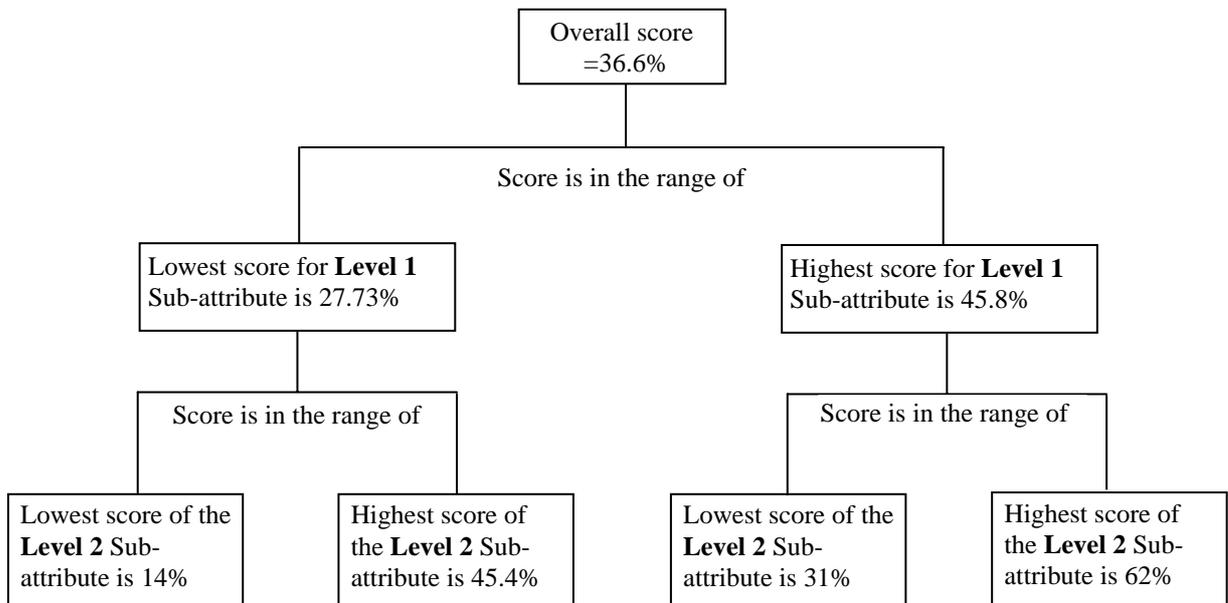


Figure 2 Scoring Breakdown

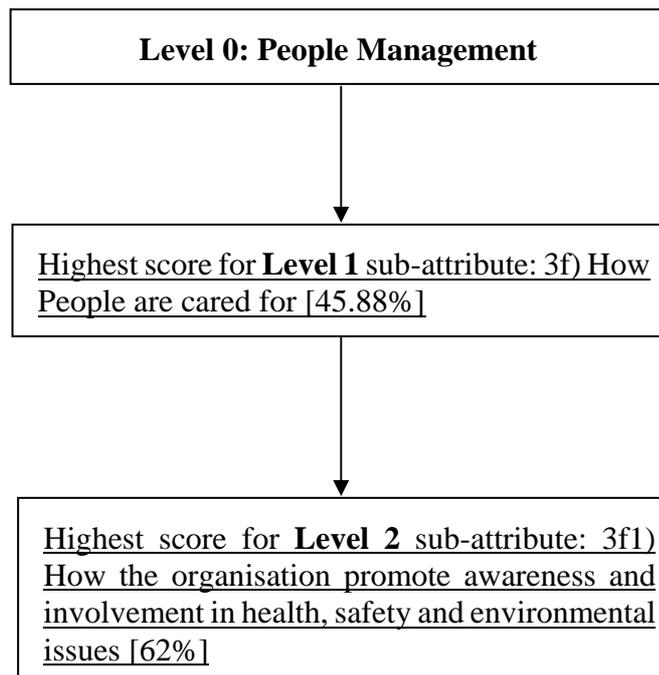


Figure 3 Identification of Strength

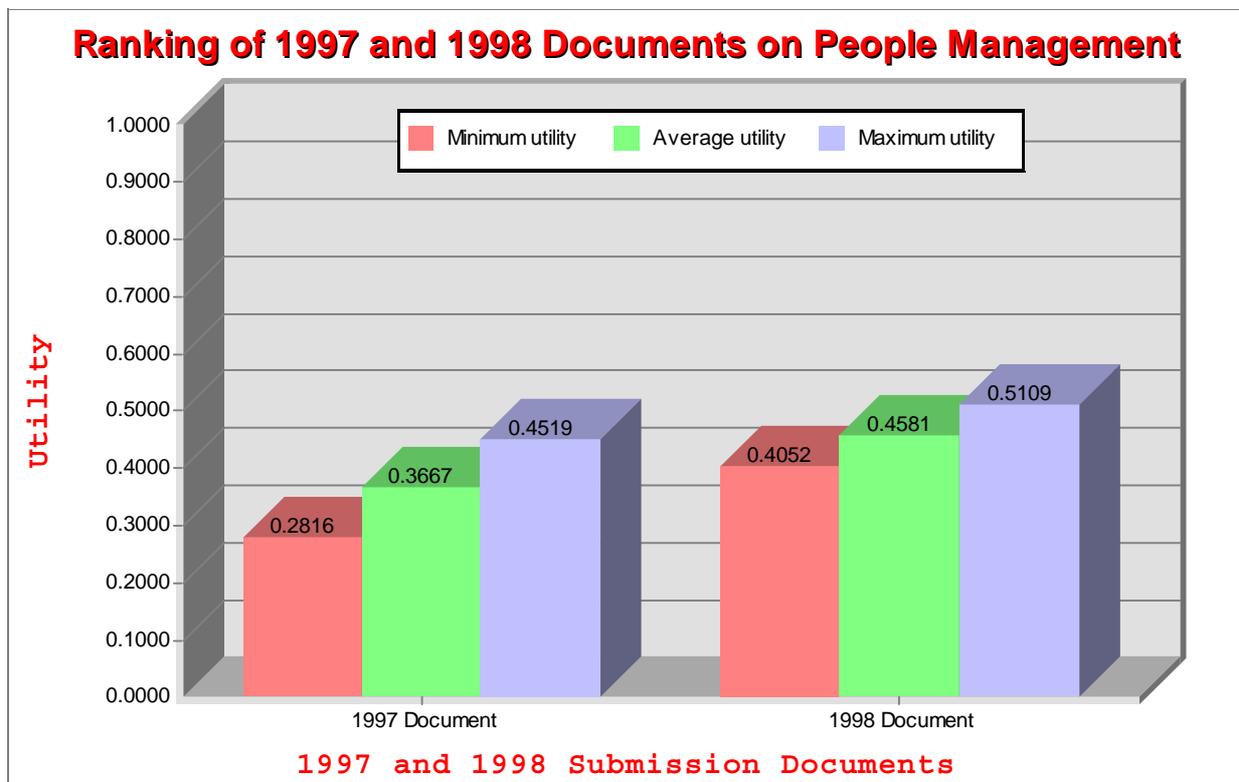


Figure 4 Ranking of the 1997 and 1998 Submission Documents for the People Management Criteria