Pergamon

0306-4573(93)E0018-J

A QUALITATIVE DISCRIMINANT PROCESS FOR SCORING AND RANKING IN GROUP SUPPORT SYSTEMS

NOEL BRYSON
Department of Information Systems and Analysis, School of Business,
Howard University, Washington, DC, U.S.A.

OJELANKI K. NGWENYAMA
Computer and Information Systems, School of Business Administration,
The University of Michigan, Ann Arbor, MI, U.S.A.

and

AYODELE MOBOLURIN
Department of Information Systems and Analysis, School of Business,
Howard University, Washington, DC, U.S.A.

Abstract — Group support systems (GSS) are increasingly being used within organizations to support group work. One area of support that is often desired is the scoring and ranking of alternatives on qualitative/subjective domains. In this article, we present a new, conceptual approach, the qualitative discriminant process, for scoring and ranking in GSS. This approach is based on well-established decision analysis techniques. It significantly advances the state of the art of GSS by addressing four common limitations: (1) the inability to deal with vagueness of human decision makers in articulating preferences; (2) difficulties in mapping qualitative evaluation to numeric estimates; (3) problems in aggregating individual preferences into meaningful group preference; and (4) the lack of simple, user-friendly techniques for dealing with a large number of decision alternatives. Our approach is easy to implement in stand-alone personal computers and GSS platforms. We illustrate this with a real-world problem on a prototype implementation.

1. INTRODUCTION

The turbulence of the modern business environment has influenced the way in which firms are organizing for decisionmaking and problem solving (Huber, 1984a). Flexible organizations, in the form of teams and autonomous work groups, seem to be the preferred approach for dealing with the complexity and diversity of everyday business problems (Finholt & Sproull, 1990). Since the early 1980s researchers and developers have responded to the challenge of supporting group work by building and studying a new class of IS, called group support systems (GSS). Early GSS focused on group decisionmaking (Huber, 1984b; Kull, 1982); more recently, however, other types of systems such as meeting support systems (MSS, Dennis *et al.*, 1988) and negotiation systems (NSS, Jarke *et al.*, 1987) have emerged which target other types of group activity. Much of the research in the field has focused on developing GSS environments and studying their effects on group processes (McLeod, 1991). Returns on investments are encouraging; several studies have reported evidence of "process gains" such as improved effectiveness of group activity (Adelman, 1988; Vogel & Nunamaker, 1988; Dennis *et al.*, 1988; Nunamaker *et al.*, 1991) and reduction in meeting lengths (Grohowski *et al.*, 1990; Nunamaker *et al.*, 1989; Vogel & Nunamaker, 1988). Some studies have shown that GSS features which support: (1) the evaluation of decision alternatives; (2) voting (Kull, 1982; Nunamaker *et al.*, 1988; Galegher, 1990); and (3) group memory retention (Grohowski *et al.*, 1990; Finholt & Sproull, 1990) contribute to group process gains. Although much research has focused on understanding how GSS might improve group processes, little attention is given to improving some of these features. For example, the areas of voting and evaluation techniques have been neglected.

In group decisionmaking, the evaluation of decision alternatives often involves scoring and ranking of the alternatives in ways that are best described as qualitative or subjective. For example, TELOS-PDM (Burch, 1992), a commonly used technique for information systems development planning, requires that each information system project be evaluated in terms of its expected contribution to a firm's goals of improving productivity, differentiation, and management (PDM). During the group decisionmaking process, each member of the IS planning committee is required to evaluate project proposals based on the three criteria and to provide a composite score for each proposal. The scores of all the committee members are then analyzed and merged into a "group ranking" of the projects after which the final decision is made. The goal of deriving a composite score for each proposal necessitates the assignment of a numeric score for each criterion, even though some criteria, such as differentiation, obviously have a qualitative/subjective domain. In a group context, this situation becomes aggravated by the fact that an acceptable "group ranking" is the desired objective. Thus it is necessary that the committee is able to define a "consensus" measure, identify when a satisfactory level of "consensus" has been achieved, and to compute a "group" score for each project proposal. There are several other examples of group activity with similar requirements for evaluation support: employee performance evaluation, candidate selection, strategic planning, product planning, negotiation, investment decisionmaking with social criteria, and so on.

Many scoring and ranking techniques commonly used in GSS platforms have been criticized in the decision theory literature for their limitations in dealing with the real difficulties of group decisionmaking: (a) mapping of qualitative evaluations to point estimates in ranking (Goddard, 1983; Weber, 1987); (b) the aggregation of individual preferences into a group preference (Kirkwood & Sarin, 1985; Dyer & Sarin, 1979); and (c) the analysis and use of point estimate data in facilitating group consensus formation (Dutta, 1980; Bropan *et al.*, 1992). Although many relevant advances have been made in voting, fuzzy set, and possibility theories, these have not significantly influenced current GSS development.

In this study, we propose a new conceptual framework for scoring and ranking that involves a multistage qualitative discriminant process. Our framework is informed by voting, fuzzy set, and possibility theories. It provides techniques that are better suited to facilitating consensus formation in group activities than currently exist in most GSSs. It offers the following advantages: (1) a clear and simple structured graphical approach to collecting data from users; (2) maps qualitative evaluations to numeric estimates; (3) allows for vagueness in preference articulations; (4) provides support for analyzing data relevant to evaluating consensus formation; and (5) ease of implementation in manual and computer supported group activities.

2. SCORING AND RANKING TECHNIQUES

The expression of individual preferences among a set of decision alternatives may appear to be simple and straightforward on the surface, but there are several issues which must be addressed. For example, the ranking of any set of objects implies the explicit or implicit scoring/rating of the objects with regard to some common basis of comparison, and an ordering of them in reference to the assigned scores or ratings. A fundamental implication here, is that relationships exist among the objects which provide the basis for comparison. In this regard, measuring the strengths of these relationships and assigning weights/scores to them is an important issue in scoring and ranking.

Current techniques for scoring and ranking can be classified into four general categories: (1) point estimates on interval scales; (2) point estimates on ratio scales; (3) interval estimates on ratio scales; and (4) interval estimates on interval scales (Fig. 1). Although there is some discussion about using category (2) techniques in GSS software (Saaty, 1989), most current systems support the first category. User familiarity and ease of implementation in software may account for the wide acceptance and implementation of these point-estimate techniques in GSS.

More recently, however, point-estimate techniques have been criticized for several limitations: (1) they do not address the fuzziness which is characteristic of many human deci-

	Interval Scales	**Ratio Scales**
Point Estimates	Category 1	Category 2
Interval Estimates	Category 4	Category 3

Fig. 1. Categories of ranking and scoring techniques.

sionmaking problems (Weber, 1987; Hurrion, 1985; Korhonen, 1985); (2) it is very difficult to map qualitative preferences, which may have a range, to point estimates (Kirkwood et al., 1985; Goddard, 1983); (3) decisionmakers often operate in situations where incomplete information makes it impossible to assign point estimates to decision alternatives (Weber, 1985, 1987; Goddard, 1983; Korhonen, 1985); (4) in voting situations where coalitions may develop, point estimates can be manipulated (Dutta, 1980; Nurmi et al., 1990); and (5) aggregating the point estimates of individual decisionmakers to determine group preferences is problematic because, although individuals may agree on the qualitative rank of an object, they often disagree on the point-estimate for it (Kirkwood & Sarin, 1985; Weber, 1985; Dyer & Sarin, 1979).

Many researchers have argued that pairwise comparison techniques can help solve many of the aforementioned problems (Weber, 1985, 1987; Goddard, 1983; Kok & Lootsma, 1985; Kirkwood & Sarin, 1985). Advances in fuzzy set, possibility, and voting theories have also contributed to improved understanding of many of these problems and the development of more sophisticated pairwise comparison techniques which use interval estimates (Zadeh, 1965; Huber et al., 1969; Yager, 1982; Nurmi et al., 1990). Some important features of pairwise comparison techniques are: (a) they help to reduce complexity escalation in situations where decisionmakers are required to simultaneously compare more than five to seven objects; (b) they are applicable to decision situations where information about decision alternatives may be incomplete; (c) subtle distinctions and relationships can be communicated by ordering alternatives along a scale without deriving "exact" numerical values.

2.1 Limitations of current GSS techniques

Current approaches to scoring and ranking in GSS are for the most part based on category (1) techniques. For example, group system's (alias PLEXSYS, TEAM QUEST), alternative evaluation, and group matrix tools both utilize a 10-point interval scale for rating evaluation criteria and decision alternatives (Nunamaker et al., 1991). Vision Quest, another well-known GSS platform also uses a 10-point interval scale for the same purpose. These GSS platforms exhibit four basic limitations: (1) Lack of structured techniques for dealing with the problem of distinguishing among large numbers of alternatives. Thus, as the number of objects and evaluation criteria increases, it becomes increasingly difficult to rank them (Korhonen, 1986, 1987, 1991). (2) They do not elicit information on decisionmakers' preferences, which could be analyzed by facilitators and help determine strategies for consensus formation. (3) They do not provide techniques for mapping qualitative criteria to numeric data for analysis. (4) They are inadequate techniques for aggregating and analyzing individual preferences to facilitate consensus formation and for deriving group preference.

In general, the ranking processes implemented in most GSS environments presents each decisionmaker with the decision alternatives (objects) and requires the decisionmaker to

order them in a manner consistent with his/her beliefs. In the case of scoring, each decisionmaker is presented with a list of alternatives, a numeric interval of acceptable scores, and a pair of qualitative categories that is associated with the bottom and top values of the interval scale. The decisionmaker is then requested to assign an integer value from the relevant interval to each alternative. Apart from the extreme values on the scale, the meanings of the other values are usually not defined. Thus, it is possible that two different decisionmakers may associate two different meanings to the same numeric value. Similarly, a decisionmaker may associate the same meaning to two different values in the interval. These factors imply that the evaluation of the responses of the group, in order both to determine the consensus response and the level of consensus, is a problematic undertaking.

Since scoring with reference to a common basis of comparison also implies the existence of relationships between the objects relative to the basis of comparison, the scores provided by two different group members could reflect expressions of the same relationships (4:2:1) between the objects even though the actual scores may be different [e.g., (12:6:3) and (8:4:2)]. Thus, the examination of the scores in order to determine the level of consensus might lead to an incorrect conclusion — for consensus involves both consensus in terms of the relationships between the objects, and consensus in terms of the score associated with the top-ranked object. This perspective of consensus might not be apparent to the decisionmaker if there is no explicit association between the scores of pairs of objects. Current approaches to scoring in GSS limit the users' perspective of consensus, and do not elicit information about the relationships among preferences. The one exception here is Option Finder, which offers a tool for eliciting this information via a pairwise comparison technique. It is, however, limited because it elicits ordinal and not cardinal preference information. Consequently, the decisionmaker has no way of stating that he/she is indifferent to the relative rank positions of a given pair of decision alternatives.

Another important limitation is the approach to aggregating individual preferences into group preferences. The common technique used in GSS is arithmetic averaging, in which a group mean score/rank is derived. The deviation of each individual score from the mean is also computed. This information is then used to stimulate dialogue and bring the group members closer to the mean. The major problem here, however, is that a group mean is meaningless. For example, if the mean score is 5 for two decisionmakers who ranked decision alternatives on a 10-point (0–9) scale, and the scores of both decisionmakers are equidistant from the mean (1, 9), the deviation of each decisionmaker from the mean does not in any way represent how far apart they are from consensus on the ranking of this decision alternative.

3. QUALITATIVE DISCRIMINANT PROCESS

The approach that we are proposing provides a process and a structured comparison technique for making qualitative distinctions among decision alternatives, which overcome the four main limitations of current GSS approaches. It also offers the option to map the qualitative distinctions to numeric estimates from vague real numbers (VRN; cf. Parik, 1983). Our approach is informed by pairwise comparison techniques (Thurstone, 1927; Saaty, 1986, 1989; Harker, 1987a,b; Bryson *et al.*, 1992; Bryson & Mobolurin, 1993), fuzzy set theory (Zadeh, 1965; Yager, 1982), theory of vague real numbers (Parik, 1983), and voting theory (Novak, 1980; Nurmi *et al.*, 1990). In the remainder of this section, we will discuss the basic concepts and definitions relating to our approach. In the next section, we will present the procedure.

3.1 *Vague real numbers*

Given the vagueness inherent in many scoring and ranking processes, the concept of a *vague real number* (VRN) appears to be relevant. Parik (1983) defines a vague real number as a tuple (s_L^i, s_U^i) where $s_L^i < s_U^i$, and both s_L^i and s_U^i are rational numbers. Also, the VRN (s_L^{iA}, s_U^{iA}) is said to be less than the VRN (s_L^{iB}, s_U^{iB}) if $s_U^{iA} < s_L^{iB}$.

We submit that, for subjective and qualitative domains, the score for an object is best represented numerically by a VRN rather than by a single point. For the score for an object

is the set of values that the given evaluator uses, implicitly or explicitly, in providing numeric interval or point estimates for the object. And in situations involving vagueness, it is likely that on occasions such a set will consist of a range of values rather than a single point. The following three remarks describe properties of VRN that are important for the remainder of this discussion.

REMARK 1. If a decisionmaker E associates the VRN (s_L^t, s_U^t) as the score of object t, then any real number $s^t \in (s_L^t, s_U^t)$ could be provided as a numeric point estimate of the score of t.

REMARK 2. If a decisionmaker E associates the VRN (s_L^t, s_U^t) as the score of object t, provides two different real numbers s^{tA} and s^{tB} as numeric point estimates of the score of object t, and if $s^{tA} \in (s_L^t, s_U^t)$ and $s^{tB} \in (s_L^t, s_U^t)$, the decisionmaker is not being inconsistent.

REMARK 3. If a pair of decisionmakers E_A, E_B associate the same VRN (s_L^t, s_U^t) as the score of object t, but provide two different real numbers s^{tA} and s^{tB} as numeric point estimates of the score of object t, if $s^{tA} \in (s_L^t, s_U^t)$ and $s^{tB} \in (s_L^t, s_U^t)$, then it would be incorrect to conclude that the decisionmakers disagree on a score for object t.

Given the importance of achieving a high level of consensus in group decisionmaking processes, then the previous three remarks are instructive. For if different numeric point estimates are representative of the same VRN, then we may incorrectly conclude that a situation of disagreement exists when in fact there is agreement. But how can we identify the relevant VRN? And, if we succeed in this task, how can we be sure that two different decisionmakers associate the same meaning to a specific VRN? For it is the meaning that is important. We believe that a qualitative discriminant process of successive rounds of categorization — from broad qualitative categories to numeric point estimates — can achieve the objective of reasonably mapping meaning to numeric estimates.

3.2 Qualitative discrimination

Let Q be a complete set of ordered mutually exclusive qualitative categories Q_i such that Q_{i1} is considered to be superior to Q_{i2} for each $i_1 > i_2$, where I is the index set of the qualitative categories. Also, let $S = (s_L, s_U)$ be a numeric interval scale such that each object $t \in T$ is to be assigned a value in S. We will associate an exclusive numeric interval $S_i = (s_{L(i)}, s_{U(i)})$ exclusively with each Q_i, where S is the union of these mutually exclusive intervals S_i. The mapping from Q_i to S_i is thus a one-to-one correspondence mapping. The numeric interval scale S is thus the union of these numeric intervals S_i, over the index set I. It should be noted that, since for each $i_1 > i_2$, that Q_{i1} is superior to Q_{i2} then $s_{L(i1)} > s_{U(i2)}$.

Each qualitative category Q_i may be further subdivided into a complete set of mutually exclusive subcategories $Q_{i,j}$, or buckets $B_{i,j}$, each with its own numeric subinterval $S_{i,j} = (s_{L(i,j)}, s_{U(i,j)})$. Again, each qualitative subcategory $Q_{i,j}$ may be further subdivided into a complete set of mutually exclusive sub-subcategories $Q_{i,j,k}$, or buckets $B_{i,j,k}$, each with its own numeric sub-subinterval $S_{i,j,k} = (s_{L(i,j,k)}, s_{U(i,j,k)})$.

Now, an object t that has been assigned by a decisionmaker to qualitative category Q_i may be assigned to exactly one qualitative subcategory $Q_{i,j}$ and exactly one qualitative sub-subcategory $Q_{i,j,k}$. Therefore, if objects t_1 and t_2 are assigned by a decisionmaker E_A to qualitative categories Q_{i1} and Q_{i2}, respectively, where $i_1 > i_2$, then s_{t1}, the numeric point estimate of the score for object t_1, is greater than s_2, the numeric point estimate of the score for object t_2. This follows from the fact that $s_{t1} \in S_{i1}$ while $s_{t2} \in S_{i2}$, and each value in S_{i2} is smaller than any value in S_{i1}. Similar relationships hold between the qualitative subcategories and the qualitative sub-subcategories.

It is also possible that two evaluators could assign object t to the same qualitative sub-subcategory, even though they disagree on the numeric point estimate of the score. Thus, any consideration of agreement between a given pair of evaluators should focus on the level

of agreement. For example, a pair of evaluators may agree on the qualitative category to which an object should be assigned, but differ as to the qualitative subcategory to which it should be assigned. Thus, they agree at one level, while they disagree at a finer level. On the other hand, if the evaluators assign the same numeric score to object t, they do not share the same mappings from Q to S, then they do not necessarily agree on the quantitative category.

Consequently, unless evaluators share equivalent sets of qualitative categories Q, numeric interval scale S, and mappings $G:Q \to S$, then it is generally meaningless to attempt to use numeric point estimates from S in order to assess agreement between decisionmakers. So how can we estimate the current level of consensus in the group? We suggest that qualitative evaluations which have "natural" meaning to the evaluators are better indicators of group consensus than numeric point estimates. In Appendix A, we define a set of consensus indicators that are based on the assignments to the qualitative categories, subcategories, and sub-subcategories. These indicators can be used to assess group consensus and tendencies for consensus formation with regard to any decision alternative(s) under consideration (for theoretical details see Ngwenyama *et al.*, 1993a).

The process of dividing each qualitative category at one level into multiple lower level (i.e., more detailed) subcategories may be repeated until the assignment of any two objects into any subcategory at the lowest level of decomposition is equivalent to assigning them the same VRN. But, how exactly could one go from a qualitative category to a VRN or even a numeric point estimate. The method of pairwise comparisons could be useful here. Pairwise comparison methods involve the comparison of each pair of objects (say t and r) in T, the set of objects. Pairwise comparisons methods are most effective when the number of objects involved is small, the variation in scores for the relevant objects involved is not wide, and the decisionmaker has a good feel for the relationship between the objects. The qualitative discriminant process (QDP) results in the assignment of objects to qualitative sub-subcategories that would, in general, satisfy these conditions. Pairwise comparisons in QDP are restricted to objects within the same sub-subcategory. The objective is to determine the ordinal relationships of the objects within the given sub-subcategory.

3.3 *Transforming qualitative assignments to numeric scores*

The intervals $S_i = (s_{L(i)}, s_{u(i)})$ associated with each qualitative category do not have to be of equal length, but for a given i_1 the subintervals $S_{i1,j}$ are of equal length, and the sub-subintervals $S_{i1,j,k}$ are of equal length. This follows from the fact that each $Q_{i1,j}$ resulted from the implicit trisection of Q_{i1} into the three subcategories, top, middle, and bottom, and each $Q_{i1,j,k}$ resulted from the implicit bisection of $Q_{i1,j}$ into the two sub-subcategories, upper and lower. Thus, if we specify values for each S_i, then we implicitly also specify the corresponding $S_{i,j}$ and $S_{i,j,k}$. This implies that if object t_1 was assigned to the qualitative sub-subcategory $Q_{i,j,k}$ by evaluator r, then x_{t1}^r, the VRN that is the score for object t_1, is included in the numeric sub-subinterval $S_{i,j,k}$. Now, since any numeric point estimate (say s_{t1}^r) of the score of object t_1 is included in the VRN x_{t1}^r, then s_{t1}^r satisfies the following relation:

$$s_{L(i,j,k)} \leq s_{t1}^r \leq s_{U(i,j,k)} \tag{3.1}$$

Now, if object t_2 were also assigned to $Q_{i,j,k}$, but were ranked as being inferior to t_1, then if s_{t2}^r is a numeric point estimate of the score of object t_2, we have:

$$s_{L(i,j,k)} \leq s_{t2}^r + p \leq s_{t1}^r \leq s_{U(i,j,k)} \tag{3.2}$$

where we associate a precision p (>0) with the numeric point estimates.

It should be noted that these two relations (3.1, 3.2) determine the VRNs that are acceptable as scores for the objects for a given evaluator. However, there does not need to be any explicit calculation of these VRNs. Our knowledge of their existence and the bounds on their values provide us with flexibility in generating legitimate individual and "consensus" numeric point estimates.

Let s_t be a "consensus" numeric point estimate of the score of object t. Ideally, s_t should be as close as possible to the numeric point estimates for object t that were provided by the individual decisionmakers. Thus, for R, the index set of the decisionmakers, we want $\sum_{(r \in R)} |s_t^r - s_t|$ to be as small as possible for object t. Since this concern would apply to all objects in T, our objective is to minimize $\sum_{(t \in T)} \sum_{(r \in R)} |s_t^r - s_t|$. This problem can be solved as a linear programming (LP) problem. The solution of this problem will provide the following:

1. A numeric point estimate vector for the objects to each decisionmaker, such that the numeric point estimates for the objects will be consistent with the evaluator's assignment and ordering of the objects.
2. A compromise numeric point estimates vector for the objects.

We suggest that these mathematical methods for deriving numeric point estimates of "consensus" should be applied after an acceptable degree of group consensus via qualitative categorization. In the event that there are "outliers" associated with the score of an object t then these could be dropped. Such a course of action would only be reasonable if the degree of consensus was sufficient but there was not unanimous agreement. In such a case, if R_{t^*} is the index subset for the decisionmakers who have a high degree of agreement on the score for t, we may use $\sum_{(r \in R_{t^*})} |s_t^r - s_t|$ instead of $\sum_{(r \in R_t)} |s_t^r - s_t|$ in the objective function of our LP problem.

The LP formulation specifies the assignment of each object to a narrow level bucket $(B_{i,j,k})$, the ordering of the objects in each of these buckets, the specification of a numeric scale $S = (s_L, s_U)$, and the consistent lengths of the numeric intervals (i.e., $D_i = s_{u(i)} - s_{L(i)}$) for each qualitative category. In cases where the user is not sure of a precise value for each D_i, a VRN should be provided. Let $(D_{L(i)}, D_{U(i)})$ be such a VRN. Then our LP could accommodate this vagueness in the specifications by including the following type of constraint:

$$D_{L(i)} \le D_i \le D_{U(i)}$$

The formulation and solution of the LP in a computer implementation of the QDP could be done in a manner that is transparent to the user, in a manner similar to the formulation and solution of the eigenvector method in the analytic hierarchy process (AHP), or the least squares method in statistical regression. It should be noted that the sizes of LP problems are generally quite small, and can be quickly solved by current commercial LP software.

4. DESCRIPTION OF THE GENERAL PROCEDURE

A tree structure (see Fig. 2), of qualitative categories is used to discriminate among the objects. Let B be the bucket that contains the set of all objects before ranking, and B_i, B_{ij}, and B_{ijk} be the buckets for objects assigned to Q_i, Q_{ij}, and Q_{ijk}, respectively. Then, B is the root of the tree and applies to the set of objects, while B_i, B_{ij}, and B_{ijk}, represent successive levels. Let B^* be the bucket that contains the ordered set of all objects after ranking. Although we have defined only three levels, the user is free to define any number of levels and subdivision deemed appropriate of the application at hand. The levels i, j, and k are, respectively, defined as *broad*, *intermediate*, and *narrow*.

Assumptions:

1. All decisionmakers will use the same set of qualitative categories Q.
2. Each qualitative category Q_i in Q is meaningful to each decisionmaker for the purpose of evaluating the objects.
3. All decisionmakers will use the same interval scale S. No attempt will be made to identify the interval, subintervals, or sub-subintervals of S that could be asso-
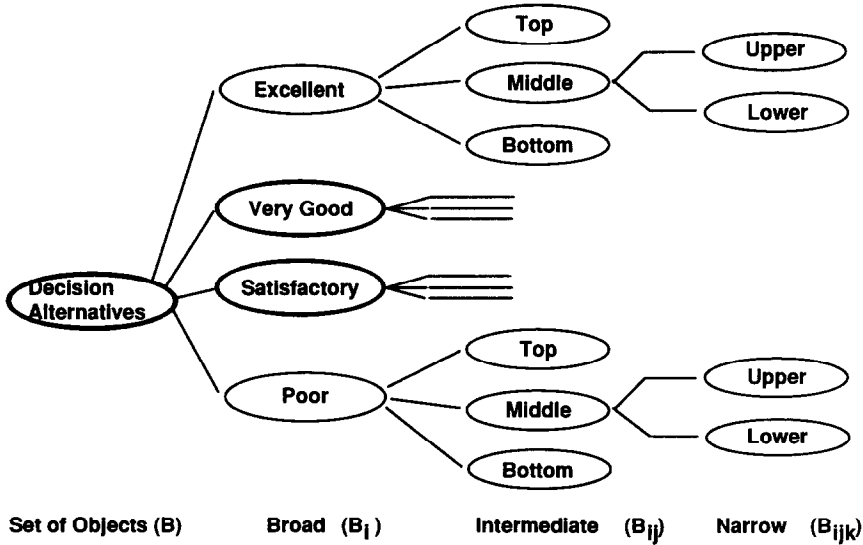
Fig. 2. Tree diagram for categorization process.

ciated with each Q_i until all objects have been assigned to their qualitative sub-subcategory.

4. All decisionmakers (i.e., evaluators) use the same explicitly specified mapping from the qualitative categories Q_i to the numeric intervals S_i.

5. If two decisionmakers assign an object to the same qualitative category Q_i, then they share a broad (but not necessarily deep) level of agreement on the value of that object with respect to the common basis of evaluation. If two decisionmakers assign an object to the same qualitative subcategory Q_{ij}, then they share a deeper level of agreement than in the previous case. If two decisionmakers assign an object to the same qualitative sub-subcategory $Q_{i,j,k}$, then they share an even deeper level of agreement than in the previous case.

The following stages outline the procedure for differentiating among the objects.

4.1 *Individual procedure for ranking*

Stage 1
   Assign each object t, $t \in T$, to an appropriate B_i.

Stage 2
   For each $i \in I$
      If B_i is not empty,
         Assign each object "t" in B_i to an appropriate $B_{i,j}$.

Stage 3
   For each $i \in I$
      For each $j \in J$
         if $B_{i,j}$ is not empty,
            Assign each object "t" in $B_{i,j}$ to an appropriate $B_{i,j,k}$.

Stage 4
   For each $i \in I$
      For each $j \in J$
         For each $k \in K$
            If $|B_{i,j,k}| \geq 2$ then,
               Order the objects in $B_{i,j,k}$ using pairwise comparisons, and place the objects in B^* based on this ordering.

4.2 Group procedure

There are four steps in the group procedure: (1) the specification of the numeric interval scale; (2) the execution of the individual procedure by each decisionmaker; and (3) the aggregation of the individual preferences into a group preference. The following is an outline of the group procedure:

Step 1: Scale definition
Specify the numeric interval scale S, and the lengths of the numeric intervals for each qualitative category (i.e., D_i, $i \in I$).

Step 2: Individual ranking
For each decisionmaker "r," $r \in R$
Perform the individual procedure.

Step 3: Aggregating group preference
(a) For each object "t," $t \in T$
The facilitator should determine if the degree of consensus is sufficient.
(b) For those objects for which the degree of consensus is not sufficient the facilitator should inform the group and request that decisionmakers repeat step 2 if they feel that another round would improve the outcome, or the facilitator should go to step 4.

Step 4: Generating point estimates
(a) The facilitator generates LP formulation and solution.
(b) Present the consistency vector to the group and present each decisionmaker with his/her vector of scores.

5. APPROACHES TO USING THE QDP

A group decisionmaking problem may involve the evaluation of the candidate alternatives in terms of a single criterion or multiple criteria. In addition, the problem may involve a single stage or multiple stages. A particular stage may involve one criterion or multiple criteria. The QDP is relevant for all of these situations. We will provide descriptions of the single criterion and multiple criteria situations.

Single criterion problem

In this problem situation, the individual and group procedures of the QDP are applied in order to generate individual and group scores for the alternatives. An example of such a problem is the popular sequencing of managerial activities (McLeod & Liker, 1992; Turoff & Hiltz, 1982). The group receives a list of management activities and instructions to reach consensus on the proper sequence of activities.

Multiple criteria problem

Multiple criteria decisionmaking is another general problem area to which QDP can be applied. In practice explicit or implicit synthesis approaches are to be used to determine *composite scores* for the alternatives. Both approaches involve the hierarchical structuring of the problem into *criteria level* and *alternatives level*, but they differ in terms of rigor. The explicit synthesis approach utilizes a formal weighting model as follows. First, the decisionmakers determine a score/weight for each criterion. Second, they determine a score/weight for each alternative with respect to each criterion. Then a composite score is computed for each alternative using an *additive value function*. Finally, the alternatives are ordered based on their composite scores. This *weighting model* approach is sometimes used by groups to evaluate projects that are to be funded. In these types of situations, the QDP can be used to generate the scores or weights for the criteria, as well as for the alternatives.

An example of the application of the weighting model is the TELOS/PDM procedure (Burch, 1992). This procedure is used by information systems steering committees for project evaluation and selection. Each proposed project is evaluated based on two sets of criteria: (a) TELOS feasibility factors, technical, scheduling, etc.; and (b) PDM strategic change factors, productivity, differentiation, and management. The score for each set of

criteria is generally determined by averaging the scores of the constituent criterion. How-
ever, it could also be calculated using a weighted sum of the scores with regard to the rel-
evant criteria. For the TELOS/PDM procedure, the QDP can be used to determine the
score for each alternative with respect to each set of criteria, and also the criteria weights.

The implicit synthesis approach to the multiple criteria decisionmaking problem does
not follow a formal model. First, the decisionmakers decide on the criteria upon which the
alternatives will be judged and then a scale is selected. No weighting model is developed
for the criteria. Each decisionmaker determines a score for each alternative with respect
to each criterion. Using these criteria scores the decisionmaker then determines a *compos-
ite score* for each alternative. In this approach, the QDP would be used to generate the
group and individual scores for all the alternatives with respect to each criterion, and may
also be used to generate the composite scores for the alternatives.

The implicit synthesis approach is sometimes used in the selection of personnel by a
search committee. For example, in the selection of a dean for a school, the search com-
mittee may decide that the important criteria are academic credentials, management cre-
dentials, and fundraising credentials. The committee members may decide not to use the
weight model for a number of reasons including lack of confidence in the model, or strong
disagreement on the importance of the different criteria. It should be noted that the QDP
can be used to determine the level of agreement in the latter situation. If the committee
members do not want to use any formal synthesis technique, but want to rate the candi-
dates with reference to each criterion, then the QDP can be used to do this at both the indi-
vidual and group levels. Other applications of the implicit approach include the ranking
of sports teams by a panel of sports pundits.

6. CASE ILLUSTRATION

The prototype implementation of the QDP is PC-based with screen-interactive data
capture and back-end data analysis. Figure 3 depicts the conceptual architecture of the sys-
tem. Each decisionmaker using the QDP support system is guided through a hierarchy of
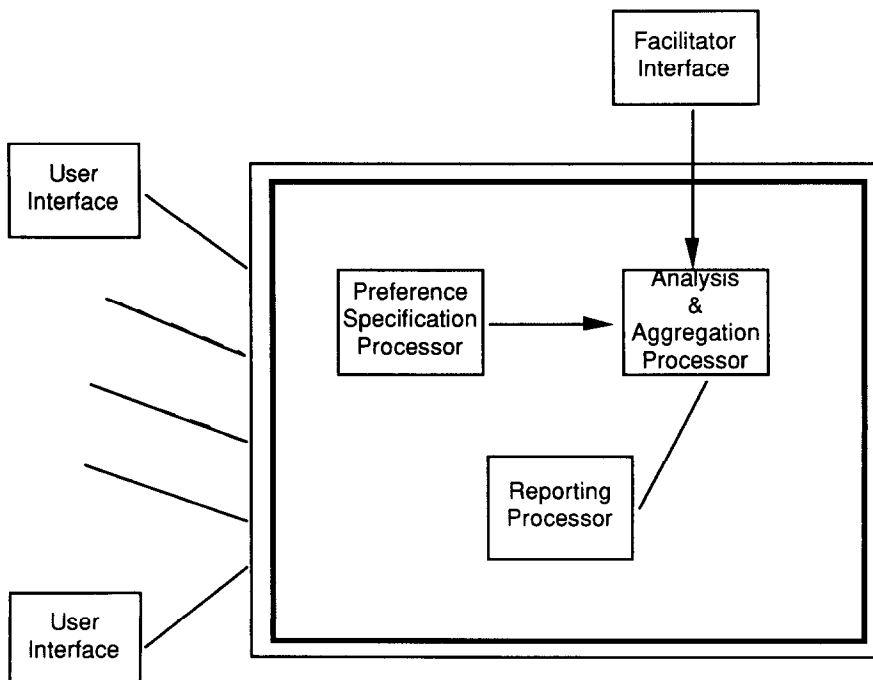

Fig. 3. Conceptual architecture of QDP support system.

three mouse-driven interactive user interface screens (Figs. 4-6), which are supported by the preference specification processor (PSP).

Ranking and scoring with the QDP system is a three-step process. In step 1, the decisionmaker performs three rounds of discriminant analysis, in which the decision alternatives are moved from broad categories on the left of the screen to successively narrower categories on the right of the screen. At the end of this step the PSP produces a file of the rankings which serves as input to the analysis & aggregation processor (AAP). In step 2, the facilitator invokes the AAP which: (a) generates the VRN scores for each set of rankings, and (b) formulates and solves the LP (see example Appendix A), which generates individual score vectors, a "consensus" score vector, and the group ranking. The results are then reported to each decisionmaker for group discussion. If there are disagreements about individual scores, decisionmakers can re-do their individual rankings and the process can continue from there until there is agreement with the results. The following is a case illustration which demonstrates the process.

6.1 *Case description*

A search committee consisting of the faculty representative, the alumni/student representative, and the corporate partners representative has been mandated to identify suitable candidates for the position of Dean of the School of Business, and to present a list of the top three candidates to the university president. The committee decided that the three criteria to be used in evaluating the candidates are academic credentials, management credentials, and fundraising credentials. They also decided to use a multistage decisionmaking process with the first stage involving an initial screening of the candidates, and that the only candidates who will be considered in the later stages are those who the group agreed had excellent academic credentials. Our illustrative example applies to this first stage. It should be noted, however, that in a multistage decisionmaking process the QDP can be applied at each stage with the output from one stage being input to the succeeding stage. Also, any stage can involve a single criterion or multiple criteria.

Fig. 4. User-interface screen 1.

Qualitative Discriminant Ranking Procedure		
Second Round Categorization		
Decision Alternatives In Broad: Excellent	**Intermediate Sub-Categories**	
Thomas Steele William Washington August St. James	**Top**	≡≡≡≡
	Middle	≡≡≡≡
	Bottom	≡≡≡≡

Select Broad Category you wish to rank:
1. Excellent; 2. Very Good; 3. Satisfactory; 4. Poor
To rank alternatives: Block out item, move cursor to Sub-Category and double click.

Press Enter for next screen.

Fig. 5. User-inferface screen 2.

6.2 *Application of the QDP*

As stated previously, the QDP system guides decisionmakers through a three-round analysis process. In **round 1**, the decisionmakers are each presented with the list of the candidates on the right of the screen, and are required to place each candidate in one of the four broad categories, *excellent*, *very good*, *satisfactory*, and *poor*, on the left (see Fig. 4).

Qualitative Discriminant Ranking Procedure		
Third Round Categorization		
Decision Alternative Intermediate: (Very Good). Top	**Narrow Sub-Categories**	
Jens Hilderbrand Deablo Jesus Thomas Isikawa Goef Codogan	**Upper**	**Jens Hildebrand**
	Lower	
		Goef Codogan

Select Intermediate Sub-Category:
1. Top; 2. Middle; 3. Bottom
To rank alternatives:
1. Use mouse to block item
2. Move cursor to appropriate Narrow Sub-Category and double click.

Fig. 6. User-interface screen 3.

The reader may recall that the QDP permits the group to redefine these four broad qualitative categories depending on the context of the valuation (e.g., for the ranking of a set of criteria, the broad categories *very high, high, average, low* might be appropriate). The process is straightforward, and yields a preliminary ranking of the alternatives. The decisionmakers may also shift the candidates among any of the categories until they feel comfortable with their assignments.

Round 2 is another cycle of the ranking process which yields a higher level of differentiation of the alternatives. In this round, the decisionmakers are required to further distinguish among the candidates in each broad category, by differentiating them into the intermediate subcategories. *top, middle*, and *bottom* (see Fig. 5). Again, it is possible to shift around the candidates until the desired assignments are achieved. For example, the decisionmaker may decide to move a candidate from *very good* to *excellent/bottom*, or *excellent/bottom* to *very good/top*. In fact, the second and successive round forces the decisionmaker to rethink the rankings of each category under consideration in order to achieve a higher degree of differentiation among the candidates.

In **round 3**, the third and final round of the ranking process (see Fig. 6), each decisionmaker is required to differentiate among each set of candidates in the intermediate subcategories, and place them into the relevant *narrow* subcategories, *upper* and *lower*.

This stage of the process allows for the further refinement of the assignments. It should be noted that for each of the four *broad* categories, there are six associated *narrow* subcategories which give the decisionmaker a wide range of options for ranking the candidates. For each *narrow* sub-subcategory, if there is more than one candidate, the pairwise comparison technique is used to determine the rank order within that *narrow* sub-subcategory. Figure 7 displays the assignments of the candidates to the sub-subcategories, and the rankings within each sub-subcategory. The symbol "≫" should be interpreted as meaning "superior to."

6.3 *Analysis and aggregation*

The procedure for assigning a score is just as simple. In the current implementation we use a scale of 1 to 100 divided into four intervals, and assign a range to each broad category (*excellent* [80-100], *very good* [60-80], *satisfactory* [30-60], *poor* [0-30]). It should be noted that the group could decide to use any other subdivision of this scale.

(i, j, k)	Faculty Representative (A)	Alumni/Student Representative (B)	Corporate Partners Representative (C)
(4, 3, 2)			
(4, 3, 1)			Washington
(4, 2, 2)	Steele ≫ Washington	Steele	
(4, 2, 1)		Washington	Steele
(4, 1, 2)	St. James		
(4, 1, 1)		Hildebrand	St. James
(3, 3, 2)		St. James	
(3, 3, 1)	Hildebrand ≫ Jesus ≫ Isikawa ≫ Cadogan	Isikawa	Jesus ≫ Isikawa
(3, 2, 2)	McIntire	Jesus	Cadogan
(3, 2, 1)	Obadio	Cadogan ≫ McIntire	Obadio ≫ Hiawata ≫ McIntire
(3, 1, 2)	Hiawata	Hiawata ≫ Obadio	Hildebrand
(3, 1, 1)			
(2, 3, 2)			
(2, 3, 1)			
(2, 2, 2)			
(2, 2, 1)			
(2, 1, 2)			
(2, 1, 1)			

Fig. 7. Individual rankings of the candidates.

When the individual ranking is completed the facilitator invokes the analysis & aggregation processor (AAP) which formulates and solves the LP. The output includes individual and group scores for each alternative, and a group ranking of the candidates. Figure 8 lists the estimated scores for each decisionmaker, the resulting aggregate (consensus) scores, and group rank. As can be seen from the results, there is no major disagreement on the rankings among these decisionmakers. It is important to note, however, that two candidates, Steele and Washington, tied for first place in the aggregated ranking, and that while St. James was assigned to the *very good* category by the alumni/students representative, his group rating was in the *excellent* category.

7. CONCLUSIONS

We have presented a qualitative discriminant process (QDP), a simple and intuitively appealing technique for scoring and ranking in GSS that is relevant to group decision-making situations that involve the evaluation of alternatives. This approach is based on well-recognized decision theory techniques and is easy to implement in any computing environment with LP software. (The prototype discussed in this article was developed for IBM-compatible workstations using VISUAL BASIC™. We are currently developing a version for the Macintosh environment.)

The QDP addresses four major limitations of current systems. These are: (1) the inability to deal with vagueness in human decisionmaking; (2) difficulties in mapping qualitative evaluation to numeric estimates; (3) problems in aggregating individual preferences into meaningful group preference; and (4) the inability to deal with a large number of decision alternatives.

We have taken the position that it is more prudent to first differentiate among the objects at increasing levels of detail using qualitative categories that have "natural" meaning to the evaluators, before attempting to assign numeric point estimates as the scores of the objects. An important benefit of this approach is that information relevant to consensus building can be obtained even before numeric values have been assigned. The QDP also provides more structure to the task of scoring and ranking than exists in current GSS. This approach is supported by the results of GSS research. For as noted by McLeod and Liker (1992): "Providing specific rules for and resources for task interaction . . . can lead to improved group performance among a number of dimensions, such as, task quality, and conflict resolution." We submit that the QDP would lead to improvements in the quality of scoring and ranking, and in consensus building and assessment.

CANDIDATE	FACULTY REP.	ALUMNI/ STUDENT REP.	CORPORATE PARTNERS REP.	CONSENSUS	GROUP RANK
Steele	90.30	90.10	89.90	90.10	1*
Washington	90.10	89.90	93.43	90.10	1*
St. James	83.43	79.90	83.23	83.23	3
Hildebrand	76.57	80.10	66.57	75.57	4
Jesus	73.83	73.23	73.83	73.83	5
Isikawa	73.63	73.63	73.63	73.63	6
Cadogan	73.43	69.90	73.23	73.23	7
McIntire	70.10	66.77	66.77	66.77	9
Obadio	67.17	66.37	67.17	67.17	8
Hiawata	66.57	66.57	66.97	66.57	10

The prototype discussed in this paper was developed and tested for IBM compatible workstations using VISUAL BASIC™. We are currently developing a version the MACINTOSH environment.

Fig. 8. Resulting scores and group rank.

The QDP was developed to support group decisionmaking tasks which require that numeric point estimates be provided for objects based on criteria whose natural domains are qualitative or subjective. Such situations are common. They include: (1) the derivation of a national ranking of college football teams by sports "experts"; (2) the selection of personnel by search committees; (3) the sequencing of activities; and (4) the selection of projects for funding by a review committee. The applications are numerous. Although the QDP is not the only procedure for addressing these tasks, it represents a new conceptual approach that has as its emphases: (1) accommodating vagueness; (2) avoiding the premature assignment of numeric scores; and (3) facilitating consensus building and assessment. Thus, it offers new possibilities for scoring, ranking,and classification in group situations. It must also be emphasized that QDP is well suited to decision situations in which only systematic qualitative evaluation is required. For example in classification tasks where assignment to categories is based on similarity of attributes, the broad categories can be defined appropriately, and the alternatives assigned to the appropriate sub-subcategories.

REFERENCES

Bryson, N., & Mobolurin, A. (in press). An approach to using the analytic hierarchy process for solving multiple criteria decision making problems. *European Journal of Operational Research*.

Bryson, N., Mobolurin, A., & Ngwenyama, O. (in press). Modeling pairwise comparisons on ratio scales. *European Journal of Operations Research*.

Burch, J. (1992). *Systems analysis, design and implementation*. Boston: Boyd and Fraser.

Cook, R., & Stewart, T. (1979). A comparison of seven methods for obtaining subjective descriptions of judgmental policy. *Organizational Behavior and Human Performance, 13*, 31-45.

Dennis, A., George, J., Jessup, L., Nunamaker, J., & Vogel, D. (1988). Information technology to support electronic meetings. *MIS Quarterly, 12*(4), 591-624.

Dutta, B. (19xx). Strategic voting in a probabilistic framework. *Econometrica, 41*, 578-601.

Dyer, J.S., & Sarin, R.K. (1979). Group preference aggregation rules based on strength of preference. *Management Science, 25*, 589-609.

Fedrizzi, M., & Kacprzyk, J. (1988). On measuring consensus in the setting of fuzzy preference relations. In K. Kacprzyk and M. Roubens (Eds.), *Non-conventional preference relations in decision making* (pp. 129-141). Springer-Verlag.

Finholt, T., & Sproull, L. (1990). Electronic groups at work. *Organization Science, 1*(1), 41-64.

Gabrielli, W., & Von Winterfeldt, D. (1978). *Are weights sensitive to the range of alternatives in multiattribute utility measurement?* (SSRI Research Report 78-6) University of Southern California, Los Angeles, CA.

Goddard, S.T. (1983). Ranking in tournaments and group decision making. *Management Sciences, 29*, 1384-1392.

Golden, B., Wasil, E., & Harker, P. (1989). *The analytic hierarchy process: Application and studies*. Berlin: Springer-Verlag.

Grohowski, R., McGoff, C., Vogel, D., Martz, B, & Nunamaker, J. (1990). Implementing electronic meeting systems at IBM: Lessons learned and success factors. *MIS Quarterly, 14*(4), 369-384.

Harker, P. (1987a) Incomplete pairwise comparisons in the analytic hierarchy process. *Mathematical Modelling, 9*, 837-848.

Harker, P. (1987b). Alternative modes of questioning in the analytic hierarchy process. *Mathematical Modelling, 9*, 353-360.

Harker, P. (1989c). The art and science of decision making: The analytic hierarchy process. In B. Golden, E. Wasil, and P. Harker (Eds.), *The analytic hierarchy process: Application and studies* (pp. 3-36). Berlin: Springer-Verlag.

Huber, G.P. (1974). Methods for quantifying subjective probabilities and multi-attribute utilities. *Decision Sciences, 5*, 430-458.

Huber, G.P. (1984a). The nature of post industrial organizations. *Management Science, 30*, 928-951.

Huber, G.P. (1984b). Issues in the design of group decision support systems. *MIS Quarterly, 8*(3), 195-204.

Hurrion, R.D. (1985). Implementation of a visual interactive consensus decision support system. *European Journal of Operational Research, 20*, 138-144.

Janis, I. (1982). *Victims of group think* (2nd ed). Boston: Houghton Mifflin.

Jarke, M., Jelassi, M.T., & Shakun, M.F. (1987). MEDIATOR: Towards a negotiation support system. *European Journal of Operational Research, 31*, 314-334.

Kirkwood, C.W., & Sarin, R.K. (1985). Ranking with partial information: A method and an application. *Operations Research, 33*, 38-48.

Kok, M., & Lootsma, F.A. (1985). Pairwise comparison methods in multiple objective programming. *European Journal of Operational Research, 22*, 44-55.

Korhonen, P.J. (1986). A hierarchical interactive method for ranking alternatives with multiple qualitative criteria. *European Journal of Operational Research, 24*, 265-276.

Korhonen, P. (1987). VIG: A visual interactive support system for multiple criteria decision making. *Belgian Journal of Operations Research, Statistics and Computer Science, 27,* 3–15.

Korhonen, P. (1991). Using harmonious houses for visual pairwise comparison of multiple criteria alternatives. *Decision Support Systems, 7,* 47–54.

Kull, D.J. (1982, May). Group decision: Can computers help? *Computer Decisions,* 70.

Lootsma, F.A. (1985). Performance evaluation of non-linear optimization methods via pairwise comparison and fuzzy numbers. *Mathematical Programming, 33,* 93–114.

McLeod, P. (1991). What we know, what we don't know, and what we think we know about GDSS. Unpublished proceedings of the Human Computer Action Workshop, University of Michigan, Ann Arbor, MI.

McLeod, P., & Liker, J. (1992). Electronic meeting systems: Evidence from a low structure environment. *Information Systems Research, 3*(3), 195–223.

Ngwenyama, O., Bryson, N., & Mobolurin, A. (1992). *Supporting consensus formation in group support systems.* Unpublished working paper.

Novak, V. (1990). On the syntactico-semantical completeness of first order fuzzy logic I, II. *Kybernetika,* 3–15.

Nunamaker, J., Vogel, D., Heminger, A., Martz, B., Grohowski, R., & McGroff, C. (1989). Experiences at IBM with group support systems: A field study. *Decision Support Systems, 5*(2), 183–196.

Nunamaker, J., Dennis, A., Valacich, J., Vogel, D., & George, A. (1991). Electronic meeting systems to support group work. *Communications of the ACM, 34*(7), 40–61.

Nurmi, H., Fedrizzi, M., & Kacprzyk, J. (1990). Vague notions in the theory of voting. In J. Kacprzyk and M. Fedrizzi (Eds.), *Multiperson decision making using fuzzy sets and possibility theory* (pp. 43–52). Dordrecht: Kluwer.

Parikh, R. (1983). The problem of vague predicates. In R. Cohen and M. Wartofsky (Eds.), *Language, logic, and method* (pp. 241–261). XXX: Reidel.

Saaty, T. (1980). *The analytic hierarchy process: Planning, priority setting, resource allocation.* New York: McGraw-Hill.

Saaty, T. (1986). Axiomatic foundations of the analytic hierarchy process. *Management Science, 32,* 841–855.

Saaty, T. (1989). Group decision making and the AHP. In B. Golden, E. Wasil, and P. Harker (Eds.), *The analytic hierarchy process: Application and studies* (pp. 59–67). Berlin: Springer-Verlag.

Shanteau, J. (1980). *The concept of weight in judgement and decision making: A review and some unifying proposals* (Technical Report 228) Center for Research on Judgement and Policy, University of Colorado, Boulder, CO.

Turoff, M., & Hiltz, S. (1982). Computer support for group versus individual decisions. *IEEE Transactions on Communications, 3,* 82–90.

Vetschera, R. (1991). Integrating databases and preference evaluations in group decision support. *Decision Support Systems, 7,* 67–77.

Vogel, D., & Nunamaker, J. (1990). Group decision support system impact: Multi-methodological exploration. *Information and Management, 18,* 15–20.

Watson, R., Alexander, M., Pollard, C., & Bostrom, R. (1991). *The use and adoption of OptionFinder: A keypad based group decision support system.* 3M Management Institute.

Weber, M. (1987). Decision making with incomplete information. *European Journal of Operational Research, 28,* 44–57.

Weber, M. (1985). A method of multiattribute decision making with incomplete information. *Management Sciences, 31,* 1365–1371.

Wright, G., & Ayton, P. (1987). Eliciting and modelling expert knowledge. *Decision Support Systems, 3,* 13–26.

Yager, R. (1982). *Fuzzy sets and possibility theory: Recent developments.* New York: Pergamon Press.

Zadeh, L. (1965). Fuzzy sets. *Information Control, 8,* 338–353.

APPENDIX A. CONSENSUS INDICATORS

Let α ($50 \leq \alpha \leq 100$) be the parameter that is used as a threshold value for testing for group consensus on an issue. Let PER_{i}^{t}, $\text{PER}_{i,j}^{t}$, and $\text{PER}_{i,j,k}^{t}$ be the percentages of group members that assign object "t" to Q_i, $Q_{i,j}$, and $Q_{i,j,k}$, respectively. Also, let:

$$C_3^t = \text{Max}\{(\text{PER}_{i,j,k}^t/\alpha) : i \in I, j \in J, k \in K\};$$

$$C_2^t = \text{Max}\{(\text{PER}_{i,j}^t/\alpha) : i \in I, j \in J\};$$

$$C_{2+}^t = \text{Max}\{(\text{PER}_{i,j}^t + \text{PER}_{i,(j-1),kU}^t + \text{PER}_{i,(j+1),kL}^t)/\alpha : i \in I, j \in J\}$$

where k_U and k_L are the highest and lowest sub-subcategories associated with subcategories $(j - 1)$ and $(j + 1)$, respectively. We also define the following consensus indicators:

$$C_1^t = \text{Max}\{(\text{PER}_{i,j}^t/\alpha) : i \in I, j \in J\};$$

$$C_{1+}^t = \text{Max}\{(\text{PER}_i^t + \text{PER}_{(i-1),jU}^t + \text{PER}_{(i+1),jL}^t)/\alpha) : i \in I\}$$

where j_U and j_L are the highest and lowest subcategories associated with categories $(i - 1)$ and $(i + 1)$, respectively.

It is clear that $C_{1+}^t \geq C_1^t \geq C_{2+}^t \geq C_2^t \geq C_3^t$. The order of these indicators represent increasing levels of consensus. By definition, $C_3^t \geq 1$ indicates that the group has reached an acceptable level (i.e., α-degree) of consensus with regard to the assignment of object t to a particular qualitative sub-subcategory. On the other hand, if $C_3^t < 1$ and $C_2^t \geq 1$, then, although the group does not have a α-degree consensus with regard to the qualitative sub-subcategory to which t should be assigned, there is a α-degree of consensus on the qualitative subcategory to which t should be assigned. The facilitator thus has multiple indicators to both inform him/her of the current situation with regard to consensus and of the tendency for consensus formation.