MILITARY DECISION MAKING USING SCHOOLS OF THOUGHT ANALYSIS – A SOFT OPERATIONAL RESEARCH TECHNIQUE, WITH NUMBERS

F. Cameron and G. Pond*

Defence Research and Development Canada – Centre for Operational Research and Analysis

Abstract

Schools of Thought Analysis (SOTA) is a decision-support tool used in group decision making by the Canadian Department of National Defence. Based on their preference rankings of a set of decision alternatives, similarities and differences between participants can be determined. This methodology is meant to identify like-thinking individuals within a group, and also the dissidents, mavericks, and rebels. If contrary schools of thoughts become apparent, the differences between them can reveal critical aspects of the problems facing the group.

Since this method was first presented at the 15th ISMOR in 1998, it has been applied to numerous problems faced by the Canadian Forces. These problems range from materiel acquisition projects to the ranking of research initiatives. In this paper, a subset of those problems will illustrate the application of SOTA. Current and planned future developments related to SOTA are also presented.

Introduction

In the 1990s, the Canadian Army was facing a budget crisis on how to deal with maintaining its existing capability and to find resources for investing in its future. Brigadier General Michael Jeffery assembled a team in Kingston, Ontario to deal with the tough decisions that would ensue. In response to an expression of interest from General Jeffery, one of the co-authors provided an assessment of decision-analysis approaches that were available at the time.

The proposals on offer dealt largely with mechanistic procedures to combine individual preferences into a single list to represent group preferences. General Jeffery found the proposals largely unappealing as they did not deal with the issues he wanted to see from group deliberations. Rather than one list in priority order to represent a recommendation from a group, he was more interested in how factions within the group saw the issues and how they understood the larger context of the problem space. He wanted to have all possible sources of information plumbed, even when those sources might disagree with the orthodoxy.

He wanted to know of potential dissention within the group, and the sources for that dissention. If participants had different points of view, which seemed inevitable, General Jeffery wanted to hear of those points of view and not to have them hidden. One of his concerns was that, if the result from a group were a single prioritized list, such a summary could mask points of contention. Such a list might merely represent the result of substantial compromises to which some members of the group had given their reluctant consent, but did not really support. General Jeffery was keenly interested in additional information that might come from mavericks or dissenters.

---

* Correspondence to Land Capability Development Operational Research Team, LFDTS/DLCD, PO Box 17000, Station Forces, Kingston, Ontario, Canada, K7K 7B4

The original proposals on decision-analysis procedures were then revised in light of General Jeffery's concerns. The result was a series of diagnostics that can be applied in most group decision-making situations regardless of the specific procedure used to generate a single preference list. The diagnostics reveal the potential fissures between components within the group. The set of diagnostics subsequently became known as Schools of Thought Analysis (SOTA).

Since its introduction in Cameron (1998), the method has been applied in a variety of decision-making situations. With the additional experience have come insights as to how best to apply the method. Following a description of the method and two case studies, some of these insights will be recounted.

The Method

The diagnostics tools in Schools of Thought Analysis consist of statistical procedures to determine the strength of a developing consensus (or its weakness) and to identify clusters or schools of thought within the larger group of participants.

For SOTA the first requirement is to have pairwise distances between the participants. For some methods used to derive a single group preference, expressions of individual preferences might constitute vectors in Euclidian n-space, where n is the number of options. If the results do correspond to vectors then matrix algebra provides the means to derive distances between the individuals.

Our experience is more to have individuals provide ranks, rather than vectors. In many practical applications, the participants have limited discrimination between the alternatives; so expecting the participants to put the alternatives in rank order, and sometimes only partial order, is all that can be expected. Even when there are some participants who feel they can discriminate between options with considerable precision, this hardly matters if there remain several participants who can provide only partial ordering and no more. When individuals give their responses as simply a personal ranking of the objects, rank correlation coefficients can be used to derive the counterpart to distances between pairs of participants.

When individuals provide ranks, Kendall's rank correlation coefficient (τ) gives a value between -1 and +1 that indicates the degree of agreement between any pair. The original τ is appropriate only when there are no ties in the data. In the presence of ties, a variant of the coefficient, τ_b, can be used. For a pair (i,j) of participants, a coefficient value of +1 indicates that the two ranks are identical and -1 indicates one of the ranks is the reverse of the other. While the coefficient on a scale from -1 to +1 is not a measure of difference, a simple mathematical transformation will provide a surrogate for distance. Calculating $d_{ij} = (1-\tau_{ij})/2$ means that if two individuals have identical rankings then the distance between them will be treated as 0, and if they are in complete disagreement ($\tau = -1$) then the distance between them will be 1.

With a matrix of such proximities (or distances) between the pairs of participants, two statistical techniques of cluster analysis and multidimensional scaling can be applied. In agglomerative hierarchical cluster analysis, we start with clusters of one element each, and so the total number of clusters at this stage equals the number of elements. The distance between the clusters is the same as the normalized distances between the original elements. The objective is to take steps to combine two of the closest clusters into a larger cluster, and to repeat this until there is just one cluster remaining that consists of all of the elements.

The algorithm is as follows:

1. Select the two clusters that have the smallest distance between them. (If there is more than one candidate pair, choose one of the pairs at random.)

2. Combine the two chosen clusters into one new cluster.

3. If the new cluster consists of all of the elements, then stop.

4. Otherwise, calculate distances between the new cluster and all of the remaining clusters and go to step 1.

For each iteration through step 1, the value of the smallest distance that determines the two clusters is termed “height” for that iteration. The reason for this term being that cluster analysis results are often presented as a dendrogram or tree diagram. It may be drawn with the roots (the elements) at the bottom and those roots combining into the “trunk” of the tree where all of the elements finally combine into one cluster. The height at which clusters get combined is the value from the distance matrix that resulted in the selection of a specific pair of clusters for combining.

SOTA uses three methods of agglomerative hierarchical cluster analysis: single linkage or nearest neighbour, complete linkage or furthest neighbour, and average linkage. For step 4 above, say the new cluster is labelled i, then the recalculated distances from i to the k^{th} existing cluster for nearest neighbour is the shorter of the distances from each of the two components of cluster i and for the furthest neighbour is the longer of the distances. The average linkage method uses a weighted average of the two distances. See Everitt and Hothorn (2010) for details.

Multidimensional scaling is a method for reducing the dimensionality of a set of data. It is frequently used, as in SOTA, to rearrange a representation of the elements into a sensible configuration that is easily absorbed by the human eye. Thus the objective in this application is to generate a map in two dimensions where individuals who are in close proximity in the original n-space are close by each other. Individuals who are far apart in the two-dimensional map should, likewise, be at considerable distance in the original n-space. In other words, the objective is that distances between the individuals in two dimensions are correlated in some sense with the distances between the representation of those individuals in n-space.

The classic procedure in multidimensional scaling starts with establishing a function of the original distances in n-space and the distances in two dimensions that is typically called ‘stress’. This stress function is a measure of how much the configuration in two dimensions is a distortion of the original n-space configuration. A value for the stress function of zero means there is a configuration in two dimensions where the distances between the individuals are exactly as they were in the original multidimensional configuration. Typically this will be found only in trivial examples.

Once a stress function has been defined, an initial configuration in two dimensions is chosen as a start point. The algorithm proceeds iteratively to change this configuration and re-evaluate the stress function. The objective is to reduce stress with each iteration. The algorithm typically stops once a maximum number of iterations has taken place or when the difference between two successive evaluations of stress is less than a specified tolerance. Details can be found in Everitt and Hothorn (2010) and Groenen (1997).

Combining Group Rankings

The decision-analysis literature provides numerous methods for combining individual preferences into a group preference. The overlapping domains of social choice and voting provide even more methods. Methods for combining individual preferences range from the simple (a weighted sum for criteria and scores) to mathematically elaborate (Saaty's Analytic Hierarchy Method or AHP). Data collected or generated by the popular methods for deriving group preferences provide all the elements that allow the addition of SOTA diagnostics, specifically they can all generate some pairwise indicator of agreement (or otherwise) between participants.

For the remainder of this paper, SOTA will be used in conjunction with the rank-sum method to determine a group ranking from rankings provided by individual members of said group. The rank-sum method is outlined by Kendall (1975). However the method predates Kendall substantially, having been offered in a related form for multi-candidate voting by Jean de Borda (1733-1799), and even earlier by Nicholas of Cusa (1401-1464).

Two Case Studies

Value of Lectures in a Military Operational Research Course

In a simple example students in two classes on wargaming applications at the United States Naval Postgraduate School were asked to rank seven earlier lectures in terms of anticipated value to their future professional careers. Each student was instructed to rank the lectures in order where the student could do so, but was allowed to submit ties for lectures that seemed of equivalent value. The lectures ranged in content from a practical exercise of playing a board game of the battle of Midway, to lectures on past history of the use of wargaming and on methods for applying wargames in operational research studies and for designing military campaigns. The data from eight students[*] in one of the two classes is:

		Selected Individual Ranks – by student initials							
Title of Lecture	Abbreviation	jrw	rmk	cl	cgk	krr	amm	wrd	bkm
1. Battle of Midway	Mid	2	3	4	1	4	4	3	6
2. Bauman's Inferno	BInf	3	5	1	6	5	3	7	2
3. History of Wargaming	His	7	2	6.5	2	7	6	2	3
4. Interim Brigade Combat Team	IBCT	5	7	3	6	6	1	4	4
5. Irregular Warfare	IW	1	6	2	6	1	2	5	1
6. Second Rise of Wargaming	2Ris	6	1	6.5	3	2	6	6	5
7. Systemic Operational Design	SOD	4	4	5	4	3	6	1	7

Most responses are in the form of a permutation of the integers from 1 to 7, with "1" indicating "most preferred". However, the respondents did include ties in their submission: for example, "cl" had "His" and "2Ris" tied in last place. For ties, the average rank of the tied lectures, had they not been tied, is used: "cl" has 6.5 for the two items that would have had ranks of 6 or 7 and "amm" has 6 for three items that

[*] Student initials have been replaced by randomized initials to protect their identities.

would have had ranks 5, 6, or 7. One way to check this arithmetically is to observe that the column sums should always be $(n+1)n/2$, or, in this case 28. This rescaling of raw ranks is called the canonical form in this paper, from Kendall (1975).

From observation, we can see that “Mid” was first for “cgk” and second or third for three other students and “IW” was first for three students, although “rmk”, “cgk”, and “wrd” did not rank this lecture very high at all. Six of the eight students put “SOD” below average, but with “wrd” ranking it first.

There certainly seems to be some internal agreement. First, as noted, most students gave “SOD” a low rank. Second, “jrw” and “cl” ranked “IW” and “BInf” high and “His” and “2Ris” low. Third, four students put “Mid” in first, second, or third place and five students put “IW” in first or second place. These examples suggest that there may be some subsets of the eight students who would have some common views on the value (or otherwise) of some of the lectures.

From the previous table, the group preference from rank-sums is:

Lecture	Rank Sum	Group Rank
IW	24	1
Mid	27	2
BInf	32	3
SOD	34	4
2Ris	35.5	5.5
His	35.5	5.5
IBCT	36	7

However some students in the class might have some issues with this as a group ranking. For instance, “rmk” and “cgk” both had “IW” in sixth place, but the rank-sum put “IW” in first place for the group. The rank-sum put “IBCT” in last place, yet “amm” personally had it in first place.

Clearly this is a compromise amongst the eight selected students. As we shall see, despite possible imperfections, four of the eight students are in general agreement with this group ranking, and the remaining four students show no significant disagreement with it. In that sense, it may represent a reasonable compromise.

There are two related statistical tests to determine if the group ranking represents a strong consensus: Kendall’s coefficient of concordance, W, and Friedman’s Test. Conover (1998, Chapter 5) shows that the two tests are closely related and that they will give results that are mutually consistent.

For this data, Kendall’s W is 0.074, which yields a p-value of 0.75. Thus these rankings demonstrate no strong consensus. At this point an analyst may be left in a quandary. Should the analyst say only there is no consensus that is statistically significant and refuse to disclose the group result? Should the analyst divulge the group ranking, but include a warning that there seems to be no consensus? Or, should the analyst investigate further? SOTA provides diagnostics for that further investigation.

The Diagnostics

The first step in SOTA diagnostics is to look at the pairwise agreement between participants. When doing this, a "surrogate participant" can be added to represent the group ranking (or consensus result), so pairwise agreement with the group preference can also be presented.

A macro was developed in Excel that provides Kendall's rank correlation coefficient (τ_b) for all pairs of participants, and for pairing each participant with the group's ranking by rank sum. The surrogate participant to represent the group preference is called "Borda" in this presentation, to honour Jean de Borda. The coefficients are colour-coded by the macro. Non-significant results are in black. The table entries show statistical significance better than 5% in orange (and italics) and better than 1% in red (and bold). Note that a negative value in red or orange means disagreement that is statistically significant.

	jrw	cl	krr	amm	bkm	cgk	wrd	rmk
cl	**0.59**							
krr	**0.43**	0.00						
amm	**0.41**	**0.63**	-0.10					
bkm	0.14	**0.39**	-0.05	**0.41**				
cgk	**-0.31**	**-0.58**	-0.10	**-0.50**	**-0.41**			
wrd	-0.14	**-0.49**	-0.14	**-0.31**	**-0.43**	**0.31**		
rmk	**-0.33**	**-0.59**	0.05	**-0.82**	*-0.24*	**0.62**	0.05	
Borda	**0.78**	**0.40**	**0.39**	*0.21*	0.20	-0.05	-0.10	-0.20

The table above shows that four of the participants "jrw", "cl", "krr", and "amm" show considerable agreement with "Borda"; note that the macro reorders the columns for the participants from high agreement to high disagreement with "Borda", hence the declining agreement with "Borda" from left to right is part of the diagnostics. Of the four who do not show agreement with "Borda" at a statistically significant level, the results, even for "rmk" do not exhibit disagreement that is statistically significant. We often use the term *ambivalence* for the relationship between pairs where the correlation coefficient is not statistically significant.

For the eight students, there are pairs of substantial agreement (e.g., "jrw" and "cl", and "cgk" and "rmk"). There are also pairs that demonstrate substantial disagreement (e.g., "amm" and "rmk"). One, "krr", is notable in having considerable agreement with one other (namely "jrw"), but ambivalence to the ranks provided by the other six students (and then substantial agreement with "Borda"). Observation of the individual rankings provided by the eight students and comparison between them provides evidence for why there is agreement, disagreement, or ambivalence between the various pairs.

SOTA provides methods, drawn from common statistical procedures, to investigate the sources agreement and disagreement further. These methods are three forms of hierarchical cluster analysis and multidimensional scaling. Figure 1 has dendrograms for single-linkage, complete-linkage, and average-linkage forms of hierarchical cluster analysis, see Everitt and Hothorn (2010, pp. 319-321). The "complete" and "average" methods produce diagrams that are largely consistent: two sub-clusters of five students around “Borda” – “jrw” and “krr”, and “bkm”, “cl”, and “amm” – and a sub-cluster some distance away of “wrd”, “cgk”, and “rmk”. The “single” diagram shows the triple of “wrd”, “cgk”, and “rmk”, with the other five students clustered around “Borda”.

In practice, representatives from the various schools of thought (clusters) can be asked to give reasons for their preferences. Experience with applying SOTA has shown this is a valuable means to draw out specialist knowledge not previously shared with all of the other participants. It may also lead to a renewed discussion of the criteria used to rank the alternatives, as statements of proposed criteria often leave a lot to individual interpretations. In this example, the participants could return to a discussion of their individual interpretations of “anticipated value to their future professional careers”. It may be that, only once the participants have applied some candidate criteria to alternatives, do they realize the interpretations by some participants are at considerable variance with the interpretation of others.

In this sense, SOTA is related to the Delphi method first developed by the RAND Corporation in the 1950s and 1960s, see Dalkey (1969). The RAND proposal to be inclusive of alternate points of view remains widely popular within the operational research community. Dalky points to the anonymity afforded to the participants as one of the key features of Delphi. SOTA, as it stands, requires the identification of individuals with their position relative to the others. Individuals are invited to comment based on their agreement or disagreement with others (anonymity is renounced).

Of course the analyst may preserve anonymity by keeping the results of SOTA confidential from the participants, and use the diagrams to select participants who should speak up (perhaps anonymously) when the time is appropriate. However, sharing the SOTA diagrams with the participants during the decision-making process can have beneficial results. When the configurations are revealed these can show some participants that they have nearby allies, and further discussion can thereby be promoted. For some individuals knowing that they are the mavericks, with no apparent nearby allies, is a motivator to inject their opinions during debate. There is no evidence to date that including anonymity in applications of SOTA would improve the decision process.

The three different hierarchical cluster analysis methods are used in SOTA to construct the trees in different ways, so can offer the analyst different perspectives. The literature on hierarchical cluster analysis provides background on why and how the various dendrograms, based on the same data, will appear different, with different membership in the various sub-clusters. Many depictions of the various cluster analysis methods recommend that, when applying cluster analysis, an investigator should include diverse methods specifically because the structures that come from the different methods can reveal different characteristics of the underlying data.

SOTA also includes a configuration of the participants based on multidimensional scaling as shown in Figure 1. This shows the very close proximity of “jrw” and the group ranking, with the text for “jrw” and “Borda” overprinting, and “krr” not far away. The configuration also shows the triple of “bkm”, “cl”, and “amm” near the group ranking and the triple “wrd”, “cgk”, and “rmk” some way off. The orientation of the axes and labelling of the tic marks in multidimensional scaling is mainly to orient the viewer for the distances between the individuals.

At this point, there appears to be a school of thought of five students (“jrw”, “krr”, “bkm”, “cl”, and “amm”) and another of three students (“wrd”, “cgk”, and “rmk”).

The data from the selected students provides an application of SOTA when the concordance between participants is not strong (or statistically significant). As we shall see, SOTA can also bring insights to a decision-making process where the participants are already in substantial agreement.

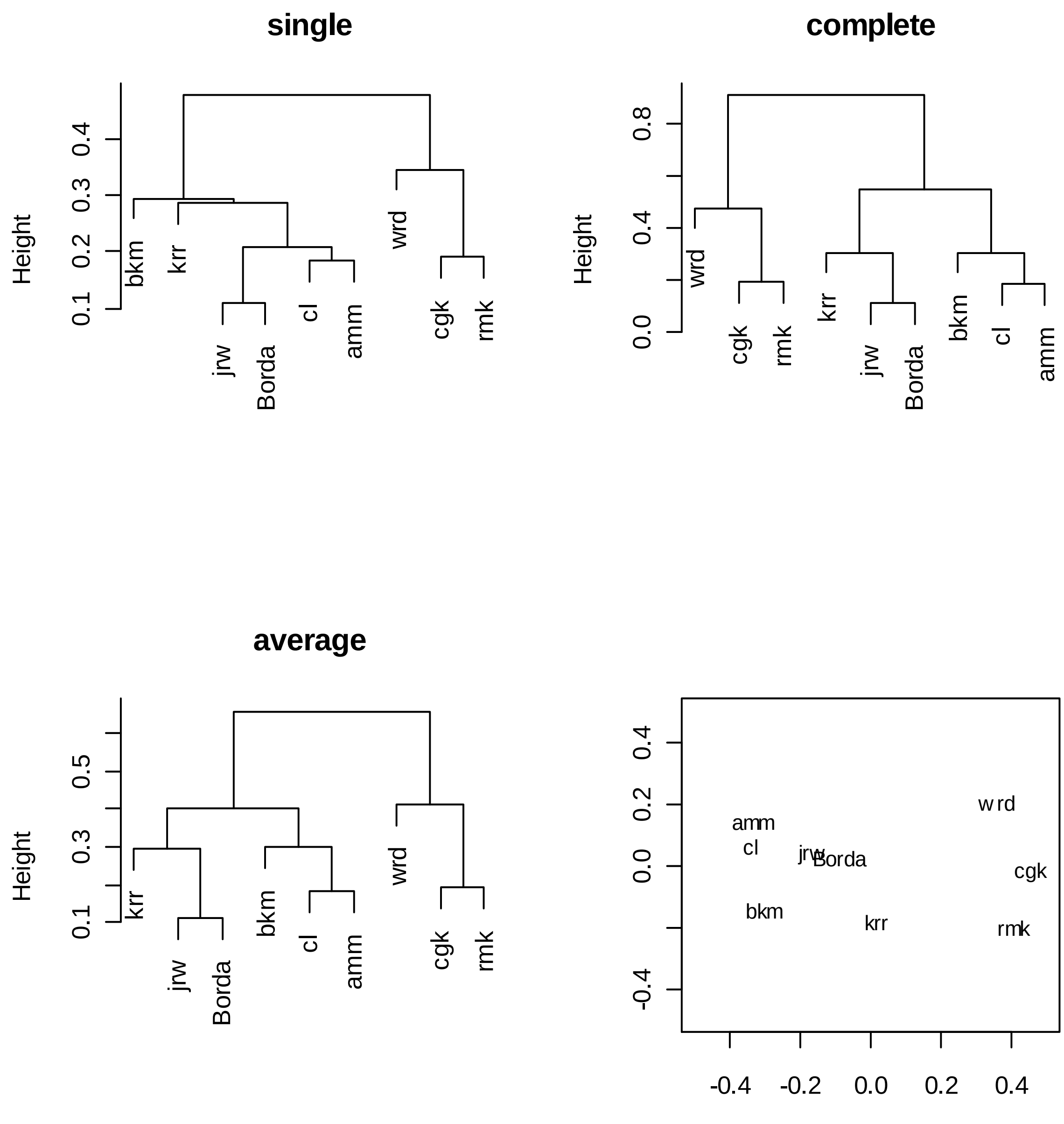

Figure 1: Configurations by three methods of cluster analysis and by multidimensional scaling

Arctic Brainstorming Exercise

The Canadian Government has outlined several strategic goals in respect to the Canadian North and the Arctic Region in particular. Among these goals, reinforcing sovereignty and national security will be of primary interest to the Canadian Forces. To support this renewed focus on Arctic operations, the defence research community is prepared to invest in a variety of new initiatives. The immediate challenge is to identify which among these various areas of research offer the greatest potential return-on-investment to the Canadian Forces.

To this end, a meeting of subject matter experts (SMEs) was convened to rank broad areas of research in terms of the potential benefit of research in that field, and the span of the existing knowledge gap in that field. The six potential areas of research were identified as:

i. Soldier Issues (e.g., clothing, diet);

ii. Mobility;

iii. Power Management;

iv. Sustainment;

v. Command and Control and Communications and Information Systems (C2-CIS); and

vi. Equipment.

The potential areas were discussed at considerable length and a widely held understanding of their meaning and their significance developed amongst the participants – more than may be suggested by the codes for the six areas shown above.

Following lengthy deliberations, the nine participants, six military and three civilian, were asked to rank each of the six areas independently in terms of benefit, and in terms of knowledge gap. Ties were allowed in each participant's rank-order. The individual ranks for military benefits, with the group ranking (called "Borda") to the right, are below.

	Participant Identifier									
Issue	Maj-C	Col	Maj-R	Maj-S	Civ-S	Civ-T	Maj-F	MWO	Civ-R	Borda
Soldier	1	1	2.5	1	1	1.5	2.5	1	4	1
Mobility	2.5	2	1	2.5	3	1.5	2.5	2	5	2
Pwr Mgmt	2.5	3.5	2.5	4.5	4	4	2.5	6	1	3
Sustain	4	3.5	4.5	2.5	2	4	6	5	6	4
C2-CIS	5.5	6	4.5	4.5	5	6	2.5	3	3	5
Equipment	5.5	5	6	6	6	4	5	4	2	6

In this example, Kendall's W was 0.42, which is statistically significant beyond a p-value of 0.001. So this group demonstrated a high level of concordance. That said, applying the diagnostics from SOTA reveals that some sources of dissention are evident.

The table of pairwise correlation coefficients below shows considerable agreement between the first five participants (Maj-C to Civ-T) with all entries showing statistical significance beyond a *p*-value of 0.001 (red).

	Maj-C	Col	Maj-R	Maj-S	Civ-S	Civ-T	Maj-F	MWO	Civ-R
Col	**0.89**								
Maj-R	**0.69**	**0.59**							
Maj-S	**0.69**	**0.74**	**0.54**						
Civ-S	**0.64**	**0.69**	**0.50**	**0.93**					
Civ-T	**0.75**	**0.89**	**0.59**	**0.59**	**0.54**				
Maj-F	**0.37**	0.18	**0.55**	0.18	0.09	0.20			
MWO	0.21	*0.28*	0.21	**0.50**	**0.33**	**0.39**	**0.43**		
Civ-R	-0.07	-0.14	-0.21	**-0.50**	**-0.47**	-0.23	*0.26*	**-0.33**	
Borda	**0.93**	**0.83**	**0.79**	**0.79**	**0.73**	**0.70**	**0.43**	**0.33**	-0.20

All three of the cluster analysis presentations in Figure 2 indicate a strong consensus amongst six of the participants, and each is closely aligned with the group ranking ("Borda"). In the MDS map, the six participants are so closely aligned that the identifiers are over-printed.

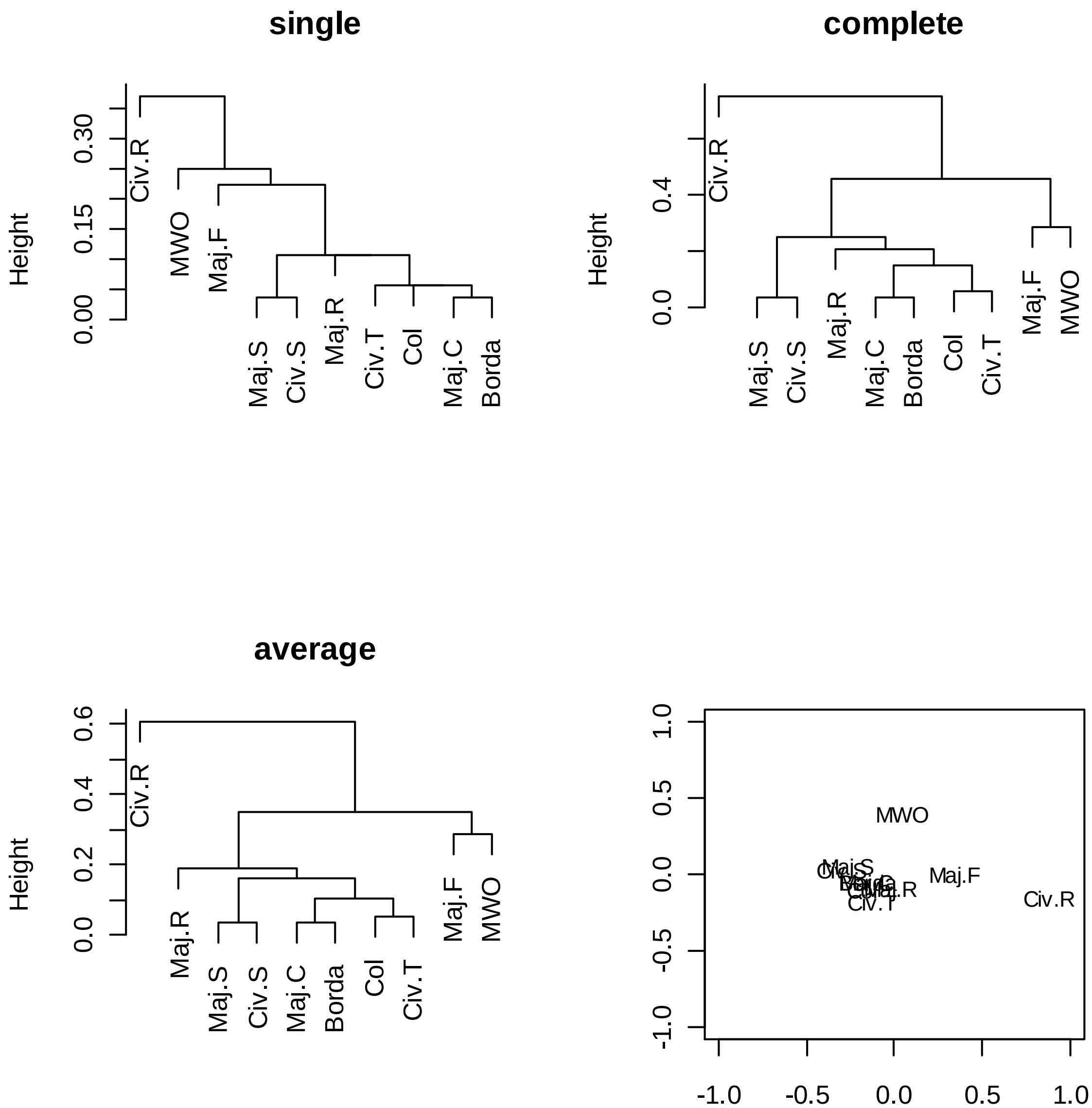

Figure 2: SOTA configurations for Arctic brainstorming

All four diagnostic figures from SOTA (three forms of cluster analysis and a map from multidimensional scaling in Figure 2) show the participant identified as “Civ.R” to be something of a maverick. “Civ.R” is the last participant to join the final cluster regardless of method and is located a long way from most other participants in the MDS map.

Discussion

With the two case studies above to illustrate SOTA, several aspects of applying methods like SOTA can be discussed. This starts with comments on statistical tests that are incorporated. Then there are some comments on exploiting contrarian views (or critical thinking) in policy making at the highest level. The following material outlines how and why contrarian views may be incorporated into more general group-decision processes, including such aspects as using devil's advocates, and dealing with individuals who may become so critical of analysis methods that they want to withdraw from the group, or renege on further participation. Clearly, choosing who should and should not participate needs to be a carefully considered process, and some comments are provided on this. Guidance is provided on the practical application of the method, particularly the time that would be required to employ SOTA in typical situations. This is followed by a proposal that SOTA (and, indeed, other decision-support techniques) should first be applied with a new group of participants in some non-contentious setting before tackling the real problem: novice participants are often unsure of how some new process might unfold and seem to benefit from seeing it applied to a test case where the issues are less emotive. Some techniques, notably the Rand Corporation's Delphi Method, provide participants with anonymity and there is a brief discussion of the relevance of anonymity within SOTA. The discussion includes some comments on creative thinking and critical thinking, including an admission that SOTA tends to provide more support for the latter than for the former. The discussion concludes with an outline of some future developments of SOTA. Following the discussion, the paper provides five lessons that were learned from applying SOTA to a wide variety of problems.

Kendall's Coefficient of Concordance and Friedman's Test

Conover (1998) provides a comparison of Kendall's coefficient of concordance, *W*, and Friedman's Test for agreement amongst judges when ranking objects. Kendall and Friedman independently developed approaches that are essentially equivalent, as Conover demonstrates. Conover goes on to demonstrate how the approach of Friedman's Test can be used to determine when a statistically significant result can or cannot be found between how objects have been ordered by a group, i.e., where partial ordering is all that can be derived. That is to say, while the group may show a preference for one object over another by group-rank, the two objects should be treated as tied since the strength of preference is not strong. The application of this to SOTA is taken up in Pond and Cameron (2010).

Kennedy-style Inclusion of Contrarians

The story is told by Janis (1989, p. 329) of a beneficial change in group decision-making from one international crisis to another in the Kennedy White House: "Such outsiders were deliberately brought into the executive committee's meetings by President Kennedy during the Cuban missile crisis and they were urged to express their objections openly. This leadership practice was quite different from that which President Kennedy adopted throughout the Bay of Pigs planning sessions…" Janis (1972) coined the term *groupthink* for pathologies he found in the Bay of Pigs decision-making process.

While the contrast of the leadership and decision-making style in these two instances is amongst the best known in the field of statecraft, many senior leaders do sincerely wish to involve many individuals in group decision making, a Kennedy trait shared by many military leaders. Indeed General Jeffery, when setting up the apparatus in Kingston to deal with the future of the Canadian Army, wished to be highly inclusive, inviting participation beyond the military profession to draw in defence scientists, the operational research community, and reaching out to allies, industry, and academia.

As Janis (1989) illustrates, the Kennedy initiative of being more inclusive resulted in considerably better decisions during the Cuban Missile Crisis than during the Bay of Pigs operation (known more pejoratively as the "Bay of Pigs Fiasco"). But a similar intent to be inclusive can also have debilitating results on group dynamics. So invitations to participate have to be handed out wisely.

The Devil's Advocate

Daft and Marcic (2009, p. 200) describe a long-standing desire in applications of management science to incorporate contrarian views, sometimes to the point of designating a devil's advocate whose mandate is specifically to challenge the assumptions and conclusions of others in the group. Much of this dates back to the early work on groupthink in Janis (1972). Recent initiatives within military circles conform to the recommendations that Daft and Marcic intended for civilian application: Craig (2007) provides a summary of her experiences with 'red teaming', procedures recently introduced in the US to designate individuals to perform a role similar to a devil's advocate and resembling Kennedy's initiative for greater inclusivity (beyond the experts) during the Cuban Missile Crisis.

Evidently, the most apparent reason to invite individuals to join a decision-making group is their expertise in the subject matter under investigation. An inclusive approach, reminiscent of that of the Kennedy White House during the Cuban Missile Crisis, will draw in participants who are prepared to take on a role as "devil's advocate" even when their subject matter expertise on some matter may be limited. Within the mythology of operational research there are stories dating from World War II of the pipe-smoking tweed-jacketed outsider at the back of the room who rightfully challenges the conventional wisdom of military staffs. Critical thinking being fundamental to good decision making, thus provides a reason to include some participants who would not qualify merely for their expertise on the technical aspects of a problem.

Decision Groupies

Apart from these two laudable purposes for drawing participants into the process (when they can provide expertise and critical thinking), there may be a few additional legitimate reasons for other individuals to join the group. However many such groups can have individuals added for other purposes, perhaps an individual merely indicates a passing interest, or perhaps someone is inserted to be a 'lobbyist' whose intent is to browbeat the others into support for a predetermined point of view. Clearly some purposes for adding participants to the group may subvert the objectives of applying rigour and objectivity.

We have coined the term 'decision groupies' for those individuals who have been added to the decision-making process for reasons other than expertise or critical thinking. Meta-decisions to include such groupies may be taken lightly, as the consequences may seem innocuous at the time[*]. But since each individual added to the process comes with considerable potential to influence the outcome, such inclusions should be carefully considered. Since one aspect of SOTA is to give voice to those who do not seem to agree with the majority, there may be opportunities for decision groupies to distort the results,

[*] The term 'meta-decision' is used here as these are decisions that will affect how other decisions will be made.

sometimes subversively. Jean de Borda was challenged at one point that his proposed voting method could be subject to manipulation by unscrupulous participants; his vexed response was: ‘My method is intended for honest men only’. SOTA should be resistant to a small number of uncoordinated ‘decision groupies’, but a widespread and coordinated manipulation or a large proportion of like-minded ‘decision groupies’ could be a problem (as they would be for most other decision-analysis procedures).

Participants Who Renege

Many of those who participate in our group activities have had little or no prior experience with activities where decision-analysis methods have been applied. This is not to say they are unfamiliar with collaborative decision making. But many of their prior collaboration activities will have involved participating in a leadership team or command team consulting specialists in various areas and with a designated leader who will combine divergent assessments to establish a way forward. Such teams would leave a final decision to the most senior officer if there were indications of dissenting opinions. Less familiar for many has been activity where all participants will be treated as equal within the process. The unfamiliarity may lead to doubts about the process. In practice, when participants lack confidence in the decision-analysis methods, they may renege and challenge the validity of the whole exercise.

From time to time, participants in group decision making have reneged, or threatened to renege, from the group-ranking process as in Mason (1995). Reports of such behaviour are rare, but it may be more common than would be supposed from the literature. That may be a consequence of more reporting of successful methods than of those that have failed or, as in Mason’s case, exhibited incipient failure. Reasons for such mutinous behaviour can be varied. But from our experience it is usually the result of a participant lacking confidence in the process. A loss in confidence may be driven for parochial reasons: a participant seeing that the group is gravitating away from his or her own preferences may feel that the process is flawed – the presumed logic may go something like: “I am right, but the group disagrees with me. I am extremely eloquent in putting my case, but they are still not coming around. I *do* respect the other participants, so any failing on their part can be ruled out. So it must be the process we are using that is wrong.”

Regardless of the possible logic that might justify reneging, an analyst must be sensitive to it. Confidence of participants in procedures used for a decision-making activity may come only with greater familiarity with those procedures.

Employment of Schools of Thought During Implementation

In most decision-making processes, at least one school of thought should be evident that would support a subsequently chosen course of action (call it the dominant opinion); such a group would be particularly evident if their preferences exactly matched those of the decision maker. Note that the chosen course of action may not necessarily be the consensus result from the group; for various reasons, an external leader may disregard the most preferred course that comes from the group and choose a different path.

So, due to external (and quite legitimate) factors, the chosen course of action may well not be the one with the greatest support from within the group. But once some course of action has been selected, the members of the school of thought most closely associated with that course could constitute a pool of potential proponents during an implementation phase, since that course should be generally appealing to members of that particular school of thought. While these proponents (chosen from the dominant school of thought) should not completely rule during the implementation stage, they can be expected generally to bring more enthusiasm and commitment to follow the chosen course.

Contrarily, participants for whom the subsequently chosen course had little appeal during the decision-analysis stage may be prepared to give that course only weak support, if that. Commanders in a military operation may well be reluctant to keep such apparent pessimists around when they really wish to have enthusiastic allies and subordinates in reaching some objective. Of course, military training generally includes injunctions to give commanders full loyalty during implementation even if the chosen course of action seems personally unappealing, inexplicable, or even life threatening.

Members of schools of thought who could see little merit in the subsequently chosen course of action may still have a role during implementation, even if they cannot be expected to give that course their uncritical support. Indeed they may make excellent candidates to monitor progress and to trigger a reconsideration of the chosen course if the intended benefits are not soon evident. Such a ‘canary in the coal mine’ is often particularly valuable to trigger a reassessment when the need for this may not be very apparent to those who are uncritical proponents.

Use of Microsoft Excel

To facilitate the application of the SOTA diagnostics, a macro was developed in Microsoft Excel to carry out some of the procedures. From a table of individuals and their individual ranks, the macro produces a new worksheet that includes the following:

- a reformulation of the raw ranks into Kendall’s canonical form;
- a re-ordering of the items by rank sum;
- all pairwise correlation coefficients, including the correlation between participants and the ranking by rank sum, colour-coded by statistical significance;
- a table of *p*-values to determine the statistical significance of each of the pairwise coefficients; and
- provision of Kendall’s coefficient of concordance, *W*, and a test of its statistical significance.

To facilitate the further processing of the data, the macro also provides a small text file with the original ranks reformulated into canonical form. This text file is easily readable by software packages used for statistical analysis. In previous applications, additional analysis was done with Systat, a commercial software package for statistical analysis, but more recently this has been performed by R, a computer language for statistical analysis, and by Matlab, a mathematical programming environment.

In application, forms are generally generated locally and specific to each session where SOTA is employed with the alternatives listed on a single sheet of paper. One form is provided to each participant, who indicates personal preferences by filling in numbers against the alternatives that are listed. In typical applications, the ranks from the participants as provided on the forms can be transcribed into Excel in a matter of minutes. The macro, including recording a text file for further processing, runs in a matter of seconds.

Use of the R Programming Language for Statistics

Cluster analysis and multidimensional scaling procedures are available in commercial products for statistical analysis like SAS, SPSS, Systat, and S-Plus. But R has since been found to be more convenient and powerful, and readily available as open-source software.

The R code to produce the configuration of students in Figure 1 is straightforward (see annex), albeit a bit opaque for those not familiar with interpreted languages like R (or with S or S-Plus). Everitt and Hothorn (2010) provide brief explanations of both the cluster analysis methods (section 18.2.2) and multidimensional scaling (section 17.2.2) and examples in the R language that demonstrate their application.

The computer code for R, shown in the annex, runs in a matter of seconds on a standard personal computer. In many brainstorming sessions, some of the output from the Excel macro and the graphics from R are then transferred to Microsoft PowerPoint for ease of presentation back to the group.

Scheduling a Group Activity that Uses SOTA

Before SOTA is applied a group is expected to have converged on criteria for preferring alternatives and to have developed a list of alternatives. There are numerous techniques for this; many have been offered in the operational research community under the topic of “creative thinking”. For scheduling of SOTA activity with the participants, the usual sequence is to pass the preference forms around to the participants and send them to a break. If the group had reached a point where both the criteria for the individual preferences and the essential characteristics of the alternatives are well understood, the completed forms are generally returned in ten minutes or less. In practice the analyst can work with Excel quickly enough that the forms are transcribed as they are returned, and the processing can start immediately once the last form has been returned and transcribed, with results from Excel and from R in a matter of minutes.

However, the analyst will generally want some time to consider the significance of some of the SOTA results, and to plan how to introduce these results back into the group deliberations. From the time that the blank forms are circulated until the analyst is prepared to present results has generally been 15 to 30 minutes, with most of that interval provided to participants to complete forms and to the analyst to consider the interpretation and significance of the results.

The objective of the schedule for applying SOTA is to give the participants a bit of time (say a 30-minute break) to consider the ranks they have submitted, and to consume some refreshing beverages if appropriate. Then they return to a group session to debate the alternatives further, with the SOTA results as a stimulant to draw out dissident opinions and alternate points of view.

If the algorithms from the Excel macro and the R source code were incorporated into a computer-based system, such as the decision lab described in Ward (2009), results should be available in less than a minute. However, in practical applications, working with pencil and paper lends considerable informality to the SOTA process and ensures it can be highly portable, as in the wine-tasting example below. Further, even if results were available in a minute or two, the analyst may still want some additional time to consider how to introduce the SOTA results back to the group so the participants can engage in meaningful subsequent discussions.

Impact on Group Dynamics

When SOTA has been applied with a group where participants have already established mutual credibility over some previous time, say from prior professional interaction on other issues, the method seems to have considerable success. Individuals who are identified as outliers or dissidents generally have retained respect from the others even when they have views that may conflict considerably with the majority of participants. Quite often the phenomenon of outliers has been treated with considerable good humour by all concerned. Furthermore the outliers are rarely reluctant to speak about the reasons for their contrary opinions, and, indeed, many even relish the role of maverick. Sometimes the identities of those mavericks will come as no surprise to them or to their colleagues – the dissidents may well have a track record within their professional circles for contrarian views or for critical thinking.

However, in one case, an individual who had only recently joined a pre-existing team was soon identified by SOTA has having divergent views. In this case, there were clear issues of embarrassment over the new participant being singled out as an apparent dissident. The group, which normally worked in close proximity on a day-to-day basis, had removed itself to an off-site location for brainstorming specifically intended for creative thinking. At the end of the off-site activity, SOTA was used to determine if there was consensus on the more valuable concepts, and, where there was less consensus, how corresponding concepts should be viewed. Several of the individuals demonstrated contrarian views, and this led to a lively and good humoured debate. All but one of the contrarians, the individual in question, had been working in the group for several months previously and had developed a solid relationship with the others. However the participant who had been with the group for only a short time was evidently somewhat self-conscious about seeming to be identified as “failing to support the team”. Subsequent to the off-site activity, this individual has gone on to gain considerable respect within the team, and is now specifically sought out for contributing a different perspective. But at the time, such an inclination for difference from the group was viewed with some embarrassment.

Participants must be aware that SOTA will focus on sources of dissention, and if the method is to contribute to improving the group’s processes, the reasons for the dissention need to be shared with other participants. In some settings, especially if conformance with the rest of a group is viewed as an imperative, the response to SOTA results, since it finds contention within the group, may be unwelcome.

Application of SOTA to a Non-contentious Issue

The ABCA Armies Program is an agreement between five armies (American, British, Canadian, Australian, and more recently New Zealand) to improve standardization within coalition land operations. In September 2006 ABCA conducted the Interoperability Gap Analysis Symposium at Tobruk Barracks in the Puckapunyal Military Area in Australia to identify the most critical gaps that coalitions were facing. Each participating army provided a dozen or more uniformed personnel who were experts in various areas. After enumerating a considerable number of potential gaps, the participants developed a prioritized list of the more critical ones. SOTA was used in conjunction with other decision-making techniques. There was some initial concern amongst participants that the decision-analysis methods would obscure their concerns, and suppress contrary opinions. Had these concerns gone unaddressed, it was possible that participants would give only weak support to the result, similar to the mutinous mood that Mason (1995) reports in another case study.

One evening, SOTA was used in the Tobruk Officers' Mess in a blind tasting of local wines. In a relatively non-confrontational setting, the participants could see how their individual judgements would be combined to provide a group ranking. They also saw how various schools of thought could be identified and that dissenters and mavericks could be allowed (even encouraged) to have their say. The results were presented with both the overall consensus ranking, and with clear identification of where there were divergent opinions. Indeed having divergent views contributed to the conviviality of the evening.

When the participants returned the following day to the main task for the group, the participants in the wine tasting had developed substantially greater confidence that the methods used by the analysts would lead to a balanced view of the contributions of all, even those contributions that might differ from the majority.

Creative Thinking and Critical Thinking

Two aspects that are fundamental to good decision making are creative thinking and critical thinking. From the material presented above, SOTA has a clear application in critical thinking as the method is intended to encourage critical debate of an existing set of alternatives, which may have already emerged from some burst of creative thinking. In particular, SOTA should encourage mavericks and contrarians to verbalize how they view the various alternatives and why they value some of the alternatives differently from other participants. In that sense, the method allows (even compels) critics of some developing consensus to speak up while that consensus is still open for further debate and adjustment. This discussion of different points of view may stimulate more creativity to refashion some of the alternatives to respond to the various contrary perspectives.

While SOTA has an apparent link in that sense specifically to critical thinking, there is another useful, if less obvious, connection between SOTA and creative thinking. SOTA (and diagnostic methods like it) can as least suggest that a creative-thinking phase may have been inadequate so far. This would be the case if testing Kendall's coefficient of concordance indicated that the ranks from the participants appeared little different from random results. Such a finding can demonstrate that further resources will be needed to revisit the creative-thinking phase. Forcing a creative-thinking phase to be revisited might then result in the generation of new alternatives or a reworking of some alternatives that were previously in contention. Without such a stimulus to say that more creative thinking is still required, the participants might blithely accept inadequate results prematurely.

If the list of revisited alternatives is later re-evaluated with the SOTA diagnostics applied in the new iteration, SOTA should show if the new group has now achieved a consensus with greater support than in the previous application. If a group were to conduct iterations of creative and critical thinking, SOTA might even give an indication that a "stopping condition" as been reached where the group's preference for the current alternatives finally has the benefit of strong and widespread support from the participants.

However the contribution from SOTA for creative thinking is necessarily limited. In some sense SOTA may be used to detect that there are some holes in the range of alternatives – it may show that the group has no strong consensus that even the most attractive of the alternatives are sufficient for the problem facing the group. But it will not in itself spin out new alternatives to fill a "gaping hole". However, it should reduce the likelihood that the group can ignore the gaping hole. The group cannot casually conclude that their creative thinking phase has produced all the alternatives that need to be considered when SOTA says there is little concord for the preferred alternatives.

Anonymity: Not the Solution for All Ills

As mentioned the Delphi method was developed by the RAND Corporation in the 1950s and 1960s and remains popular. A key component of the early Delphi process was anonymity for participants. Dalky (1969) reported a comparison of face-to-face discussions with a process that enforced anonymity and concluded: 'The results indicated that, more often than not, face-to-face discussion tended to make the group estimates less accurate, whereas, more often than not, the anonymous controlled feedback procedures made the group estimates more accurate'.

The objective of the anonymity is to ensure that issues like rank or assertiveness do not colour the relevance of an individual's arguments. Since interventions from participants cannot be associated with specific identities, it is hoped that those contributions will be assessed by others only on the merit of the content.

Currently in applications of SOTA, participants do not retain their anonymity. Indeed, one purpose of SOTA is to encourage dissenters to speak up once they see they have potential allies who have similar views. Thus each individual needs to know where he or she fits with other members of the group, and each needs to know how close (or distant) he or she may be to others.

Analysts applying SOTA techniques can use the configurations to find individuals who may be a spokesperson for some contrary opinions. Such spokespersons may be invited to speak even when they may otherwise feel reticent. In this respect SOTA is intended to pre-empt situations where participants may remain silent thinking they are 'voices in the wilderness', or may withdraw since they judge that their contributions are seen by others to be without value. If such participants come to feel their contributions were unwelcome or overlooked, they may later give only weak support to the result since they could feel the process was flawed from the beginning.

Future Development

SOTA has been implemented with an Excel macro and the addition of some simple commands in the R programming language. In this form it has been a valuable addition to other group decision-making procedures and efficacious in application. However these procedures will be revised and streamlined so they can be applied by other analysts without requiring that they should have much specialized training.

As reported by Conover (1998) a version of Friedman's Test can be applied in post hoc analysis of the various objects that judges have ranked individually. This application is for ranks and is a non-parametric procedure in statistics. There is a corresponding procedure in parametric statistics that is widely used. Analysis of variance (ANOVA) may indicate rejection of a hypothesis that the means are the same for mutually exclusive sub-groups of some data. Rejecting this hypothesis, the analyst still faces the problem of determining which sub-groups appear to have the same mean and which have different means. So, having rejected the overall hypothesis on means, post-hoc analysis with tests like those in Scheffé's method can be used to determine when the means of some of the sub-groups seem the same and when (with some statistical significance) they should be accepted as different. Friedman's Test for post hoc analysis of contrasts in the non-parametric case will be added to SOTA's repertoire when the preferences of individuals are given as ranks.

Ward (2007) describes the Queen's Executive Decision Centre in Kingston as Canada's first electronic group-decision support laboratory. Discussions have taken place with faculty of the Queen's University School of Business over adapting procedures in the decision laboratory to include methods from SOTA.

Lessons from a Decade of Schools of Thought Analysis

Lesson 1: Acknowledge that there will be dissenting opinions and contrary points of view in any decision-analysis or decision-making group. Appreciate that the contributions of alternative views will strengthen the decision process. Apply appropriate methods to determine the sources of differences and seek to give all views a platform.

Lesson 2: Apply a method to determine the level of consensus or concordance for any specific proposal. When ranks are used by individual participants, Kendall's coefficient of concordance, *W*, or the Friedman Test provide statistical techniques to determine if they differ from random returns (and thus represent some concordance). Report the level of consensus, or the lack of it, to sponsors so they can treat recommendations from a group with appropriate levels of confidence.

Lesson 3: Anticipate that many participants may be reluctant to engage in a decision process with which they are unfamiliar. Engage them in the process, if possible, by applying the techniques in a non-threatening case study, perhaps applying the techniques first to a topic unrelated to the main task.

Lesson 4: Appreciate that *making* a decision is not accomplishing the ultimate objective: implementation of a decision. In many applications, the identification of schools of thought during the planning or decision making can identify those who might take on specific roles during implementation. For example, those who gave strong support to the chosen course of action might be assigned tasks to achieve the chosen objectives. Those who dissented could be given roles to alert the others if or when the chosen course of action is apparently no longer working.

Lesson 5: Advocate the use of critical thinking during the implementation phase. Applying critical thinking during the implementation process brings tremendous value. Like a canary in the coal mine, members of the group who are already sceptical of the group's preferences will be sensitive to circumstances during the implementation phase that indicate that implementation is encountering problems. In management science, the use of devil's advocates in corporate decision making was verified some decades ago as a considerable asset as reported by Daft and Marcic (2009, p. 220). In recent years, military professional debate has proposed measures to institutionalise critical thinking within command teams, popularizing the notion of red teaming described by Craig (2007).

Summary and Conclusion

The Use of Diagnostics in Group Decision Making

This paper describes some diagnostics from Schools of Thought Analysis (SOTA) that have been used in group decision making. For illustration, it provides two cases studies of the application of SOTA.

But more important than the specific diagnostic methods is the importance to an analyst of having *some* form of diagnostic tools that can be applied methodically. Other forms of scientific enquiry, notably those reliant on statistical methods, have a battery of diagnostics that can determine if various prerequisite assumptions are valid. This is particularly critical when the invalidity of some assumption that has not been detected can thereby lead to an erroneous conclusion. For example, many statistical tests rely on an assumption that the data adhere to a normal distribution, so there are diagnostic statistical tests for the data that will determine first if such a critical presumption is valid.

Many decision-analysis procedures dwell on mathematical algorithms for combining individual preferences into a single ordering that will be called the group's preference. However, analysts involved in the process should also employ diagnostics to determine if applying the algorithms is the right approach, or not. If their intended objectives of using rigorous and objective methods have been overwhelmed by underlying pathologies, these analysts need to know.

References

Borg, I. and Groenen, P. (1997) *Modern Multidimensional Scaling*. New York: Springer.

Cameron, F. (1998) *Contrary Schools of Thought within Military Decision-Making Groups*. Schrivenham, UK: International Symposium on Military Operational Research http://ismor.cds.cranfield.ac.uk/ismor1998.htm

Conover, W. (1998) *Practical Nonparametric Statistics*. New York: Wiley. Third Edition.

Craig, S. (2007) 'Reflections from a Red Team Leader' in *Military Review*, March-April. TRADOC Combined Arms Center, Ft Leavenworth, Kansas. 57-60

Daft, R.L., and Marcic, D. (2009) *Understanding Management*. Mason, Ohio: South-Western Cengage Learning, Sixth Edition.

Dalkey, N.C., (1969) *The Delphi Method: An Experimental Study of Group Opinion*. Santa Monica: RAND RM-5888-PR

Everitt, B.S., and Hothorn, T. (2010) *A Handbook of Statistical Analyses Using R*, London, UK: Chapman Hall/CRC, Second Edition.

Janis, I.L. (1972) *Victims of Groupthink*. Boston: Houghton Mifflin Company

Janus, I.L. (1989) *Crucial Decisions: Leadership in Policymaking and Crisis Management*, New York: Free Press

Kendall, M.G. (1975) *Rank Correlation Methods*, New York: Oxford University Press, Third Edition.

Mason, D.W. (1995). *Application of the Consensus Decision Support (CDSP) Methodology in Prioritizing the 1995/96 Major Development Program*, DLOR-RN-9502, Department of National Defence, Ottawa.

Pond, G., and Cameron, F (2010). 'The mathematics of school of thought analysis', in progress.

Ward (2007) 'A new twist on the old talking stick' at http://business.queensu.ca/centres/qedc/ accessed 7 July 2010

Annex: Computer Code in R to Produce Figures 1 and 2

```
> f1 <- file.choose()                  # Choose file produced by Excel
> rr <- read.table(f1,header=TRUE)     # Read table in .txt format
> dd <- as.dist((1 - cor(rr,method="kendall"))/2)
>                                      # Calculate Kendall's tau
>                                      # and convert to distances
> cs = hclust(dd,method="single")      # Cluster analysis - single
> cc = hclust(dd,method="complete")    # Cluster analysis - complete
> ca = hclust(dd,method="average")     # Cluster analysis - average
> isomdsloc = isoMDS(dd,maxit=100,
+   tol=1e-10)                         # Multidimensional scaling
> x = isomdsloc$points[,1]             # Extract x coordinates
> y = isomdsloc$points[,2]             # Extract y coordinates
> windows()                            # Open graphics window
> layout(matrix(c(1,2,3,4), 2, 2,
+        byrow=TRUE))                  # Four plots in one window, 2 by 2
> plot(cs, main="single",sub="",xlab="")
> plot(cc, main="complete",sub="",xlab="")
> plot(ca, main="average",sub="",xlab="")
> plot(x,y,type="n",xlab="",xlim=c(-0.5,0.5),
+          ylab="",ylim=c(-0.5,0.5))  # Draw frame for MDS configuration
> text(x,y,rownames(isomdsloc$points),cex=0.8)
>                                      # Draw identifiers for MDS
```