

AN EMPIRICAL COMPARISON OF TWO CHOICE AGGREGATION METHODS IN THE ANALYTIC HIERARCHY PROCESS

Sibs von Solms

Tel: +27 35 902-2214; Fax: +27 35 902-2229

Email: sibs_von_solms@mhs21.tns.co.za

Kirti Peniwati

PPM School of Management

Jl. Menteng Raya 9, Jakarta 10340, Indonesia

Email: kirti@cabi.net.id

Keywords: Analytic Hierarchy Process, Domination, Group Dynamics, Multi-criteria Decision-Making, Consensus, Geometric Mean.

Summary: *This paper builds on the theoretical discussion presented in an earlier paper (this issue) on the importance of and several problems related to participative group decision-making. A variation of how the Analytic Hierarchy Process (AHP) can be applied to address the problems and aid participative decision-making is suggested. An empirical study in which the suggested method was tested is discussed.*

1. INTRODUCTION

In the earlier paper, the importance of participative decision-making was discussed. The value of involving a larger amount of employees in decision-making was argued along with the suggestion that this involvement should not be limited to decisions of a 'local' or minor nature, but also bigger, strategic decisions should ideally be based on a participative approach. Both the value of and the problems associated with this type of approach were discussed from a theoretical perspective. This paper wishes to carry the discussion forward by (i) Highlighting an additional problem, (ii) Suggesting a possible variation of the AHP approach as an attempt to practically address the problems, (iii) Presenting an empirical study in which the suggested methodology was used and (iv) Discussing the study outcomes and possible future research.

2. PROBLEM FORMULATION

Accepting the theoretical reasons for adopting a more participative approach to decision-making, results in being confronted with the inevitable problems associated with the involvement of larger numbers of participants in the decision-making process. Collective choices will always be complicated by conflicting values and the notion of justice demands that the preferences of all stakeholders be treated equal [Davos, 1989]. Clearly, we need to confront the problems and search for solutions. The main problem discussed in the earlier paper was that in only the simplest decision-making situations and cases of unitary contexts do the use of debated consensus make sense. In all circumstances where perceptions, values, seniority, power and interests differ among the group members, the possibility of domination or coercion exists and consensus could increasingly be biased towards the member (or minority) with 'power' in whatever form. Not only is this possibility a reality in most organizational decision-making situations, but it increases in magnitude when, as was argued in the earlier paper, larger numbers of employees, from various seniority levels, participate in the process. Investigating an alternative way of preference aggregation was suggested.

The effects groups have on their members were not addressed in the earlier paper, mainly due to reasons of length. It is, however, clear from the group dynamics literature that the mere fact of being part of a group affects the way in which individuals react and the way in which they make decisions. The literature, both theoretical and empirical, on this subject is vast and only a very brief discussion can be attempted here.

The fact of the influence of groups on the actions of individuals was established empirically as early as the end of the nineteenth century and work done during the first two decades of the twentieth century confirmed this. The extent and results of the influence, under different circumstances, formed an important part of group dynamic research during the mid 1900s. These factors included the effectiveness of different groups in problem solving [Shaw, 1932], [Taylor & Faust, 1952]. The results from this line of research are summarized in the findings of Marquart [1955] that, in the majority of cases, groups performed better at problem solving than individuals, but that the group performance rarely exceeded that of the best group member, suggesting that the group performance is based on the presence of one superior individual. Problem solving in groups, however, is not a simple *groups-are-better* situation. The quantity and quality of a group's performance can be influenced by, what is called 'socio-emotional' factors. These factors often distract the attention of the group members away from the group task. In work done by Bales [1952], was found that group members divided their time between actions directed towards achieving group goals (57 %) on one hand and socio-emotional goals (43 %) on the other.

A particular factor influencing group effectiveness is that of group size. The size influences several aspects of group behavior including the amount of participation in the group process, the satisfaction experienced by individual group members with their contribution and the effectiveness of groups achieving problem solving and other task goals. One aspect influenced by group size, directly relevant to the current discussion, relates to consensus. An important finding regarding group size is that the smaller the group the higher its influence on the group's members and the higher the degree of consensus achieved by the group's members. This was demonstrated in a classical experiment involving 150 Boy Scouts [Hare, 1952]. The work of Hare, and others following him, stresses the *advantages* of a greater degree of consensus. But, as discussed in the earlier paper, this does not necessarily follow, as consensus can often represent the view of a minority 'forced' on the group through a process of power rather than through reasoned debate.

The work of Hare [1952] and Asch [1958] (inter alia) highlighted a group decision-making problem of serious proportions, namely that of group pressure via group norms. In the first paper, the problem of domination was discussed, showing that individuals within a group can use (misuse) their seniority or

perceived superior knowledge to dominate the group and control the outcome to predominantly reflect their own views. The second aspect, discussed here, is that the group structure exerts an influence of its own, over and above any domination based on seniority, personality or knowledge. It was graphically illustrated by the experiments of Asch [1958] where subjects had to compare the length of 3 lines to that of a test line to choose the line matching the test line. Asch [1958] placed a naïve subject in a group of research assistants who were instructed (unknown to the subject) to deliberately select the wrong line. Faced with a unanimous majority of wrong replies, 37 % of the subjects followed the (wrong) group choice, even in cases where the difference between the test line and the chosen line was between 3.5 and 5 cm. The amount of group pressure influence increased in more unclear situations (ie where the difference in line lengths were < 2 cm), and in groups with a high degree of cohesion or where the subject perceived his belonging to the group would be beneficial to him. Group norms are further often internalized resulting in an even greater degree of conformance, by group members, to these norms and hence, group pressures.

3. SUGGESTED SOLUTION

A possible solution to the problems of domination and group pressures distorting outcomes could be to use AHP in a way where all group members rate the criteria or alternatives of the hierarchy individually without discussing or divulging their choices. Through confidential ratings the influence of domination and group pressures can be eliminated. It is, however, not advisable because of the loss of both the bad as well as the good influences inherent in group discussions. Eliminating the harmful effects of domination and group pressures would not represent an improvement if it were offset by the loss of those aspects of group influence causing group performance to exceed individual performance [Shaw, 1932]. A particularly strong argument for the value of deliberation to overcome social choice aggregation problems, including the conditions set by Arrow [1970], is presented by Dryzek & List [2000]. Further, the value of both dissent *and* consensus in strategic decision-making is argued by Dooley & Fryxell [1999] and it seems important that whatever decision-making approach is suggested should be able to allow for positive group influences, deliberation and a balance between the two seemingly opposing aspects of dissent and consensus.

A variation of the application of AHP is suggested here. The variation implies that the group does three ratings on each pair in the hierarchy. First, the group members perform a full set of ratings *individually* without conferring with the other participants. In the second phase the group discusses these individual ratings and comes to a *consensus* rating, for the group, recorded by each group member. In the last phase each member again provides pairwise ratings *individually* (and anonymously), having both his first individual rating and the group consensus rating available to 'guide' his final preference. The principle is that this third phase incorporates - in the final individual ratings - both a protection against domination (being confidential) and openness to group and deliberation influences. Each participant is open to accept or reject any information, or view, expressed during group discussion and to express this view individually, without pressure, afterwards. In summary, this variation attempts to address all requirements presented by (i) allowing positive group dynamic influences, (ii) providing deliberation and debate, (iii) providing a mechanism to balance dissent and consensus while (iv) maintaining a marked degree of autonomy.

4. EXPERIMENTAL DESIGN

An empirical study was undertaken to test the feasibility of the 3-phase AHP approach suggested above. Thirty volunteers participated in 7 groups of 4 or 5 members each. The group members were arranged in a way to ensure a mix of seniority levels. One team consisted of members of 'equal' seniority while the other teams had a senior member who out ranked the rest of the group by varying degrees of seniority, from 1

managerial level to a maximum of 5 managerial levels. This was done to create diverging dynamics within the groups and to ensure varying degrees of ‘domination’. Each group was asked to perform pairwise comparisons on 2 sets of items. In the first set, 7 geometric figures (**Appendix 1**) were displayed and the participants were asked to estimate the areas of the figures using pairwise comparisons. In the second set a hierarchy with 2 levels were used (**Appendix 2**). 7 Criteria were suggested as necessary to determine the significance of a Safety, Health and Environmental (SHE) incident. The participants were asked to determine the relative importance of these 7 criteria in determining significance. The two sets were used to test group outcomes in both tangible (Figure Areas) as well as intangible (Risk Factors) rating areas. In both cases the 3-phase approach was used implying that members first rated the 7 elements individually, then the elements were discussed to obtain group consensus ratings. In the last phase each participant re-evaluated each element and again provided individual ratings. Each participant recorded his ratings by using a questionnaire-like form. Responses were subsequently processed using TeamEC™, the resulting preferences exported to MS Excel™ worksheets and processed further. Each participant also provided, via questionnaire, information on the degree of domination experienced and the amount of group participation, using a 5-point Likert-type Scale. Subsequent to the calculation of the results, 21 of the participants were interviewed using a semi-formal interview approach. The objective of these interviews was to obtain a fuller and more personal view of the participant’s experiences during the exercise and their reasons for agreeing or disagreeing with the group’s consensus ratings.

5. EXPERIMENTAL RESULTS

The first problem was to find a way by which any two sets of preferences could be compared. The author was, at this time, not aware of the Compatibility Index [Saaty, 1996] and, hence, developed a comparison based on the mean square difference. For the 7 elements in 2 sets P_1 and P_2 , p_{1i} and p_{2i} respectively ($i = 1$ to 7), squared differences (**d**) were calculated as:

$$d_i = (p_{1i} - p_{2i})^2 \quad (i = 1 \text{ to } 7)$$

And the Comparison Index (**C**) was calculated as:

$$C = \sqrt{\overline{d_i}} \cdot 100$$

After becoming aware of the Compatibility Index, the Comparison Index was compared to the Compatibility Index for 30 preference sets, spanning the whole range of Comparison Indexes found. A linear correlation, with correlation coefficient $r = 0.908$, was found. This is highly significant as an r-value of only 0.572 already represents a 99.9 % confidence that the relationship found was not due to chance [Sutton, 2000]. Due to the ease of calculating the Comparison Index in MS Excel and the fact that the original Matrixes cannot be exported from TeamEC and would have had to be manually re-entered to achieve the calculation of Compatibility (a total of 163 matrixes), the Comparison Index was used.

Many evaluations and comparisons of the ratings were performed, but we will concentrate here on the results relevant to the stated problems of domination and group pressures. To identify the results the following system is used:

- * Group: $g = 1$ to 7
- * Group Member: $m = 1$ to 4 or 5 (as applicable)
- * Group Consensus Ratings: C_g
- * Geometric Means of Individual Ratings before Discussion: I_{1g}
- * Geometric Means of Individual Ratings after Discussion: I_{2g}

- * Comparison Index between Individual Ratings before Discussion and the Group Consensus: δ_{1gm}
- * Comparison Index between Individual Ratings after Discussion and the Group Consensus: δ_{2gm}
- * Comparison Index between Geometric Means and Group Consensus: Δ_{1g} and Δ_{2g} respectively

The results are graphically shown in **Figures 1 and 2**:

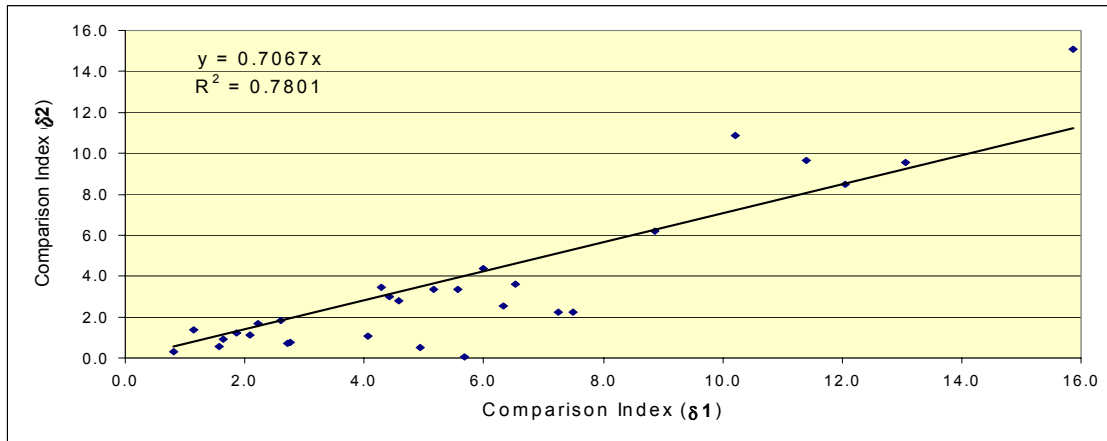


FIGURE 1: COMPARISON OF δ_{1gm} & δ_{2gm} FOR THE FIGURE AREAS

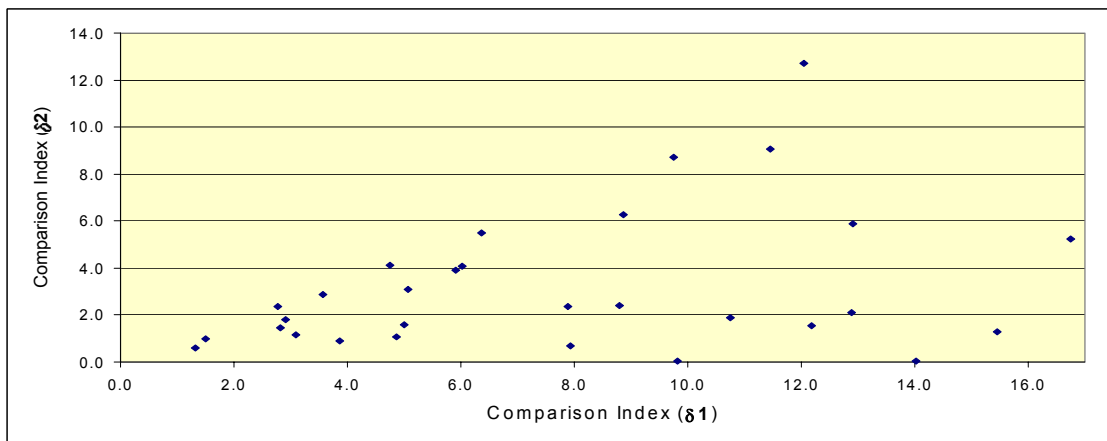


FIGURE 2: COMPARISON OF δ_{1gm} & δ_{2gm} FOR THE RISK FACTORS

6. DISCUSSION OF RESULTS

The main problem investigated (as discussed) above is that of eliminating the influence of domination and group pressures while maintaining the advantages of debate. The theoretical basis for using the 3-phase rating procedure was that it would allow – in the final individual ratings – the participants to express preference without domination due to it being done *confidentially* while simultaneously allowing for group influence because it follows the group discussion. The value of the first individual ratings was expected to be two-fold. First, it would provide each participant with a reasoned rating to contribute to the discussion and second, it could serve as a ‘counter point’ to use when weighing the group rating before making his final individual rating.

Both the stated aims are confirmed by the data. The comparison of δ_{1gm} to δ_{2gm} (**Figures 1 & 2**) shows this. The positive influence of the discussions is evident in the consistently lower δ_{2gm} values as compared to the corresponding δ_{1gm} values. This is very evident in both the case of the Figure Areas as well as the Risk Factors, with the paired differences showing significant differences (δ_{2gm} lower) at 99.95 %. Clearly, the majority of participants allowed the group consensus ratings to influence their final ratings.

There is evidence that the final ratings also incorporated a degree of *resistance* to the group view with a final rating often reverting back close to the original (first) individual preference. The results for the Figure Areas (**Figure 1**) show this clearly with a significant linear relationship between the δ_{1gm} and δ_{2gm} results. In the case of the Risk Factors the data is not so clear, with some participants showing more ‘resistance’ while other conformed to a great degree with the consensus ratings. This is in line with the findings of Asch [1958] that more subjective or intangible issues result in higher conformance to group views. The raters, clearly, are less sure of their own views and hence adopt the group view more readily. The interviews confirmed this with many participants saying that ‘they felt unsure of themselves’ in evaluating the Risk Factors. The resistance to the group view can be confirmed by presenting the results in a slightly different way. If the relative degree of conformance (**R**) is calculated as: $R = (\delta_{1gm} - \delta_{2gm})/\delta_{2gm}$ thus relating the amount of change in view from *before* to *after* deliberation to the difference recorded after debate. The results are shown in **Table 1** and the high degree of ‘dissent’ confirms the fact that resistance to the group view could be, and was, expressed. It also illustrates the higher degree of conformance to group consensus in the more subjective Risk Factor ratings.

TABLE 1: RELATIVE DEGREE OF CONFORMANCE

	Figure Areas	Risk Factors
Dissent ($R < 1$)	65.5 %	46.7 %
Neutral ($1 < R < 3$)	27.6 %	23.3 %
Conforming ($R > 3$)	6.9 %	30.0 %

Linking the domination experienced by the participants to the degree of conformance is not very rewarding. The reaction to *high* domination varied greatly – from 99.5 % relative conformance in one case down to 1.6 % in another. At the same time 3 participants registered conformance in excess of 90 %, while indicating that they experienced *very low* domination. If, however, the Domination and Participation dimensions are each divided into 2 groups (High/Low) to form a 4-block matrix and the average conformance calculated per block, the block averages show some – but not statistically significant – trends. For both the Figure Areas and Risk Factors the order of conformance (descending) is: High/Low; High/High; Low/High; Low/Low (Domination shown first). Similarly the only statistically significant difference (by T-test) was between the differences of Low Domination to High Domination. The differences were significant at 95 % Confidence level for the Figure Areas and at 85 % for the Risk Factors. The combination of the High/High and High/Low blocks showed higher conformance indicating, albeit not convincingly, that the high domination experienced did result in a tendency to accept (and conform to) the group consensus view. The reasons why domination was not always resisted, were investigated during the interviews. Three participants in particular reported that they had experienced very high domination, one was a white woman and the other two black (Zulu) men. The reason given was cultural in all three cases. In the first case it became clear that the female participant prefers ‘strong’, dominating men, who she sees as ‘in charge’, ‘in the know’ and consequently worthy of ‘respect’. The other two cases hinged on an apparent combination of culture and the remnants of our apartheid past. The two black participants indicated their ‘respect’ for domination. It is a characteristic of the Zulu culture that the ‘strong man is king’ irrespective of whether you agree with him or not and, subsequently you do not express any

dissenting views. This cultural view could have been enhanced by the fact that many blacks still regard themselves as ‘inferior’, a remnant from apartheid that will take some time to disappear. This said, it must be noted that in the majority of cases the domination was actively resisted and several participants expressed their appreciation (during the interviews) for the AHP indicating that they could use the final individual ratings to express their own view regardless of the dominator’s view. The fact that the reaction to domination is not psychologically homogeneous, ie that different people accept or reject group pressures for different psychological reasons, is in line with the findings of Asch [1958] who distinguished three main categories each for the *independent* (resisting) and *yielding* (conforming) subjects. What must be noted is that the AHP variation used here does not ‘favour’ any particular psychological reaction to group pressure but allows each individual full freedom to react to the group influence in any way he deems appropriate.

A final comparison between Consensus and the Geometric Mean as choice aggregation methods would be to look at the ‘quality’ of the outcome, ie which method gives the best results. It should be clear that in the normal sense of ‘better’ we would not be very successful. Only objective outcomes can be compared in this way and, as we have argued before, only trivial decisions are of this kind. The experiment did include the Figure Areas and, as the author knew the true areas, the different group ratings could, in this case, be compared to the *actual* areas presenting the opportunity to compare the methods in the sense of ‘more accurate’. **Figure 3** shows the comparisons between the seven groups’ I_{1g} , I_{2g} and C_g . Interestingly, I_{1g} ($g=1$ to 4) were closer to the actual areas than the corresponding C_g . Group 7 registered near identical results for I_{17} and C_7 . Only in the remaining 2 groups did the debate improve the outcome with $C_g < I_{1g}$ ($g=5$ to 6). The I_{2g} ($g=1$ to 7) followed the original ratings in I_{22} , the consensus ratings for 4 groups (3, 4, 6 and 7) and split the difference for I_{21} and I_{25} . The means of the group scores are, however, not significantly different. Neither Consensus nor the Geometric Mean can claim victory, particularly as all the means ($I_{1g} = 3.295$, $I_{2g} = 3.798$ and $C_g = 3.325$) are bigger than 2.75 which is the Comparison Index equivalent of a Compatibility Index value of 1.10 - regarded as the level of compatibility to be sought [Saaty, 1996]. Further research is needed to clarify the current results as they do not confirm the research arguing for the superiority of groups in problem solving [Shaw, 1932] or the superiority of debated consensus [Hare, 1952].

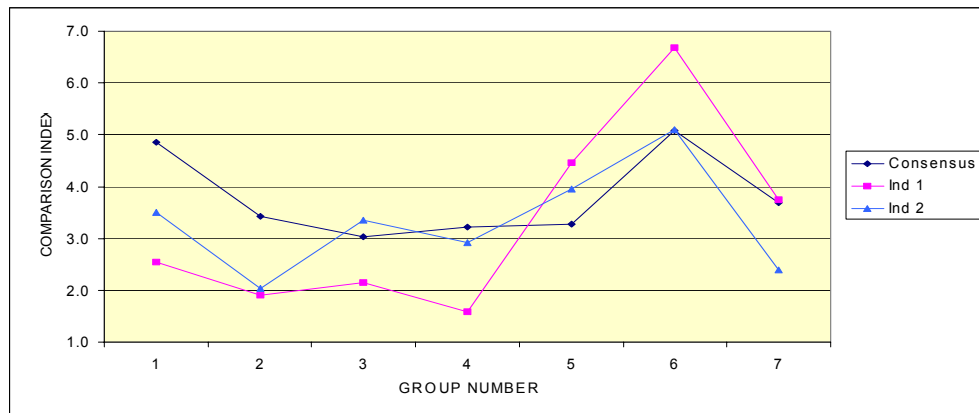


FIGURE 3: COMPARING GROUP AGGREGATES TO ACTUAL FIGURE AREAS

Another basis for comparison of choice aggregation methods is that of the criteria expected of such methods, from a social choice perspective. Saaty, et al [1996] show that the criteria (conditions) as proposed by Arrow [1970] can be summarized into four, ie *Decisiveness*, *Unanimity*, *Independence of Irrelevant Alternatives* and *Non-dictatorship*. They further supply a compelling mathematical prove

showing that by using the AHP and the Geometric Mean all four these conditions are met at appropriate levels of consistency. Here we would like to focus on the last one, which includes both Arrow's [1970] Condition 4 – *Citizen's Sovereignty* and Condition 5 – *Non-dictatorship*. This paper has concentrated on the issues of domination and group pressures as forms of *imposing* choices on participants - of which *dictatorship* is the ultimate form. How, then do the two methods investigated compare in this regard? The results from the empirical study do not provide any indication either way but from a purely theoretical perspective, the interesting conclusion is that consensus, hailed as the ultimate group choice aggregation method, fails the test when evaluated against Condition 4 & 5. In any consensus-seeking situation the possibility of some form of domination is real and this could lead to a decision *imposed* on the participants, not reflecting the group preference. The AHP with choice aggregation via the Geometric Mean can not only be shown to comply mathematically to these two conditions [Saaty, et al, 1996] but also (and importantly), using the 3-phase approach, allows freedom (sovereignty) in expressing preferences.

7. CONCLUSION

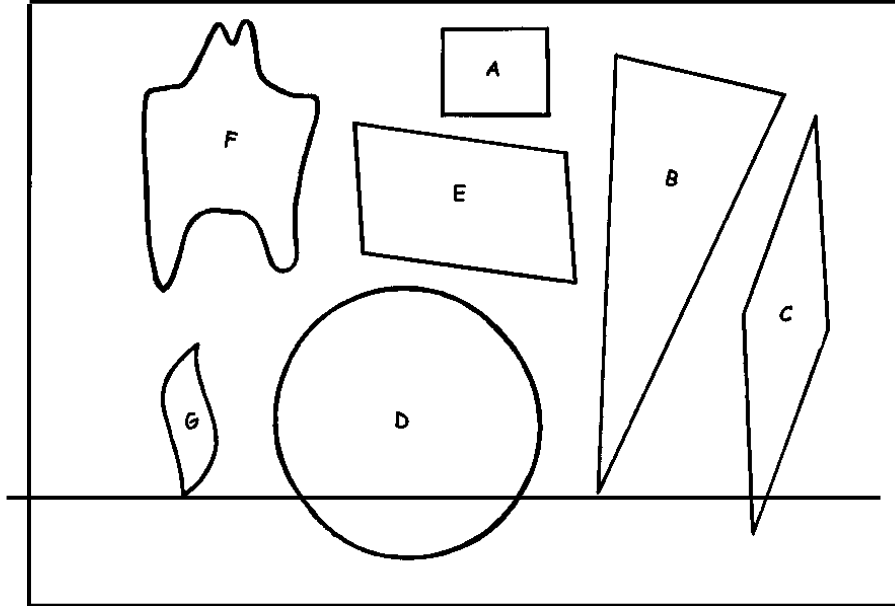
This paper indicated that individual preferences can be distorted in different group decision-making situations either through a dominator present in the group, the pressure exerted by groups on their members or other socio-emotional factors. A modification of the AHP procedure was suggested to allow group influence to benefit the outcome but simultaneously giving some protection against undue domination and group pressures. A preliminary empirical study, in which the usefulness of this modification was tested, was described and the results discussed. Many requirements can be listed for group choice aggregation methods. We presented four in this paper; (i) Allowing the positive group dynamic influences, that cause group performance to exceed individual effort, to play its role, (ii) Providing adequate deliberation and debate to allow group learning and preference structuration, (iii) Providing a mechanism to balance the two seemingly opposing aspects, dissent and consensus while (iv) maintaining a marked degree of autonomy. Evidence was found to indicate that the 3-phase approach does allow the group dynamics and debate to influence individual ratings. The ratings in the third phase included some agreement with and conformance to the groups' consensus views. Simultaneously, however, these ratings also showed that total acceptance of the group view could be resisted. The geometric mean of the individual ratings after discussion was found to embody *both* the influence of the debate and of each participant's own preferences. Clearly, there are a lot more to be done to enhance this approach, but its value, particularly in situations where wide participation is sought, was demonstrated.

REFERENCES

- Asch, SE; (1958); *Effect of Group pressure Upon Modification and Distortion of Judgements*; in Maccoby, EE, Newcombe, TM & Hartley, EL (Eds); *Readings in Social Psychology*; 3rd Edition; Holt, Rinehart & Wilson; New York, NY
- Bales, RF; (1952); *Some Uniformities of Behavior in Small Social Systems*; in Swanson, GE, Newcombe, TM & Hartley, EL; *Readings in Social Psychology*; Henry Holt & Co; New York, NY
- Davos, CA; (1989); *Group Environmental Preference Aggregation: The Principle of Environmental Justice*; *Journal of Environmental Management*; Vol 22/1; 55-65
- Dooley, RS & Fryxell, GE; (1999); *Attaining Decision Quality and Commitment From Dissent: The Moderating Effects of Loyalty and Competence in Strategic Decision-Making Teams*; *Academy of Management Journal*; Vol 42/4; 389-402
- Dryzek, JS & List, C; (2000); *Social Choice Theory and Deliberative Democracy: A Reconciliation*; Working Paper W6 in Social and Political Theory; Australian National University; June

- Hare, AP; (1952); *A Study of Interaction and Consensus in Different Sized Groups*; American Sociological Review; Vol 17/3; June; 261-267
- Marquart, DI; (1955); *Group Problem Solving*; Journal of Social Psychology; Vol 41; 103-113
- Saaty, TL; (1996); *Decision-Making with Dependence and Feedback: The Analytic Network Process*; RWS Publ; Pittsburgh, PA
- Saaty, TL, Peniwati, K & Vargas, LG; (1996); *The Group Choice Problem with Cardinal Preferences*; Proceedings the Fourth ISAHP; Vancouver, Canada;
- Shaw, ME; (1932); *A Comparison of Individuals and Small Groups in the Rational Solution of Complex Problems*; American Journal of Psychology; Vol 44; 491-504
- Sutton, P; (2000); *Tips on Interpreting Mill Process Trial Data: How to Interpret r-Values Obtained Through Linear Regression*; TAPPI Journal; Vol 83/1; Jan; 58
- Taylor, DW & Faust, WL; (1952); *Twenty Questions: Efficiency in Problem Solving as a Function of Size of Group*; Journal of Experimental Psychology; Vol 44; 360-368

APPENDIX 1: FIGURE AREAS



APPENDIX 2: HIERACHY FOR THE RISK FACTORS

