Operational Research Techniques for Options Analysis in Defence Organizations

W.J. Hurley

Department of Business Administration
Royal Military College of Canada
PO Box 17000, Station Forces
Kingston, Ontario CANDAD
K7K 7B4

DRDC CORA CR 2009-002
November 2009

Defence R&D Canada
Centre for Operational Research & Analysis

National Defence / Défense nationale

Canada

Operational Research Techniques for Options Analysis in Defence Organizations

W. J. Hurley

Department of Business Administration
Royal Military College of Canada
PO Box 17000, Station Forces
Kingston, Ontario CANADA
K7K 7B4

Defence R&D Canada – CORA

Contract Report

DRDC CORA CR 2009-02

November 2009

Principal Author

W. J. Hurley

Approved by

D. Haslip
Section Head, Land & Operational Commands OR

Approved for release by

D. Reding
Chief Scientist

Abstract

This paper examines approaches to options analysis in a defence setting. These include optimization techniques for problems with many options, a modification of Pugh's Method for multi-criteria decision analysis, and median rank voting. The idea is to present techniques that are not well known to the military operational research community. In addition, I discuss the special organizational characteristics of defence decisions.

Résumé

Ce document examine des méthodes qui servent à analyser des options dans le domaine de la défense. Il s'agit notamment de techniques d'optimisation dans le cas de problèmes présentant de nombreuses options, d'une modification de la méthode de Pugh pour l'analyse d'une décision en fonction de critères multiples et de la votation en vue d'un classement médian. On souhaite ainsi présenter des techniques qui ne sont pas bien connues dans le milieu de la recherche opérationnelle militaire. En outre, j'aborde les caractéristiques organisationnelles spéciales des décisions touchant la défense.

Executive summary

Operational Research Techniques for Options Analysis in Defence Organizations

W. J. Hurley; DRDC CORA CR 2009-02; Defence R&D Canada – CORA; November 2009.

This paper is about the application of operational research (OR) techniques to options analyses within a defence organization. An *options analysis* is one where a decision-maker (DM) must select the best option or a best subset of the options under consideration.

I conceive the defence management problem as two ongoing parallel tasks:

1. *Force Employment.* This set of decisions – political, strategic, and tactical – is about how to use the existing force. At one end of the decision spectrum we have the political decision about what operations to initiate and continue; at the other, we have commanders on the ground assessing how to best implement a superior's direction. Options analysis is certainly a useful decision tool in these endeavours. For instance, it is one of the critical steps in the operational planning process.

2. *Force Generation.* This set of decisions includes all of the resource decisions to put the future force structure in place. It includes not only the immediate resource requirements for ongoing operations but also the ongoing analysis to put the correct future force structure in place for five, ten, and twenty years out. Options analysis is useful for these kinds of decisions, particularly those that look at future weapon systems and doctrine.

In this monograph, I intend to focus on options analysis techniques that are useful in the force generation task. In this regard, I restrict myself to two kinds of problems that I think are particularly appropriate for option analysis:

1. *Problems where there are many options.* Most defence problems have the characteristic that there only a small number of options under consideration. However, I have run into a few where the number is quite large. In some of these it is possible to use mathematical programming techniques to come up with good solutions. Here I detail a couple of DND resource allocation problems that can be solved as a knapsack problem.

2. *Problems where there are a limited number of options.* In the case where the number of options is small, one or both of the following two techniques can be useful:

(a) *Voting Methods*. When a group of knowledgeable experts is available and/or time is limited, a voting technique can be used. Each expert would rank the alternatives in order of preference. The problem, then, is to aggregate these individual rank-orders into a consensus rank-order. In some cases, this technique can be used to limit the number of options in preparation for further, more detailed study. For instance, a voting method could be used to finalize the options to be considered in a Multi-Criteria Decision Analysis. In this report I discuss the usefulness of median rank as an aggregation tool. It is rarely considered in the military operations research literature, yet it has some very useful properties.

(b) *Multi-Criteria Decision Analysis (MCDA) Techiques*. This is a very common technique in defence option analysis. The DM would define a set of options and criteria. The relative importance of the criteria and the options on each criterion are measured and subsequently aggregated into relative values for the options. The DM can then choose the best option or a best subset of the options depending on the nature of the original problem. In this report, I detail three methods that have received little attention in the operations research literature: Preference Function Modelling, Pugh's Method, and the Method of Even Swaps. In addition I discuss a number of general aspects of MCDA techniques that are particularly important in some options analysis problems.

Sommaire

Operational Research Techniques for Options Analysis in Defence Organizations

W. J. Hurley ; DRDC CORA CR 2009-02 ; R & D pour la défense Canada – CARO ; novembre 2009.

Le document porte sur l'emploi de techniques de recherche opérationnelle (RO) pour analyser des options au sein d'une organisation de la défense. Dans une analyse des options, un décideur choisit la meilleure possibilité ou le meilleur sous-ensemble parmi les options envisagées. En fait, je présente un résumé de ma réflexion sur la façon dont on peut avoir recours à ces différentes méthodes dans le domaine de la défense.

J'envisage le problème de la gestion de la défense sous la forme de deux tâches parallèles permanentes :

1. Emploi de la force. Ces décisions, politiques, stratégiques et tactiques, portent sur la façon d'utiliser la force existante. À une extrémité du spectre de décision, on retrouve la décision politique sur les opérations à lancer et à poursuivre ; à l'autre extrémité figurent les commandants sur le terrain qui déterminent la meilleure façon de mettre en œuvre la directive de leurs supérieurs. L'analyse des options est certes un outil décisionnel utile à ce chapitre. Par exemple, c'est l'une des étapes critiques du processus de planification opérationnelle.
2. Mise sur pied de la force. Cet ensemble de décisions comprend toutes les décisions sur les ressources qui visent la mise en place de la structure future de la force. Il englobe non seulement les besoins immédiats en ressources pour les opérations en cours, mais également l'analyse constante qui sert à mettre en place une structure de force adéquate pour les 5, 10 et 20 prochaines années. L'analyse des options est utile pour les décisions de la sorte, particulièrement pour celles qui portent sur les systèmes d'arme et la doctrine de l'avenir.

Dans cette monographie, j'aborde les techniques d'analyse des options qui sont utiles à la mise sur pied de la force. À cet égard, je me restreins à deux types de problèmes qui, selon moi, conviennent tout particulièrement à une analyse des options.

1. Problèmes comportant de nombreuses options. La plupart des problèmes de défense ont une caractéristique en commun, soit le fait que seul un petit nombre d'options est envisagé. Toutefois, j'ai constaté que quelques problèmes présentent un nombre élevé d'options. Pour certains de ces problèmes, on peut faire appel à des techniques de programmation mathématique pour en arriver à des solutions

adéquates. Je décris donc deux problèmes d'attribution de ressources du MDN qu'on peut régler à titre de problème d'empilement.

2. Problèmes présentant un nombre restreint d'options. Dans les cas où le nombre d'options est restreint, l'une ou l'autre des techniques ci-après, ou les deux, peuvent s'avérer utiles :

   (a) Méthode de votation. Lorsqu'un groupe d'experts chevronnés est disponible et/ou si le délai dont on dispose est limité, on peut avoir recours à une méthode de votation. Chaque expert classe alors les solutions possibles en ordre de préférence. Puis, le problème consiste à regrouper ces différents rangs de classement en un classement de rangs qui fait consensus. Dans certains cas, on peut faire appel à cette technique pour restreindre le nombre d'options en vue d'une étude supplémentaire plus en profondeur. Par exemple, on peut se servir d'une méthode de votation pour parachever les options à envisager dans le cadre de l'analyse d'une décision en fonction de critères multiples. Dans le rapport, je discute de l'utilité du classement médian comme outil de regroupement. On l'envisage rarement dans la documentation sur la recherche portant sur les opérations militaires, mais il présente des aspects fort utiles.

   (b) Techniques d'analyse d'une décision en fonction de critères multiples. Il s'agit d'une technique très courante pour l'analyse des options dans le domaine de la défense. Le SM précise un ensemble d'options et de critères. L'importance relative des critères, et des options de chaque critère, est évaluée et par la suite adjointe aux valeurs relatives des options. Le SM peut choisir la meilleure option ou encore un sous-ensemble des options, selon la nature du problème initial. Dans le rapport, je décris trois méthodes qui ne sont que très peu abordées dans la documentation sur la recherche portant sur les opérations : la modélisation de la fonction de préférence, la méthode de Pugh et la méthode d'Even Swaps. Je discute en outre de différents aspects généraux des techniques de l'analyse d'une décision en fonction de critères multiples qui sont particulièrement importants pour certains problèmes d'analyse des options.

Table of contents

List of figures

List of tables

This page intentionally left blank.

1 Introduction

This monograph is about the application of operational research (OR) techniques to options analyses within a defence organization. An *options analysis* is one where a decision-maker (DM) must select the best option or a best subset of the options under consideration. For instance, consider a high-level defence committee charged with deciding which capital projects to initiate in the coming fiscal year. This is a difficult decision for a number of reasons. Typically these projects involve significant investments in weapon systems, investments that will not come on stream for years. There are always more projects put up by DND than the government has funds to procure, so given this rationing, the committee must select a subset of these projects to fund. In order to select the best projects, most defence organizations use a numerate approach tempered by sound judgement. This monograph, then, is a summary of my thinking on how these numerate approaches can be applied in a defence setting.

By way of some remarks on heritage, the discipline of operations research has its origin in military problems. Before and during World War II, scientists were asked to apply their methods of analysis to defence operational problems. While there were a significant number of scientists who worked on these problems, the prominent names include Patrick Blackett, Hal Waddington, Philip Morse, George Kimball and Canadians Harold Larnder and Omond Solandt. After the war, if was felt that these same scientific techniques could also be applied to problems in industry. However there were difficulties in establishing academic respectability for the new discipline. These are most evident in the travails of C West Churchman, Russell Ackoff and, to a lesser extent, Peter Checkland (Soft Systems Analysis). Nonetheless, the application of science to defence problems has a rich history and, in my view, these methods are quite appropriate for options analysis in a defence setting.

I conceive the defence management problem as two ongoing parallel tasks:

1. *Force Employment.* This set of decisions – political, strategic, and tactical – is about how to use the existing force. As at July 2009, the Canadian Forces is involved in 18 international operations including Afghanistan. At one end of the decision spectrum we have the political decision about what operations to initiate and continue; at the other, we have commanders on the ground assessing how to best implement a superior's direction. Options analysis is certainly a useful decision tool in these endeavours (for instance options analysis is one of the critical steps in the operational planning process).

2. *Force Generation.* This set of decisions includes all of the resource decisions to put the future force structure in place. It includes not only the immediate resource requirements for ongoing operations but also the ongoing analysis to put the correct future force structure in place for five, ten, and twenty years

out. Options analysis is useful for these kinds of decisions, particularly those that look at future weapon systems and doctrine.

In this monograph, I intend to focus on options analysis techniques that are useful in the force generation task. In this regard, I restrict myself to two kinds of problems that I think are particularly appropriate for option analysis:

1. *Problems where there are many options.* Most defence problems have the characteristic that there only a small number of options under consideration. However, I have run into a few where the number is quite large. In some of these it is possible to use mathematical programming techniques to come up with good solutions. Here I detail a couple of DND resource allocation problems that can be solved as a knapsack problem.

2. *Problems where there are a limited number of options.* In the case where the number of options is small, one or both of the following two techniques can be useful:

   (a) *Voting Methods.* When a group of knowledgeable experts is available and/or time is limited, a voting technique can be used. Each expert would rank the alternatives in order of preference. The problem, then, is to aggregate these individual rank-orders into a consensus rank-order. In some cases, this technique can be used to limit the number of options in preparation for further, more detailed study. For instance, a voting method could be used to finalize the options to be considered in a Multi-Criteria Decision Analysis. In this report I discuss the usefulness of median rank as an aggregation tool. It is rarely considered in the operations research literature, yet it has some very useful properties as we shall see.

   (b) *Multi-Criteria Decision Analysis (MCDA) Techiques.* This is a very common technique in defence option analysis. The DM would define a set of options and criteria. The relative importance of the criteria and the options on each criterion are measured and subsequently aggregated into relative values for the options. The DM can then choose the best option or a best subset of the options depending on the nature of the original problem. In this report, I detail three methods that have received little attention in the operations research literature: Preference Function Modelling, Pugh's Method, and the Method of Even Swaps. In addition I discuss a number of general aspects of MCDA techniques that are particularly important in some options analysis problems.

2 Optimization Approaches

In some cases, the DM will have to consider a large number of options. Obviously looking at such a large number of options using a full-scale MCDA would be time-consuming simply because generating values for each criterion for each option would take considerable time. However there are ways to approximate values for options that economize on the measurement problem. In section 2.3, I show how this can be done in the context of using standard integer programming approaches for capital rationing. A parametric approach is taken as follows. Suppose that the options are first rank-ordered. Rather than have the DM assign a value for each option, I assume that these values are based on a function of rank-order with well defined parameters. It is then simply a matter of having the DM estimate the function parameters rather than values for all options.

In my experience in the Canadian defence environment, integer programming is rarely used for capital rationing. The usual approach is to rank projects in order of priority and then pick as many projects at the top of this list as the budget allows. This is a reasonable approach when the value of a project is highly correlated with its investment (expenditure). However, in a defence context, more often than not, value is only loosely tied to investment. Hence it makes sense to consider approaches that take into account both value and expenditure.

To demonstrate the approach, I use the example of how the Canadian Air Force allocates a fixed budget for Miscellaneous Requirements (MR) projects (defined as projects that are under 5 million dollars and can typically be completed within the current fiscal year). Projects are originated at the base/unit level and sent up to 1 Canadian Air Division, the parent organization. Over a three-day conference, representatives from the 12 bases and units, 1 Canadian Air Division, and the Directorate of Air Requirements determine which projects will be funded. They do this in the following way: projects are first put in order of priority based on agreed criteria and without regard for the expenditure a project requires; once projects are ordered, those at the top of the list are selected until the budget is exhausted. I term this approach the *Priority/Rationing Method*.

This is certainly not the only instance where a Priority/Rationing Method is used in a Canadian defence setting. For instance, I have seen it employed in the planning for Army Individual Training courses where courses are first put in order of priority and then those at the top of this list are selected until either the supply of training staff or ammunition is exhausted.

In principle, the extension of a Priority/Rationing Method to an optimization method should be easy. Once projects are rank-ordered by value, a second step would be to attach an explicit value to each project. With these values in hand, planners could then solve a knapsack problem where value is maximized subject to a budget constraint. However, in the context of the Miscellaneous Requirements allocation problem, there are some practical difficulties. One is that there are a large number of projects to value (142 MR projects were considered for 2006). Putting an explicit value on all 142 projects would just take too much time, especially in a room where there are vested interests. Hence, in this report, I suggest a simple way to transform the ordinal rankings of projects to values.

The second contribution is to compare the increase in value that an optimization technique provides over the Priority/Rationing Method. In the context of the MR selection problem, I conservatively estimate the percentage increase in value to be about 20%.

The literature on the application and solution of knapsack problems to project selection is now immense and dates to some very good early papers. One is by Weingartner (1966). Martello and Toth (1990) emphasize solution algorithms. In addition, there is a significant literature on the application of optimization to resource allocation problems within the US military. Brown, Dell and Newman (2004) have reviewed this work. A small sampling would include Baker et al (2002), Brown et al (1991), Brown et al (1994), Brown et al (2003), Evans and Fairbairn (1989), Loerch et al (1999), and Ewing et al (2006). All of these papers model defence resource allocation problems with binary choice variables. By comparison, the use of these types of models by Canadian defence planners is relatively sparse.

2.1 The Knapsack Approach

Considering the MR budget allocation problem described above, suppose v_{ij} is the value of project i originated by base/unit j. Let E_{ij} be the expenditure associated with this project. Suppose there are m bases/units and n_j projects are originated at base/unit j. Let $x_{ij} = 1$ if project i originating in base/unit j is selected and 0 otherwise. Suppose a total of B dollars are available for these projects. We assume there are no interdependencies among projects. If there are, these can be handled in the standard ways. For instance if projects 1 and 2 at base j are mutually exclusive, then we would impose the constraint $x_{1j} + x_{2j} = 1$.

Under these assumptions the knapsack program to solve for the highest value portfolio is

$$\begin{aligned} \max \quad & \textstyle\sum_{j=1}^{m} \sum_{i=1}^{n_j} v_{ij} x_{ij} \\ \text{s.t.} \quad & \textstyle\sum_{j=1}^{m} \sum_{i=1}^{n_j} E_{ij} x_{ij} \leq B \\ & x_{ij} \text{ binary for all } i, j. \end{aligned}$$

The objective function is the portfolio's value. The constraint requires that the expenditures on the complete portfolio cannot exceed the total budget. We term this approach to MR project selection the *Knapsack Approach*. Let the optimal solution be x^*_{ij} and let

$$V^* = \sum_{j=1}^{m} \sum_{i=1}^{n_j} v_{ij} x^*_{ij} \tag{1}$$

be the optimal value of the objective function at this solution.

Due to the political nature of the MR project funding process, it would be useful to identify a number of desirable solutions. One way to do this would be to solve for the second-best solution:

$$\begin{aligned} \max \quad & \textstyle\sum_{j=1}^{m} \sum_{i=1}^{n_j} v_{ij} x_{ij} \\ \text{s.t.} \quad & \textstyle\sum_{j=1}^{m} \sum_{i=1}^{n_j} E_{ij} x_{ij} \leq B \\ & \textstyle\sum_{j=1}^{m} \sum_{i=1}^{n_j} x^*_{ij} x_{ij} \leq \sum_{j=1}^{m} \sum_{i=1}^{n_j} x^*_{ij} - 1 \\ & x_{ij} \text{ binary for all i,j.} \end{aligned}$$

The second constraint assures that the first-best portfolio cannot be chosen. More generally, suppose we want to find the K-th best solution. Let $s^{(p)}_{ij}$ be the p-th best solution. To get the K-th best solution, I would solve

$$\begin{aligned} \max \quad & \textstyle\sum_{j=1}^{m} \sum_{i=1}^{n_j} v_{ij} x_{ij} \\ \text{s.t.} \quad & \textstyle\sum_{j=1}^{m} \sum_{i=1}^{n_j} E_{ij} x_{ij} \leq B \\ & \textstyle\sum_{j=1}^{m} \sum_{i=1}^{n_j} s^{(p)}_{ij} x_{ij} \leq \sum_{j=1}^{m} \sum_{i=1}^{n_j} s^{(p)}_{ij} - 1, \text{ for } p = 1, 2, ..., K-1 \\ & x_{ij} \text{ binary for all } i, j. \end{aligned}$$

Again the set of constraints, $\sum_{j=1}^{m} \sum_{i=1}^{n_j} s^{(p)}_{ij} x_{ij} \leq \sum_{j=1}^{m} \sum_{i=1}^{n_j} s^{(p)}_{ij} - 1$, makes sure that none of the first $K-1$ best solutions can be selected. The optimal solution of this program is $x^*_{ij} = s^{(K)}_{ij}$. We term this the *K-th Best Solution Program*.

For knapsack problems it is well known that small increases in budget can lead to large increases in objective function value. See, for example, Eilon (1987). Hence, in the context of the MR budget allocation problem described above, it makes sense to determine whether a small violation of the budget constraint will add significant value. To do this, I solve

$$
\begin{aligned}
\min \quad & \Delta \\
\text{s.t.} \quad & \sum_{j=1}^{m} \sum_{i=1}^{n_j} E_{ij} x_{ij} - \Delta \le B \\
& \sum_{j=1}^{m} \sum_{i=1}^{n_j} v_{ij} x_{ij} \ge V^* \\
& \sum_{j=1}^{m} \sum_{i=1}^{n_j} x^*_{ij} x_{ij} \le \sum_{j=1}^{m} \sum_{i=1}^{n_j} x^*_{ij} - 1 \\
& \Delta \ge 0, x_{ij} \text{ binary for all } i, j.
\end{aligned}
$$

This program will produce a higher value portfolio of projects at minimal increase in the supply of funds, Δ. I term this program the *Relaxed Knapsack Approach*.

As for complexity, it is well known that Knapsack Problems are NP Hard. That said, the problem I study here has some 150 binary variables and is easily solved on a PC with a commercial integer code.

2.2 A Simple Example

Suppose we are considering five projects with the characteristics shown in the following table:

Project	**Expenditure**	**Value**
1	50	125
2	35	105
3	10	100
4	8	60
5	5	20

In addition, suppose we have 90 to spend. Since the projects are in order of their values, the Priority/Rationing Method would choose projects 1 and 2 with total expenditure 85 and total value 230.

On the other hand, if the integer program of the Knapsack Approach is solved, the solution is to adopt projects 1, 3, 4, and 5 with expenditure 73 and value 305.

2.3 Transforming a Rank Order to Value Quickly

As described above, the current MR budget allocation process includes a step where all projects are put in a priority ordering: Project 1 is at least as important as Project 2, Project 2 is at least as important as Project 3, and so on. We need to take this ordering and map it into a set of cardinal values in order to solve the MR Knapsack Problems defined above. Moreover, we will demonstrate how this can be done fairly quickly with minimal input from planners.

Suppose there are n projects with values $v_1, v_2, ..., v_n$ in rank-order so that

$$v_1 \geq v_2 \geq ... \geq v_n. \quad (2)$$

To measure these values, we now turn to the theory of value functions as presented in Keeney and Raiffa (1976) and developed in Dyer and Sarin (1979). A *measurable value function* is one that measures the strength of preference between pairs of alternatives. For our purposes, we define a *relative value function* as one based on the DM's ability to measure the ratio of any two project values. That is, given the values of projects i and j, v_i and v_j respectively, the DM is able to measure the ratio v_i/v_j. This, of course, is a fundamental measurement property of the Analytic Hierarchy Process (Saaty, 1980) and other decision analysis techniques.

Suppose the DM assesses enough of these ratios to arrive at a vector of values,

$$\mathbf{v}^T = (v_1, v_2, ..., v_n) \quad (3)$$

where the superscript T denotes transpose. For instance the DM could assess the ratios

$$\frac{v_1}{v_2} = a_2, \frac{v_1}{v_3} = a_3, ..., \frac{v_1}{v_n} = a_n \quad (4)$$

and then fix one of the values, say $v_1 = c$, in order to get the values

$$v_i = c/a_i \text{ for } i = 2, 3, ..., n. \quad (5)$$

With these values, the DM intends to solve a knapsack problem with objective function

$$\mathbf{v}^T\mathbf{x} = \sum_i v_i x_i \quad (6)$$

where

$$\mathbf{x}^T = (x_1, x_2, ..., x_n) \quad (7)$$

is a vector of binary decision variables. Then any positive linear transformation of these values, say $\mathbf{u} = b\mathbf{v}$, where $b > 0$ is a constant, will produce the same solution for the knapsack problem since

$$\sum_i u_i x_i = \sum_i b v_i x_i = b \sum_i v_i x_i. \quad (8)$$

This property allows us to fix one of the project values at any value we want (as we did above for v_1) and not change the solution of the knapsack problem.

But given that the MR budget allocation problem involves a large number of projects and that we wish to take minimal input from the DMs, we need to come up with a simple relationship between project value and project rank,

$$\widehat{v}_i = f(i), \quad i = 1, 2, ..., n. \quad (9)$$

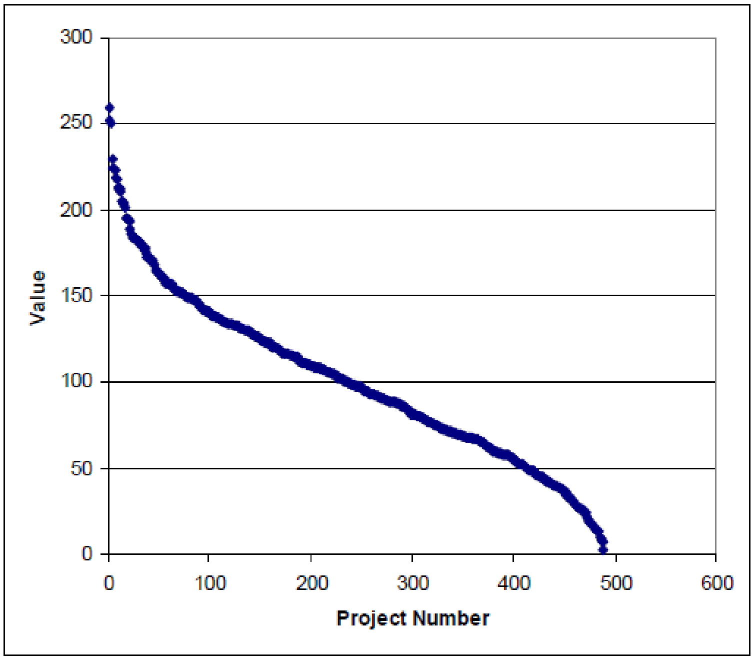

Figure 1: A value curve based on the normal distribution.

The economy is that, rather than measuring $n-1$ ratios, we simply measure the parameters of the function f.

Clearly, f must be a non-increasing function. Within the theory of measurable value functions, Kirkwood and Sarin (1980) suggest an exponential form. In the context of the MR budget allocation problem, we might make the following argument that such a form is reasonable. Suppose that the generation of n MR project values is equivalent to drawing an iid sample of size n from a particular probability distribution. These projects could then be ordered by value and the shape of the resulting curve could be examined.

Suppose values are drawn from a normal distribution. The resulting value curve is shown in Figure 1. Note that it takes a backwards "S" shape and that, for the highest value projects, the curve decreases quite sharply and then levels out. Another possibility is the exponential distribution. Its value curve is shown in Figure 2. Again, note the steep decline in value followed by a leveling out. Both of these curves are consistent with the reality that it ought to be increasingly difficult to find high value projects. In fact, if projects are drawn from a probability density function with a shrinking right tail (as with the normal and exponential distributions), the functional form will follow a downward sloping convex form. For these reasons, I propose the

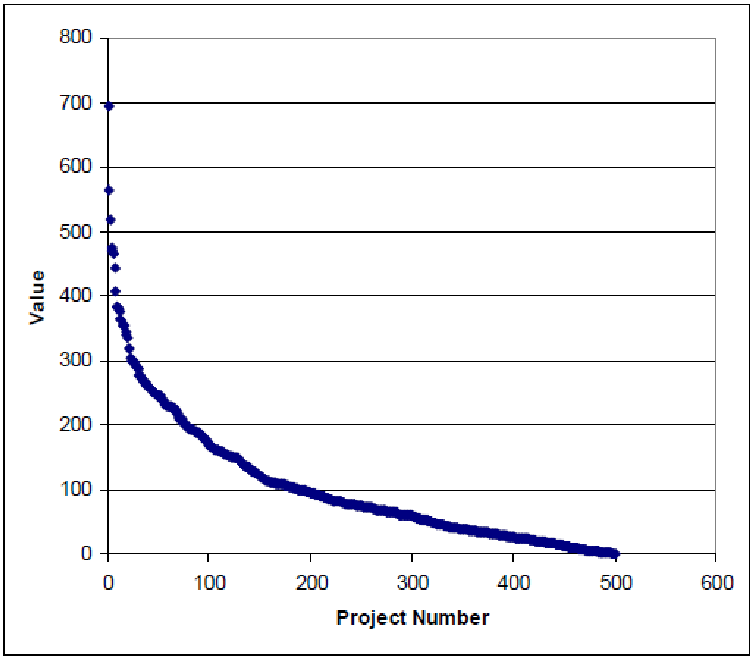

Figure 2: A value curve based on the exponential distribution.

following *power value function*:

$$v_i = v_1(1-d)^{i-1} \qquad \text{for } i = 1, 2, 3, ..., n \text{ and } 0 < d < 1. \tag{10}$$

An example of this curve is shown in Figure 3. There are many other value functions that would do. I have chosen this one because it has a simple parametric structure that is easy to interpret.

There are a number of ways we could estimate the parameters v_1 and d depending upon how much input the DM is prepared to give. Assuming as we have done previously that the DM wants minimal input, we first ask him to assume that the median project (we denote its index i_{med}) has a value of 100. We then ask him to consider the highest value project relative to the median project. In particular, we ask him to estimate the ratio $v_1/v_{i_{med}}$. Suppose he responds that $v_1/v_{i_{med}} = a$. Then the parameter estimates consistent with $v_1/v_{i_{med}} = a$ and $v_{i_{med}} = 100$ are

$$d = 1 - \left(\frac{1}{a}\right)^{\frac{1}{i_{med}-1}} \tag{11}$$

$$v_1 = 100a. \tag{12}$$

Another approach would be to ask the DM for a number of these ratio judgements.

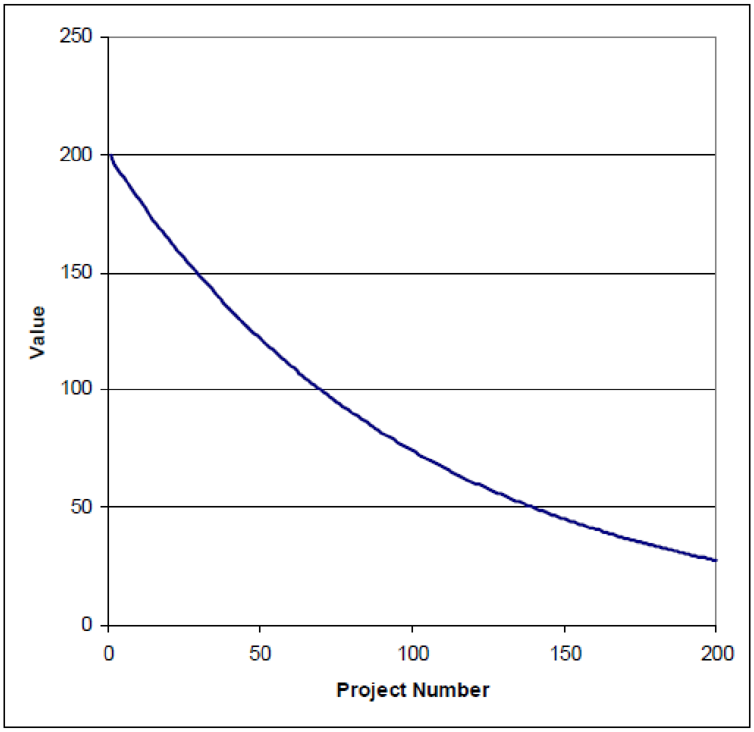

Figure 3: An example of the power functional form.

If two or more are solicited, the system is overspecified and a least squares procedure can be used to obtain estimates for d and v_1.

Finally, we point out that a comparison of the percentage change in objective function values for the Knapsack and Priority/Rationing methods is meaningful in the context of relative value functions. Let $\mathbf{x}^*$ and $\mathbf{x}^P$ solve these problems respectively. For a given vector of values, $\mathbf{v}$, the percentage increase in objective function value that the Knapsack Method provides depends directly on the ratio

$$\frac{\mathbf{v}^T\mathbf{x}^*}{\mathbf{v}^T\mathbf{x}^P}. \tag{13}$$

But any other vector of relative values, $\mathbf{u}$, satisfying $\mathbf{u} = b\mathbf{v}$, $b > 0$, produces the same solution $\mathbf{x}^*$. Note that the percentage increase in value under $\mathbf{u}$ depends on

$$\frac{\mathbf{u}^T\mathbf{x}^*}{\mathbf{u}^T\mathbf{x}^P} = \frac{(b\mathbf{v})^T\mathbf{x}^*}{(b\mathbf{v})^T\mathbf{x}^P} = \frac{\mathbf{v}^T\mathbf{x}^*}{\mathbf{v}^T\mathbf{x}^P}. \tag{14}$$

Hence the Knapsack Method gives a percentage increase in value that is invariant to the particular relative value function used.

2.4 The Air Force MR Example

In total, 142 MR projects were considered for fiscal 2006. These ranged in expenditure from $4,500 to $1.96 million. Undertaking all 142 projects would require an expenditure of $28.5 million.

In examining the priority of these projects as assigned by the conference, it is interesting to note that the costs of projects are not highly correlated with priority. The following table presents information on the projects ranked 45 through 47:

Priority	Description	Cost
45	Operational Training Equipment	$19,147.00
46	Maintenance Stands	$657,000.00
47	Tooling Acquisition	$173,450.00

Note that the cost required for the rank 45 project is considerably less than that for rank 46, yet 45 is ranked higher. The overall correlation of expenditure with the rank order is just under 30%, indicating in general that the expenditure on a project is not highly correlated with its value.

We were unable to get the exact amount allocated for MR projects but we estimate it to be about $20 million. Under this assumption, the Priority/Rationing Method would select projects ranked 1 through 84 and these 84 projects would cost $19,993,798.

We then used the Knapsack Approach with the value function developed in the previous section. Initially we did this selection with $a = 2.0$, a conservative assessment based on our professional judgement and experience. The results are shown in the following table along with those for the Priority/Rationing Method:

	Priority/Rationing	Knapsack
Total Value	11,463	14,422
Total Expenditure	$19,993,798	$19,993,874
Number of Projects Adopted	84	134

Obviously, the Total Value is going to be higher using the Knapsack Method. With $a = 2.0$, Total Value is some 25% higher. In addition, the Knapsack Method results in more projects being adopted (84 for Priority/Rationing and 134 for Knapsack). Of the 84 projects that the Priority/Rationing Method recommends, the Knapsack Method selects 80 of these. The expenditures required for the 4 projects that the Knapsack Method drops are then used to fund 54 more projects. It is worth noting that this increase in the number of projects is likely to please a majority of the MR conference participants.

2.5 Sensitivity

The Knapsack Method gives a total expenditure that is just under the $20 million budget. We solved the Relaxed Knapsack Program to see how things would change if the budget constraint were minimally violated. For an additional $7,774, the Total Value goes to 14,433, a small increase. This assures us that a small violation of the budget constraint will not result in a large increase in value.

Obviously, it is important to do a sensitivity analysis on the value of a. One could take the following approach with the DM. Suppose we ask him to specify an uncertainty interval about his estimate a and suppose he responds with $[a - \Delta, a + \Delta]$. The Knapsack program could then be solved for the two end-points of this interval to see how the solution changes. For the instance defined above, we employed $\Delta = 0.5$, and again, this assessment is based on our professional judgement and experience.

The optimal portfolio of projects changed minimally for $a - \Delta$ and $a + \Delta$. At $a - \Delta = 1.5$, the objective function value for the Knapsack Method was some 38% higher than it was for the Priority/Rationing Method; at $a + \Delta = 2.5$, it was 19% higher. Assuming that the true value of a is somewhere in the range $[1.5, 2.5]$, the Knapsack Method provides at least a 20% increase in value over the Priority/Rationing Method for this instance of the MR budget allocation problem.

2.6 Concluding Remarks

In the case where there are a large number of options from which to select, a DM should consider using an integer programming approach rather rank-ordering the options and then selecting those at the top of the list for which there is enough budget. In the case of Air Force Miscellaneous Requirements projects, I showed that the approach integer programming can save approximately 20% over the ad hoc approach.

This approach does a couple of other things. Suppose we term the set of options which solves the knapsack problem an optimal *portfolio*. Then I show how to generate second-best, third-best, up to mth-best portfolios. The mth best solution is obtained by adding constraints that do not allow the program to choose the first-best, second-best, up to the $(m - 1)$th-best portfolios. This can be helpful when, for good reasons, the DM chooses not to model all of the actual constraints.

Finally, I developed a rough-and-ready technique for translating a rank-order of the options into a valuation for each option. This does not take much effort on the part of the DM to implement. Moreover, the robustness of the solution can be checked with some sensitivity analysis.

3 The Ordinal Consensus Ranking Problem

When the number of options under consideration is finite and small, sometimes a group of experts can be asked to rank-order the options on the basis of some sort of organized discussion of options, criteria and values. These individual rank-orders are then aggregated to determine which option is best. This problem is known as the Ordinal Consensus Ranking Problem. The interesting methodological issue is how to aggregate these individual rank-orders. Many methods have been suggested. In this section I argue that the median-rank is an interesting alternative to more computationally intensive techniques. In addition, it is quite robust to strategic voting. Most of the ideas in this section are taken from Hurley (2002).

This is certainly not a new problem. It dates to the work of Condorcet (1785) and Borda (1781). The literature is now immense. Important contributions include the work of Kendall (1962), Cook and Seifort (1982), Cook, et al (1997), Kemeny and Snell (1962), Inada (1969), Cook and Kress (1985), and Beck and Lin (1983). In most of these methods, the problem is to find a consensus order which, by some metric, is as close as possible to the individual rank-orders.

Of these, the contribution of Emond and Mason (EM, 2003) is particularly important. EM generalize Kendall's τ to their so-called τ_x measure. This enables them to handle input orderings that allow ties, that are incomplete, and that allow weighted input rankings. Their methodology is supported by a DND software implementation called Fundamental Investigation of Defence Objectives (FIDO). The one difficulty with their approach is that the computational requirements can be severe if the number of options is large.

In practise, these consensus ranking problems have two important characteristics. The first is measurement error: I assume that the options under consideration have an objective true value and our experts can only observe an imperfect signal of this value. Hence our expert judges may arrive at different rank-orders due to measurement error. The second characteristic, which may or may not be present in a particular application, is strategic voting: some members of the group may give "their" alternative a better ranking than one they know is in the best interest of the organization.

There are any number of organizational settings where this kind of strategic voting is a problem. For instance, take the example of judging an international figure skating competition. Does the scoring system mitigate the effects of those judges who have a preference, either conscious or subconscious, for their country's skater? Another example is the way in which the Canadian Army determines annually which of its Miscellaneous Requirements projects will be undertaken. A board made up of officers from the various Land arms is convened and charged with the task of rank-ordering all projects. There is certainly ‘ownership’ of projects at this table. An artillery officer will not see armored projects the same way an armored officer sees them.

Related to the issue of strategic voting, the work of Bassett and Persky (BP, 1994) is particularly interesting. BP study the way figure skating competitions are judged. They point out that the rank-order of the skaters is determined by median rank voting, a system that has some interesting properties in the presence of strategic voting. One is that it is ‘resistant to manipulation by a minority subset of judges ...’ (page 1075). In view of the fact that other sports use different aggregation techniques (for instance, diving and gymnastics use a trimmed mean), it is reasonable to ask which of these trimmed mean vote aggregation techniques is preferred in the case where there is strategic voting. Yaniv (1997) also considers trimmed mean as an aggregation tool but does not consider strategic voting explicitly.

So, in this section, I undertake a Monte Carlo study of the performance of Trimmed Mean Consensus Ranking Procedures in the presence of measurement error and strategic voting. I denote by $Trim(k)$ the trimmed mean rank-order which throws out the best k ranks and the worst k ranks for each alternative. Note that the Average Rank and Median Rank measures are special cases of a trimmed mean. Average Rank, denoted $Trim(0)$, is the case where no observations are thrown out. At the other extreme, Median Rank throws out all but one observation. For instance, if there are 7 rank observations (7 judges), $Trim(3)$ is the median rank-order.

The simulation results suggest that, in the presence of strategic voting, the Median Rank procedure is superior to other trimmed mean procedures including the Average Rank procedure. And in the absence of strategic voting, it does no worse than these other procedures. This evidence suggests that Median Rank Aggregation is a reasonable aggregation tool.

These results have an analog in the theory of robust estimation. For instance, consider a random sample from an unknown symmetric distribution and suppose it is necessary to estimate the median of the distribution. If the underlying distribution were normal, then a maximum likelihood estimate of the median is the sample average. However, if the underlying distribution has fat tails (such as the Cauchy), a trimmed mean will produce a lower mean square error than the sample average. Hence, for some fat-tailed distributions, a trimmed mean is preferred to a sample average. The interested reader is referred to Degroot (1975) for a discussion of this point.

3.1 Median Rank Aggregation and Strategic Voting

To give some idea of the way median rank voting handles strategic voting, consider the following example taken from Hurley (1998). Suppose that all-stars in a football conference are selected by representatives of the teams making up the conference.

	Rank Average	*Truncated Average Rank*	*Median Rank*
a	1.857	1.8	1
b	1.571	1.6	2
c	2.571	2.6	3

Table 1: Results for the various aggregation rules

Each team ranks the players nominated at a position, and the best player is to be selected on the basis of median rank. Suppose the following rank-orders have been submitted for the quarterback position:

	Teams						
QBs	*A*	*B*	*C*	*D*	*E*	*F*	*G*
a	1	1	1	3	3	1	1
b	2	2	2	1	1	2	2
c	3	3	3	2	2	3	3

where a is team A's quarterback, b is team B's, and c is team C's. Moreover, suppose that these rankings are each team's true unbiased rankings. The first step is to simply write the ranks for each quarterback, in order, from highest to lowest:

QBs				*median*			
a	1	1	1	1	1	3	3
b	1	1	2	2	2	2	2
c	2	2	3	3	3	3	3

The median-rank, then, is the rank of the middle observation, or, in this case, the fourth highest number in the list. Hence, QB a has a median-rank of 1, QB b has a median-rank of 2, and QB c has a median-rank of 3. Therefore, the median rank-order is $a \succ b \succ c$. If it happened that two players had a median rank of 1, there are a number of tie-breaking mechanisms. Some of these are set out in BP.

To see how the median-rank scheme handles strategic voting, suppose that team B feels that the rank order is $a \succ b \succ c$, but, in the interests of having their quarterback be the all-star, they submit the ranking $b \succ c \succ a$. The results for the median rank aggregation technique and two others (average rank and trimmed mean rank where a player's best and worst ranks are thrown out) are shown in Table 1. The median rank scheme returns the correct rank order, $a \succ b \succ c$, whereas the average-rank and truncated-average-rank return $b \succ a \succ c$. This is no more than the basic BP result that Median Rank aggregation is immune to manipulation by a minority of judges.

3.2 The Simulation Design

In the absence of strategic voting, the simulation works in this way. Suppose there are N objects and M judges. The objective, true values of the objects, $v_1, v_2, ..., v_N$, are selected randomly from the unit interval, $[0, 1]$. The *unbiased assessment* of judge j for alternative i is then drawn from a normal distribution with mean v_i and variance σ_i^2. Once judge j obtains an unbiased assessment for all objects, he can rank-order the objects. We do this for each of the judges. These are then aggregated using $Trim(k)$ for various values of k. Effectively we are modelling measurement error on the part of the judges.

To measure how well each aggregation technique does, we measure the frequency with which each picks:

1. (*Unordered Objective*) the best L alternatives *without* regard for rank-order;
2. (*Ordered Objective*) the best L alternatives in the proper rank-order.

For the *Unordered Objective*, we are only concerned that a vote aggregation technique pick the top L objects; it does not matter what order these top L objects end up in, only that they make the top L spots in the list. The *Ordered Objective* is more specific. Not only do the top L objects have to be picked, but they also have to appear in the correct order.

We introduce strategic voting as follows. We assume that each judge will vote strategically with probability p. Effectively p controls the "amount" of strategic voting: the expected number of judges voting strategically is Mp. If a judge is going to vote strategically, he will do so in an extreme way. Suppose the judge's unbiased assessment of three objects is $A \succ B \succ C$ but, for other reasons, he would prefer that object B is selected (he is an artillery officer and A is an armored project). Then we assume that he will submit the ranking $B \succ C \succ A$. That is he puts his object first, and the remaining objects in reverse order. He reasons that this gives his project the best chance of succeeding. We term this rank-order the *biased assessment,* and in the case at hand, the judge is said to have ownership of object B (the personally preferred choice). While this form of strategic voting is particularly perverse, we only introduce it in this way to give an extreme version. In reality strategic voting is likely to be more subtle, especially if the judge has a reputation to maintain.

In summary, the steps of the simulation are:

1. Generate a set of true values, $v_1, v_2, ..., v_N$, randomly from $[0, 1]$.
2. Generate an unbiased assessment for each judge.

3. For each judge, generate a biased assessment with probability p. (If a judge is biased, ownership of an alternative is assigned randomly.)
4. Aggregate the assessments using $Trim(k)$ for various values of k, and for each, determine whether the Unordered and Ordered objectives are satisfied.

Before reporting our experimental results, it is important to clarify why judges vote strategically. As we have argued above, organizational DMs are sometimes of two minds on what the best course of action is. For instance, in the context of choosing military projects, an armoured officer may be presented with objective, rational evidence that an artillery project has higher value than the armoured project he has put up. A part of him says that he ought to vote for the artillery project because it is best for the force structure as a whole. Another part says he ought to vote for the armoured project because it is "his" project. In addition to this motivation for strategic voting, each judge observes an imperfect, unbiased signal of each alternative's value. Non-strategic-voters are assumed to submit a rank-order based on these signals; strategic voters submit the perverse rank-order described above.

3.3 Results

For the first set of simulations, we fix the following parameters:

Number of Objects	7
Number of Judges	7
Standard Deviation (σ_i)	0.10
Number of Simulation Iterations	100,000

What we vary is the probability, p, that judges vote strategically.

The frequencies for the *Ordered Objective* for various values of the strategic voting parameter, p, are shown in Table 2 and for the *Unordered Objective*, in Table 3. For instance, consider the element for $p = 0$, $L = 2$, and $Trim(0)$ in Table 2: .714. This means that, in 71.4% of the 100,000 iterations, the average rank aggregation procedure picked the top 2 elements in the correct order, when there was no strategic voting. Similarly, in Table 3, for $p = .20$, $L = 3$, and $Trim(3)$, the element is .731. In this case, the median rank aggregation procedure is able to pick the top 3 elements without regard for order in 73.1% of the 100,000 iterations, when there was a 20% probability of strategic voting.

We first consider the ability of each aggregation method to find the best object ($L = 1$). In the absence of strategic voting ($p = 0$) we would expect the Average Rank procedure ($Trim(0)$) to perform the best. However it does not. In fact, it appears as though all of the aggregation procedures are statistically equivalent. When strategic

L	1	2	3	4	5	6	7
				$p = 0$			
$Trim(0)$	.844	.714	.600	.503	.420	.354	.354
$Trim(1)$	.843	.710	.596	.498	.414	.348	.348
$Trim(2)$	.844	.710	.595	.497	.412	.347	.347
$Trim(3)$	.844	.710	.595	.497	.413	.347	.347
				$p = .20$			
$Trim(0)$	.717	.541	.423	.338	.272	.221	.221
$Trim(1)$	.746	.572	.449	.363	.295	.245	.245
$Trim(2)$	.772	.596	.465	.376	.306	.255	.255
$Trim(3)$	.781	.603	.470	.380	.309	.257	.257
				$p = .40$			
$Trim(0)$	.492	.299	.208	.157	.122	.096	.096
$Trim(1)$	.527	.329	.232	.179	.141	.115	.115
$Trim(2)$	.573	.360	.253	.198	.159	.131	.131
$Trim(3)$	.612	.392	.273	.213	.171	.141	.141

Table 2: Results for the Ordered Objective

voting is introduced ($p = .20, .40$), the trimmed mean procedures outperform the Average Rank procedure. Of these, the Median Rank procedure is best. To give some idea of the error in these numbers, a 95% confidence interval for median rank frequency of .781 ($Trim(3), p = .20$) is $[.778, .784]$. Note that when there is excessive strategic voting ($p = .40$), $Trim(3)$ picks the best alternative 61.2% of the time and $Trim(0)$ picks it 49.2%. In the presence of strategic voting, the Median Rank procedure is clearly better.

Now consider the case where the objective is to identify the best L objects without regard for order. For instance, a voting method might be used to prescreen a set of investment alternatives: the top L alternatives will be given further consideration. To assess the methods on this objective, consider the output in Table 3. Note that, in the absence of strategic voting (the panel with $p = 0$), the trimmed means do as well as Average Rank for all values of L. On the other hand, in the presence of strategic voting ($p = .20, .40$), the trimmed means outperform Average Rank, and of these Median Rank does the best.

I have done an extensive sensitivity analysis of this example, varying the number of judges, the number of alternatives, and the standard deviation, σ_i. In each case the same result obtains: with only measurement error, Median Rank does as well as the other trimmed mean and Average Rank measures; with measurement error and strategic voting, Median Rank outperforms these other aggregation procedures. It is important to understand that this example is in no way a proof that Median Rank

L	1	2	3	4	5	6	7
				$p = 0$			
$Trim(0)$	.844	.837	.836	.838	.835	.844	1.000
$Trim(1)$	.843	.834	.834	.835	.833	.843	1.000
$Trim(2)$	.844	.833	.832	.833	.831	.842	1.000
$Trim(3)$	.844	.833	.832	.832	.831	.842	1.000
				$p = .20$			
$Trim(0)$	.717	.708	.726	.749	.785	.807	1.000
$Trim(1)$	.746	.727	.733	.755	.794	.828	1.000
$Trim(2)$	.772	.736	.731	.761	.798	.834	1.000
$Trim(3)$	.781	.742	.731	.760	.800	.834	1.000
				$p = .40$			
$Trim(0)$	.492	.449	.475	.527	.656	.774	1.000
$Trim(1)$	.527	.471	.483	.531	.659	.795	1.000
$Trim(2)$	.573	.488	.482	.542	.666	.810	1.000
$Trim(3)$	.612	.512	.486	.540	.679	.820	1.000

Table 3: Results for the Unordered Objective

would always be superior. Our results are based only on a simulation for a specific parameter set. Future research will look at the issue more generally.

3.4 Conclusions

I have presented results of a Monte Carlo simulation which suggests that, in the presence of strategic voting, Median Rank aggregation works better than other trimmed mean procedures and the Average Rank procedure. In view of the fact that all of these aggregation procedures are about the same when there is no strategic voting, the firm conclusion of this simulation example is that Median Rank voting is a robust method for aggregating individual rank orders. It would be interesting to compare Median Rank voting to a broader set of aggregation methods. In addition, it needs to be extended to be able to handle incomplete rankings, weighted rankings, and input orderings that allow ties.

4 Multi-Criteria Methods

As discussed in the introduction, it is rarely the case that decision options are evaluated on one criterion, especially in military settings. For instance, where military planners are determining which configuration of an armoured combat vehicle (ACV) to recommend to a senior decision board, a number of criteria must be considered. These might include such things as the ACV's firepower, its protection (armour), mobility and cost. If a single ACV were best on all of these criteria, the decision would be easy. However this is rarely the case. Generally, to get more of a particular criteria, you have to give something up on one or more of the others. For instance to get a more mobile ACV, you might have to reduce the vehicle's overall weight by reducing its armour. To examine these kinds of tradeoffs, DMs sometimes employ a MCDA.

There has been much discussion in the operations research literature about the language of MCDA. For my purposes here, I will define it as a process comprising three stages:

1. *Model Formulation Stage*: The DM begins with a problem (ie its time to replace the LAV III). Next a number of important criteria (ie must be mobile, must provide protection, etc) and options that could solve the problem are generated. The end result of this exercise is a list of *options* and a list of *criteria*. This is generally considered to be the toughest part of the MCDA process. One of the primary benefits of this stage is that it helps structure and clarify a decision. As the characteristics of various options and the criteria with which we are going to assess these options are examined, insights into the problem are developed that otherwise wouldn't. Bringing an ordered, well-reasoned logic to the selection of a best option is useful in most organizational settings, but it is particularly useful in defence planning for the following reason. Consider, for instance the force structure problem within DND. Decisions on which main military assets to purchase are the result of two long, parallel processes. One is formal, the other is informal. The formal process is largely designed to achieve consensus among the various stakeholders who, for the most part, have conflicting interests. A formal options analysis, usually with an explicit MCDA justifying the best option, is a very useful tool to help achieve consensus.

2. *Model Measurement Stage*: Once the options and criteria have been defined, the measurement required by the MCDA technique we have selected is initiated. There are now a strikingly large number of MCDA methods. The major techniques would include Multi-Attribute Utility Theory (MAUT, Keeney and Raiffa (1976)), Simple MultiAttribute Rating Technique (SMART, Edwards (1977)), the Analytic Hierarchy Process (AHP, Saaty(1977, 1980, 1986, 1987), ELECTRE (a family of methods), PROMETHEE, MACBETH and the Method

of Even Swaps (Hammond et al (1998, 1999)), Preference Function Modelling (PFM, see Barzilai 1997, 1998, 2008) and Pugh's Method (1990). Summaries of the most recent research on these and other methods can be found in the compendium by Figueira et al (2005), Belton and Stewart (2002), and the paper by Wallenius et al (2008). Almost all of these methods require the following steps:

(a) *Relative Importance of the Criteria*: All of these methods require the DM to make an assessment of the relative importance of the criteria. Typically, this involves a process that arrives at a set of weights. Criteria with higher weights are judged to be more important than those with lower weights.

(b) *Relative Importance of the Options*: It is usually the case that the relative value of all options are assessed on each criterion, a single criterion at a time.

(c) *Aggregation*: The measurements in the previous two steps are aggregated to arrive at a set of relative values for the options. Obviously, these values will depend on the MCDA method used.

3. *Communication of Results*: The results of a MCDA model are not of much use if they cannot be communicated clearly. I will have more to say about this in a subsequent section.

4.1 The Model Formulation Stage

In my experience there are many kinds of defence problems where a MCDA technique can be employed. The two that readily come to mind are the Course of Action (COA) analysis in the Operational Planning Process and the analysis of a capability deficiency within the capital acquisition process. Since the principles in all of these problems are pretty much the same, I am going to focus on the generation of options to satisfy a capability deficiency because it is arguably the most important defence problem.

Any discussion of a capability deficiency necessarily requires an analysis of the present and future security environment. A good analysis for the CF can be found in Roy and Chapman (2004) who suggest that the identification of deficiencies is sometimes dependant on the pace of technological change. For instance, suppose a defence contractor is suddenly able to produce an anti-armour missile with a range two kilometers longer than anything else on the market. Clearly this technological advance has created a deficiency. On the other hand, sometimes defence planners will look at different combinations of existing assets and discover a deficiency. For instance, suppose a country's land force currently has tracked main battle tanks and its planners are wondering whether there is a need for a lighter, wheeled ACV with much better mobility that could respond more quickly to a wider range of tasks. Fundamental analysis, such as war-gaming, might indicate that such a vehicle seriously enhances the force's fighting ability over a wider range of scenarios.

4.2 Generating Criteria

4.2.1 Value Focussed Thinking

Over the past decade, military operations researchers working in the MCDA area have become quite enamored of the ideas in Keeney's (1992) book, *Value Focused Thinking* (VFT). Keeney suggests that the DM must first explore and define values (criteria) rather than define options. He goes on to argue that this effort on value analysis will lead to more creative options. I'm not sure that I agree with him. Based on the MCDA analyses that I have seen within DND, the discussions at the model building stage don't usually follow the "discuss-options-first" approach – the approach that Keeney critiques. Typically they go back and forth between options and criteria, with DMs using differences in the options to develop criteria. In fact, one of the techniques that Keeney describes for motivating the discussion of values is a discussion of options (VFT, page 58):

> "Often one begins to think hard about a decision situation only after at least two alternatives have become apparent. The question then is which of the alternatives is best. This raises the issue of what makes one alternative better than another. An articulation of the features that distinguish existing alternatives provides a basis for identifying some objectives for the decision problem."

As this quote indicates, it seems to me that any reasonable discussion of values must necessarily consider possible options.

This concern aside, there is much that Keeney has offered in VFT that makes sense. For instance, he suggests that a group approach to value analysis should begin with individuals working independently to make a list of criteria. Once these lists have been made, the group can assemble and discuss each list one after the other. As Keeney suggests (VFT, page 57): "If general discussion began immediately, it would be too easy for group members to anchor on the ideas presented by the first speakers."

4.2.2 Desirable Properties for a Set of Criteria

A good set of criteria has three desirable properties:

1. completeness;
2. nonredundancy; and

3. small size.

To be complete, the criteria must specify everything that is important to the DM. If there is some important criterion left out, then clearly the DM may choose an inferior option. The values should also be nonredundant. If two criteria measure the same thing, then the DM will put too much emphasis on the actual criterion he or she wants to measure. I have seen an example of this within DND. In a contractor selection using an MCDA, two of the criteria used were Cost (a contractor's total cost measured by the hourly rate times the number of hours to complete the work) and Hourly Rate (a contractor's rate per hour). These two criteria are clearly dependent and led to the DM putting more emphasis on Cost then he intended. Elimination of the redundant criterion caused another contractor to be selected. Finally the set of criteria should be as small as possible. The smaller the tree, the fewer the assessments the DM will have to make.

4.2.3 Buede and Bresnick

For military operations researchers who specialize in the application of MCDA, the paper by Buede and Bresnick (BB, 1992) is particularly useful. The authors' discuss the generation of four value trees for four different stages of the weapon system acquisition process in the United States. These examples are good because they develop fairly comprehensive lists of appropriate criteria for such problems. By way of illustration, I discuss one of their examples.

In 1980, BB were asked to define the requirements for a new program called the mobile protected weapons system (MPWS). This concept called for an armoured vehicle with the following characteristics:

1. it had to be helicopter transportable; and
2. it required a direct-fire gun or missile system to support marines landing on a beach with subsequent inland movement towards an objective.

In a first-pass at the problem, BB suggested the following top-level criteria to DMs:

1. Firepower
2. Mobility
3. Survivability
4. Reliability/Availability/Maintainability (RAM)
5. Helicopter Transportability

In their subsequent discussion with DMs in order to define relative values for these criteria, it became obvious that these relative values depended on the role the MPWS would play. For instance, when asked about the trade-off between firepower and helicopter transportability, the DMs said that it depended on what they were doing. Consequently, BB defined three scenarios where the asset would be used: assault support (during an amphibious landing), blocking position (establishing defensive blocking positions if required), and subsequent operations (inland operations subsequent to a successful landing).

The complete value hierarchy is shown in Figure 4. Note that criteria are quite detailed. For instance, the criterion Firepower has five sub-criteria:

1. Lethality
2. Accuracy
3. Servicing Rate
4. Target Acquisition
5. Stowed Kills

These in turn have further sub-criteria. For instance, Lethality is measured against Tanks, Light Armoured Vehicles, Materiel, Personnel, and Helicopters. All in, this is a very detailed criteria set. DND planners could certainly use it as a template, adding or subtracting criteria as they see fit depending on the application.

4.3 Multi-Criteria Systems

Based on my reading of the MCDA literature, there are three systems that aren't mentioned much: Jonathan Barzilai's Preference Function Modelling (PFM); Pugh's Method; and the Method of Even Swaps. In the next section I discuss PFM. In a subsequent section, I present a modification of Pugh's Method that I think would be useful in a defence setting. A final section looks at the Method of Even Swaps.

4.3.1 Preference Function Modelling

To be frank, I am not sure what is "under the PFM hood." Barzilai has written a series of papers describing principles of measurement in the social sciences. Some of these are available from his website

http://www.scientificmetrics.com/publications.html.

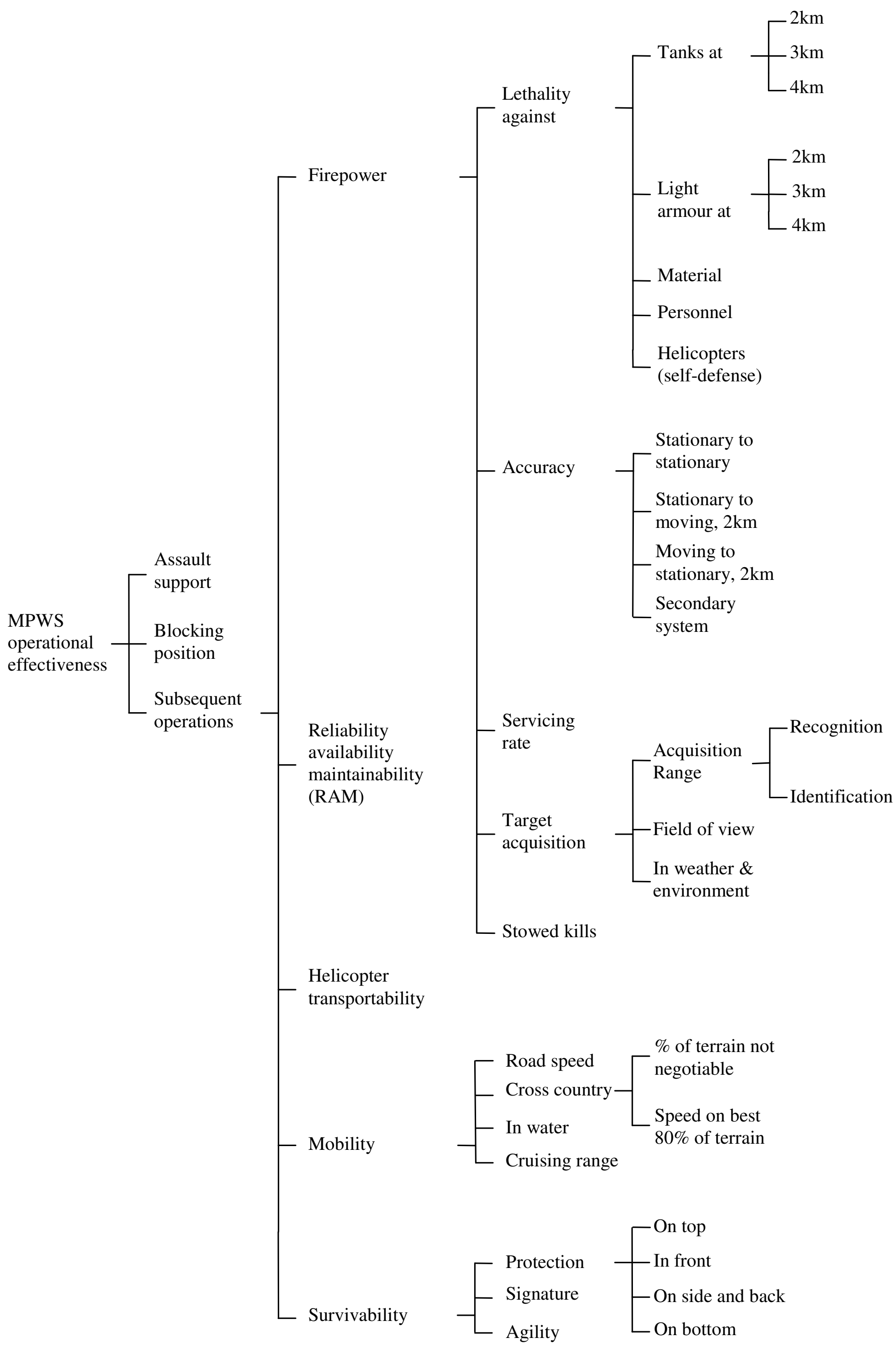

Figure 4: The BB value hierarchy for the MPWS system.

But, to my knowledge, he has not revealed exactly how his system works in any of these.

I have used his software implementation of PFM, *Tetra*, and I find it to be easy to learn how to use. Moreover, I suspect that with a little ingenuity, one could use the software to determine exactly what calculations are being made. To this point, I have not had the time and even if I discovered its inner workings, it would not be something I would feel comfortable disseminating.

These concerns aside, I know Jonathan Barzilai and, more importantly I am familiar with most of what he is published. I consider his work to be outstanding. I have been urging him to write a book on the topic of measurement in the social sciences but he continues to be preoccupied with pushing the theory and disseminating his results.

4.3.2 A Modification of Pugh's Method

The interested reader is referred to Pugh (1990) for a presentation of his method. What I have done here is modify it so that it incorporates a number of aspects of other systems I like. In addition to the basic system, I develop a simple way to do a sensitivity analysis. In my view, this method could easily be used in a defence options analysis.

Suppose that a DM has narrowed the discussion to four options including a reference option. Label these *RefOption*, *Option1*, *Option2*, and *Option3*. The *RefOption* is some well-defined base option, possibly the existing option. In addition, suppose there are five criteria labelled $V_1, V_2, ..., V_5$. We begin by setting these objects up in an empty data table:

	Options			
Criteria	*RefOption*	*Option1*	*Option2*	*Option3*
V_1				
V_2				
V_3				
V_4				
V_5				

Next the DM would use the following verbal scale with its associated numeric scale to assess each of the options on each of the criteria:

Verbal Scale	**Numeric Scale**
Option is *much worse* than the RefOption	-2
Option is *worse* than the RefOption	-1
Option is the *same* as the RefOption	0
Option is *better* than the RefOption	$+1$
Option is *much better* than the RefOption	$+2$

Suppose the DM filled in the following values:

Criteria	Options			
	RefOption	*Option1*	*Option2*	*Option3*
V_1	0	+2	0	+1
V_2	0	+1	−1	+1
V_3	0	0	+2	−1
V_4	0	−2	+1	−1
V_5	0	0	−1	+2

We could interpret this table in the following way. For instance, *Option1*'s measure of +2 on V_1 means that *Option1* is much better than the *RefOption* on criterion V_1. *Option3*'s measure of −1 on V_3 means that *Option3* is worse than the *RefOption* on criterion V_1.

To get the overall score for an option, we would simply add the numbers in each column. This is shown in the following table:

Criteria	Options			
	RefOption	*Option1*	*Option2*	*Option3*
V_1	0	+2	0	+1
V_2	0	+1	−1	+1
V_3	0	0	+2	−1
V_4	0	−2	+1	−1
V_5	0	0	−1	+2
Overall Score	0	+1	+1	+2

In this case *Option3* would be judged to be the best because it has the best overall score. Note that all three options are better than the reference option since their overall scores exceed 0.

But the difficulty with this approach is that it ignores the relative value of the criteria. It might be that some criteria are more important than others. For instance, for just about any weapon system, we ought to put a higher weight on protection than on the cost of the system. To implement heterogeneous relative values for the criteria, suppose we associate a *weight* with each criterion. The following table gives a hypothetical set of weights for the five criteria:

		Options			
Criteria	**Weights**	*RefOption*	*Option1*	*Option2*	*Option3*
V_1	0.30	0	$+2$	0	$+1$
V_2	0.25	0	$+1$	-1	$+1$
V_3	0.20	0	0	$+2$	-1
V_4	0.15	0	-2	$+1$	-1
V_5	0.10	0	0	-1	$+2$
Overall Score		0.00	0.55	0.20	0.40

To get the overall score for *Option1*, we would calculate

$$S_1 = 0.30(+2) + 0.25(+1) + 0.20(0) + 0.15(-2) + 0.10(0) = 0.55.$$

The scores for the other options are computed in the same way. Note that *Option1* is now the best. Note as well that all three have positive overall scores meaning that they are preferred to the reference option.

More generally, let the score of option j on criterion i be s_{ij} and let the weight of criterion i be w_i. Then the overall score of option j is

$$S_j = \sum_i w_i s_{ij} \tag{15}$$

where

$$0 \leq w_i \leq 1 \text{ for all } i \text{ and } \sum_i w_i = 1. \tag{16}$$

4.3.2.1 Estimating the Criteria Weights

To estimate criteria weights for any level of a value hierarchy, I like the following approach. First have the DM put the criteria in order from the most important to the least important. If the DM feels that two criteria are equally important, that's fine. Once the criteria are in order, one of two approaches can be used to fix an initial set of weights.

Approach 1: Use one of the following two formulas to fix an initial set of weights (see Stillwell et al (1981)):

$$\text{Formula 1:} \qquad w_j = \frac{1/r_j}{\sum_k 1/r_k} \tag{17}$$

$$\text{Formula 2:} \qquad w_j = \frac{n - r_j + 1}{\sum_k (n - r_k + 1)} \tag{18}$$

where r_j is the rank of the j-th criterion and n is the number of criteria. For instance, if there are 5 criteria, the weights would be as shown in the following table:

Weight	Formula 1	Formula 2
w_1	0.4380	0.3333
w_2	0.2190	0.2667
w_3	0.1460	0.2000
w_4	0.1095	0.1333
w_5	0.0876	0.0667

Once one of these initial sets of weights is selected, the DMO could then be asked to adjust the weights as required.

Approach 2: This approach uses ratio scale measurement supplemented with a verbal scale that allows the facilitator to assess the consistency of the set of weights that the DM derives. Again, the first step is to rank-order the criteria from most important to least important. By way of a concrete example, suppose option configurations of an ACV are being assessed using the criteria of Firepower, Mobility, and Survivability. Upon being asked to rank-order the criteria from most important to least important, suppose the DM specifies

$$\begin{array}{l}\text{Firepower}\\ \text{Survivability}\\ \text{Mobility}\end{array} \tag{19}$$

The next step is to begin to extract weights. We begin with the bottom two criteria and ask the DM to answer the following question: "Suppose we attach the number 100 to the absolute importance of Mobility. You have assessed that Survivability is more important. What I am interested in is how much more important. Would you agree that Survivability is 20% more important than Mobility?" It is unlikely that the DM will agree. When he does not, you would ask him for a percentage that he feels comfortable with. Suppose he responds 30%. What you are trying to get is the ratio of the absolute weights for Survivability and Mobility. Let these by w_S and w_M respectively. Then by responding 30%, the DM is assessing that

$$a_{SM} = \frac{w_S}{w_M} = 1.3. \tag{20}$$

Next we get the DM to assess Firepower and Survivability by asking the same type of question: "Suppose now we attach the number 100 to the absolute importance of Survivability. You have assessed that Firepower is more important than Survivability. What I am interested in is how much more important. Would you agree that Firepower is 20% more important than Survivability?" Again, it is unlikely that the

DM will agree and when he does not, you would ask him for a percentage that he feels comfortable with. Suppose he responds 50%. This gives

$$a_{FS} = \frac{w_F}{w_S} = 1.5. \tag{21}$$

We are now in a position to derive the weights. We have that

$$\frac{w_F}{w_S} = a_{FS} = 1.5 \tag{22}$$
$$\frac{w_S}{w_M} = a_{SM} = 1.3. \tag{23}$$

But these two equations are not enough to solve for the three weights. The additional equation we need is

$$w_F + w_M + w_S = 1. \tag{24}$$

To solve this system of three equations in three unknowns, note that

$$w_S = \frac{1}{a_{FS}} w_F \tag{25}$$
$$w_M = \frac{1}{a_{FS} \cdot a_{SM}} w_F. \tag{26}$$

Substituting these into (24) gives

$$w_F = \frac{1}{1 + \frac{1}{a_{FS}} + \frac{1}{a_{FS} \cdot a_{SM}}} = 0.4588. \tag{27}$$

Substituting this result back into (25) and (26) gives

$$w_M = \frac{1}{a_{FM}} \cdot \frac{1}{1 + \frac{1}{a_{FS}} + \frac{1}{a_{FS} \cdot a_{SM}}} = 0.2353. \tag{28}$$
$$w_S = \frac{1}{a_{FS} \cdot a_{SM}} \cdot \frac{1}{1 + \frac{1}{a_{FS}} + \frac{1}{a_{FS} \cdot a_{SM}}} = 0.3059 \tag{29}$$

and the system is solved. The weights are

$$w_F = 0.4588$$
$$w_S = 0.3059$$
$$w_M = 0.2353.$$

The questions and calculations for getting the weights associated with n criteria are straightforward. Suppose the criteria weights are denoted $w_1, w_2, ..., w_n$ and we ask the DM to assess

$$\frac{w_1}{w_2}, \frac{w_2}{w_3}, ..., \frac{w_{n-1}}{w_n}. \tag{30}$$

Suppose the DM responds

$$\begin{aligned} \frac{w_1}{w_2} &= a_{12} \\ \frac{w_2}{w_3} &= a_{23} \\ &\ldots \\ \frac{w_{n-1}}{w_n} &= a_{n-1,n} \end{aligned} \tag{31}$$

Define

$$\begin{aligned} b_2 &= \frac{1}{a_{12}} \\ b_3 &= \frac{1}{a_{12}a_{23}} \\ &\ldots \\ b_n &= \frac{1}{a_{12}a_{23}\ldots a_{n-1,n}} \end{aligned} \tag{32}$$

Then

$$\begin{aligned} w_1 &= \frac{1}{1+b_2+b_3+\ldots+b_n} \\ w_2 &= \frac{b_2}{1+b_2+b_3+\ldots+b_n} \\ w_3 &= \frac{b_3}{1+b_2+b_3+\ldots+b_n} \\ &\ldots \\ w_n &= \frac{b_n}{1+b_2+b_3+\ldots+b_n} \end{aligned} \tag{33}$$

This technique enables us to assess the DM's consistency. To see how this can be done, reconsider the example where we're assessing the military effectiveness of an ACV using the criteria Firepower, Mobility, and Survivability. Above we describe a situation were we asked the DM to assess the ratio of the importance of Firepower to Mobility, w_F/w_M, and Mobility to Survivability, w_M/w_S, and the DM responds with a_{FM} and a_{MS} respectively. Suppose now we ask him for a third assessment, w_F/w_S and he responds a_{FS}. How, then, should a_{FS} be related to a_{FM} and a_{MS}?

First note that

$$a_{FM} \cdot a_{MS} = \frac{w_F}{w_M} \cdot \frac{w_M}{w_S} = \frac{w_F}{w_S} = a_{FS} \tag{34}$$

or in other words, if the DM is consistent, it should be that

$$a_{FS} = a_{FM} \cdot a_{MS}. \tag{35}$$

In the case where $a_{FS} = a_{FM} \cdot a_{MS}$ we say the DM is *perfectly consistent*. For the problems we are considering, it is unlikely that a DM will ever be perfectly consistent. Hence it is a matter of determining the level of inconsistency that is acceptable. In the case where a DM is not consistent enough, he must reconsider his assessments.

In the general case, there are any number of ways we can check the DM's consistency. Suppose the DM has responded $a_{12}, a_{23}, ..., a_{n-1,n}$. Then we can ask him to specify $a_{i,i+2}$ realizing that, if he is to be consistent, it must be that

$$a_{i,i+2} = a_{i,i+1}a_{i+1,i+2}. \tag{36}$$

Given that DM has been asked for the three assessments, a_{FM}, a_{MS}, a_{FS}, and these assessments are not perfectly consistent but consistent enough, how do we calculate the weights? We could use the method described above using only a_{FM} and a_{MS}. However that would ignore the information the DM has given us in his response a_{FS}. One way of averaging these three assessments is to do the following optimization:

$$\begin{array}{ll} \min & (w_F/w_M - a_{FM})^2 + (w_M/w_S - a_{MS})^2 + (w_F/w_S - a_{FS})^2 \\ \text{s.t.} & w_F + w_M + w_S = 1 \end{array} \tag{37}$$

This program minimizes the squared deviation of the assessments from the weights they purport to measure.

4.3.2.2 Sensitivity With Imploding and Exploding Criteria Weights

In the case where the DM is certain about the rank-order of criteria but not their weights, the following sensitivity analysis can be done. Suppose we reconsider the opening example of this section where the DM's initial assessment of the weights is reproduced below:

Criteria	**Weights**
V_1	0.30
V_2	0.25
V_3	0.20
V_4	0.15
V_5	0.10

Suppose we denote the weight of criterion i with w_i. Let the new weight for criterion i be given by

$$w_i(\alpha) = \frac{w_i^\alpha}{\sum_i w_i^\alpha} \tag{38}$$

where the exponent α will control the way the weights will change. Initially we set $\alpha = 1$ so that $w_i(\alpha) = w_i$ for all criteria. The following table shows values of the weights for selected values of α:

Criteria	$\alpha = 0.5$	$\alpha = 0.8$	$\alpha = 1.0$	$\alpha = 1.2$	$\alpha = 1.5$
V_1	0.2491	0.2796	0.30	0.3204	0.3507
V_2	0.2274	0.2416	0.25	0.2574	0.2668
V_3	0.2034	0.2021	0.20	0.1970	0.1909
V_4	0.1762	0.1606	0.15	0.1395	0.1240
V_5	0.1438	0.1161	0.10	0.0857	0.0675

Note that the transformation preserves the rank-order of the criteria. Also note that, for values of α which exceed 1, the weights "explode" meaning that the weights that exceed 1/5 (in the general case $1/n$ where n is the number of criteria) get larger and those that are 1/5 or less get smaller. On the other hand, for values of α less than 1, the weights implode. That is, all weights get closer to 1/5 (in the general case, $1/n$). In fact, for n criteria, we have that

$$\lim_{\alpha \to 0} w_i(\alpha) = \frac{1}{n}. \tag{39}$$

The use of this idea in sensitivity analysis is straightforward. Once the initial weights have been determined and a best option has been tentatively selected, the analysis could then see how the best option changes as the explosion parameter α is changed. For the initial example outlining Pugh's Method, the results are shown in the following table:

	Overall Scores			
Value of α	*Option1*	*Option2*	*Option3*	**Best Option**
0.5	0.3734	0.2117	0.3847	3
0.8	0.4796	0.2071	0.3907	1
1.0	0.55	0.20	0.40	1
1.2	0.6193	0.1715	0.4129	1
1.5	0.7203	0.1715	0.4376	1

Note that *Option1* is best in the base case ($\alpha = 1$). Moreover that is so for just about all other values of α. However, once α is sufficiently lower than 1, *Option3* becomes the best option. That is, as the weights get closer together (the DM sees not much difference in the relative values of the weights), *Option3* is best.

4.3.3 The Method of Even Swaps

The Method of Even Swaps (Hammond, Keeney, and Raiffa (1998,1999)) arrives at a preferred option through a series of simple comparisons and judgements. It has the characteristic that most intelligent DMs can easily understand it.

I motivate the method with an example involving an internal DND analysis of the selection of an Unmanned Aerial Surveillance and Target Acquisition System (UASTAS). A typical UASTAS system comprises ground stations and unmanned airframes. Airframes are programmed to fly over areas forward of the battlefield. They have extremely sophisticated on-board sensor systems which transmit almost real-time information (including video) from their field of view back to a ground station. The problem will be to choose the best system given assessments of each system on a number of important criteria.

Suppose there are four off-the-shelf systems under consideration. These are labelled A, B, C and D. After consultation with fellow staffers, you arrive at the following set of assessment criteria:

RANGE: the effective range of an airframe measured in kilometers.

FIELD OF VIEW: an aggregate measure of the range of sensors measured in square kilometers.

RESPONSE TIME: the time in minutes to mission plan and launch a back-up airframe in the event in-flight reprogramming is not possible.

SURVIVABILITY: the probability an airframe completes a mission in a defined hostile environment. It was estimated with error by Land Operational Research Staff.

COST: the sum of the Procurement Cost and the discounted future costs of Personnel, Operation, and Maintenance (PO&M), measured in present-year dollars.

The system works in the following way. First we summarize the raw data for each system on each criterion in what is termed a Consequences Table:

Consequences Table				
	System A	**System B**	**System C**	**System D**
RANGE	56	60	60	55
FIELD OF VIEW	1.00	1.18	1.15	1.15
RESPONSE TIME	62	60	50	70
SURVIVABILITY	0.85	0.90	0.90	0.92
COST	52	49	42	55

With this Consequences Table in hand, the first thing to do is rank-order the systems for each assessment criterion:

Consequences Table (Rank-Order)				
	System A	**System B**	**System C**	**System D**
RANGE	3	1	1	4
FIELD OF VIEW	4	1	2	2
RESPONSE TIME	3	2	1	4
SURVIVABILITY	4	2	2	1
COST	3	2	1	4
Rank-Sum	17	8	7	15

The bottom row of this table is the sum of the column rank-orders. Clearly we would like to look at systems with low rank-order sums. In this case Systems B and C have low rank-sums; Systems A and D have high rank-sums.

We first look for *dominated* options. Consider Systems A and B. Note that B has a lower rank-order for every assessment criterion. Hence A is dominated by B and we can exclude System A from further consideration.

Now we examine System B and System D. System B is at least as good as System D for every assessment criterion except **SURVIVABILITY**. Given that **SURVIVABILITY** is measured with error and the two measures of **SURVIVABILITY** are close anyway (.90 for System B and .92 for System D), the decision is made that System B is preferred to System D. We can now eliminate System D.

The reduced Consequences Table is as follows:

Consequences Table		
	System B	**System C**
RANGE	60	60
FIELD OF VIEW	1.18	1.15
RESPONSE TIME	60	50
SURVIVABILITY	0.90	0.90
COST	49	42

Note that these two systems have identical measures on **RANGE** and **SURVIVABILITY**. We can now eliminate these criteria to further reduce the Consequences Table to:

Consequences Table		
	System B	**System C**
FIELD OF VIEW	1.18	1.15
RESPONSE TIME	60	50
COST	49	42

Now we are in a position to do an *even swap*. Consider System C and ask what we are prepared to pay to get the **FIELD OF VIEW** up to 1.18 square kilometers. Suppose we are prepared to pay an additional 2 million dollars. Then the revised Consequences Table is

Consequences Table		
	System B	**System C**
FIELD OF VIEW	1.18	1.15+.03
RESPONSE TIME	60	50
COST	49	42+2

Note that to get the extra field of view (1.15 to 1.15 + .03) we must be prepared to pay an additional dollar amount. Note as well that we could have chosen to adjust the **RESPONSE TIME** rather than the **FIELD OF VIEW**. With this even swap we can now eliminate the **FIELD OF VIEW** criterion giving:

Consequences Table		
	System B	**System C**
RESPONSE TIME	60	50
COST	49	44

Now the decision is clear. System C has a lower **RESPONSE TIME** and a lower **COST**. Hence System C is the preferred system.

It is worth pointing out that the Method of Even Swaps has its origin in the advice Benjamin Franklin offered Joseph Priestly:

> "...To get over this, my way is to divide half a sheet of paper by a line into two columns; writing over the one Pro, and over the other Con. Then, during three or four days consideration, I put down under the different heads short hints of the different motives, that at different times occur to me, for or against the measure.
>
> When I have thus got them all together in one view, I endeavor to estimate their respective weights; and where I find two one on each side, that seem equal, I strike them both out. If I find a reason pro equal to some two reasons con, I strike out the three. If I judge some two reasons con, equal to three reasons pro, I strike out the five; and thus proceeding I find at length where the balance lies; and if, after a day or two of further consideration, nothing new that is of importance occurs on either side, I come to a determination accordingly. And, though the weight of the reasons cannot be taken with the precision of algebraic quantities, yet when each is thus considered, separately and comparatively, and the whole lies before me, I think I can judge better, and am less liable to make

> a rash step, and in fact I have found great advantage from this kind of equation, and what might be called moral or prudential algebra ...”

Hammond, Keeney and Raiffa have extended Franklin’s logic in a very creative way to the case where there are more than two options. It is worth pointing out that some of the methods discussed in Edward de Bono’s books on creative thinking bear a striking resemblance to Franklin’s “moral algebra”.

4.4 The Problem of Rank Reversal

There are a large number of MCDA systems available for a DM to use. Some are subject to the methodological problem of rank reversal. Consider, for instance, the Analytic Hierarchy Process (AHP), one of the more popular MCDA techniques. Belton and Gear (1984, pp. 228–230) proposed an instance of an AHP multicriteria decision with 3 options, labelled A, B, and C. Using the AHP evaluation procedure, they found that

$$A \succ B \succ C, \tag{40}$$

that is, option A was preferred to option B and option B was preferred to C. Next they introduced a fourth option, D, which was an exact copy of C. Most would agree that a reasonable method should return

$$A \succ B \succ C \sim D, \tag{41}$$

where $C \sim D$ indicates that options C and D are equivalent. Unfortunately the AHP returns

$$B \succ A \succ C \sim D. \tag{42}$$

Options C and D are equivalent, but note that option B is now the preferred alternative. This phenomenon is referred to as *rank reversal* and most researchers consider it to be a serious methodological flaw.

However the AHP is not the only system subject to rank reversal. Consider the following example. Some time ago, a DLR desk officer working on the Clothe the Soldier project came to me with the following Bid Evaluation Scheme for the Wet Weather Boot competition. He proposed that competing bids be scored first on three human factor criteria: Comfort, Support, and Donning and Doffing. Each boot was scored on a 9-point scale where 1 was poor and 9 very good. To get an overall human factors performance, the scores for each boot would then be aggregated taking into account the relative importance of the human factors criteria. To see precisely how the system would have worked, consider the following example:

	Points	**Boot A**	**Boot B**
Comfort	9	7	4
Support	5	5	5
Donning and Doffing	1	4	7
HF Score		92	68

To get the HF Score for Boot A, the computation is

$$9(7) + 5(5) + 1(4) = 92.$$

To incorporate cost, the officer proposed the following normalization. Suppose the HF Scores and Cost for three boots are as follows:

	Boot A	**Boot B**	**Boot C**
HF Score	92	68	78
Cost	100	90	60

To get the normalized elements in the HF row, each of the elements will be divided by the lowest element of the row. To get normalized elements for the Cost row, the lowest element in the row is multiplied by the inverse of an element's cost. This gives the following normalized scores:

	Boot A	**Boot B**	**Boot C**
HF Score	1.35	1.00	1.15
Cost	0.60	0.67	1.00

Finally these scores are weighted by the respective weights of the HF Score and Cost:

	Weights	**Boot A**	**Boot B**	**Boot C**
HF Score	.40	1.35	1.00	1.15
Cost	.60	0.60	0.67	1.00
		0.90	0.80	1.06

By this simple calculation, the Bid Evaluation Scheme would determine that Boot C is best. After explaining his procedure he wanted to know what I thought of it.

Jonathan Barzilai has shown that any MCDA system that includes a normalization of raw scores on options is subject to rank reversal. Consequently, I explained to the officer that his system was subject to rank reversal. I explained to him that I could find characteristics of a fictitious, inferior boot that would cause options A and C to reverse their position. Here is an example. Suppose I add Boot D with the charactersistics as shown in the last column of the following table:

	Boot A	**Boot B**	**Boot C**	**Boot D**
HF Score	92	68	78	10
Cost	100	90	60	100

The calculation to get each boot's final score is shown in the following table:

	Weights	Boot A	Boot B	Boot C	Boot D
HF Score	.40	9.2	6.8	7.8	1.0
Cost	.60	.60	.67	1.0	.60
		4.0	3.1	3.7	0.8

Note first that Boot D is inferior. More importantly, note that Boot A and Boot C have reversed their ranks. Now Boot A is best.

Using such flawed methodologies has the potential to put contractors in a position where they could claim that the Bid Evaluation Scheme is unfair. For instance consider the above example and suppose that the contractor submitting Boot C claimed that if Boot D were ignored, the same scheme would rank his boot best. This has the potential to end up in court and lead to significant delays in the acquisition of a boot. Ultimately, there are any number of schemes that are not subject to rank reversal and are as reasonable as this one for suggesting to contractors the relative importance of the criteria. Hence Bid Evaluation Schemes that are subject to rank reversal should be not be used.

4.5 An MCDA Best Partition

Suppose that a MCDA analysis produces the following overall scores for five options:

$$\begin{aligned} w_1 &= 0.31 \\ w_2 &= 0.29 \\ w_3 &= 0.15 \\ w_4 &= 0.13 \\ w_5 &= 0.12 \end{aligned}$$

Option 1 is best but Option 2 is not too far off. Clearly, both Options 1 and 2 are significantly better than Options 3, 4, and 5. This naturally leads to a partition of these outcomes. The best partition is Options 1 and 2. Given the uncertainty in the estimates of MCDA input, it is probably the case that Options 1 and 2 are indistinguishable. Moreover, let us suppose that the weights for Options 1 and 2 are drawings from probability distributions and the DM is uncertain which has the highest mean. Then it doesn't matter which one is picked because picking the wrong one does not result in a significant loss in value. In my opinion, it comes down to judgement (not measurement) when one has to select the best option from the best partition.

4.6 An Important Implicit Assumption

The whole point of MCDA is to break a complex options analysis into its component parts. The idea is that, if we do a good job measuring the component parts, then we necessarily do a good job making the complex assessment. But, to my knowledge, there has been no systematic examination of this assumption.

By way of example, suppose we can measure C directly with a relative error of up to ϵ each way. That is, the measured value of C lies in the interval $[c(1-\epsilon), c(1+\epsilon)]$ where c is the true value. Alternatively, suppose C can also be measured with the product AB. That is, we could measure C by first measuring A and B and then taking their product. This would be a good thing to do if it yields a lower relative error than the case where C is measured directly. Suppose we can measure both A and B with a relative error of δ. Again this means that the measured value of A lies in the interval $[a(1-\delta), a(1+\delta)]$ where a is the true value of A. Similarly the interval for B is $[b(1-\delta), b(1+\delta)]$. Then the interval for the product is

$$[ab(1-\delta)^2, ab(1+\delta)^2] \tag{43}$$

and this will be narrower than $[c(1-\epsilon), c(1+\epsilon)]$ as long as

$$ab(1+\delta)^2 - ab(1-\delta)^2 < c(1+\epsilon) - c(1-\epsilon). \tag{44}$$

Note that the left-hand side of this inequality is the length of the interval $[ab(1-\delta)^2, ab(1+\delta)^2]$ and the right-hand side is the length of $[c(1-\epsilon), c(1+\epsilon)]$. Taking into account that $c = ab$ by assumption and simplifying the inequality gives the condition

$$\delta < \epsilon/2. \tag{45}$$

So we have to be able to measure both A and B with approximately double the accuracy of C in order for the decomposition approach to be preferable.

Now suppose that C can be measured by A/B and that the same intervals of uncertainty apply for a, b, and c. Then the length of the interval of uncertainty for measuring C with A/B is

$$\frac{a(1+\delta)}{b(1-\delta)} - \frac{a(1-\delta)}{b(1+\delta)}. \tag{46}$$

The first term in this expression is the highest that A/B could be and the second term is the lowest it could be. This expression simplifies to

$$\frac{a}{b}\frac{4\delta}{1-\delta^2} \tag{47}$$

This will be a smaller interval than the direct measurement of C as long as

$$\frac{a}{b}\frac{4\delta}{1-\delta^2} < c(1+\epsilon) - c(1-\epsilon) \tag{48}$$

or, simplifying, as long as

$$\epsilon > \frac{2\delta}{1-\delta^2}. \tag{49}$$

Note that if δ is small, $1-\delta^2 \simeq 1$, and we have again that

$$\delta < \epsilon/2. \tag{50}$$

So again we have to be able to measure both A and B with approximately double the accuracy of C in order for the decomposition approach to be preferable.

Obviously, MCDA is a much more complex calculation than a simple product or ratio. Nonetheless, these simple examples suggest that you have to be able to measure decomposed parameters with greater accuracy in order for the decomposition to be superior to a more direct approach. And again, this is the implicit measurement assumption we make when we do MCDA. We assume that our decomposed measurements are more accurate than a direct, holistic approach.

4.7 The Work of Dijksterhuis and His Colleagues

Recently psychologists have done some interesting work on how emotion and subconscious thought affect the quality of decisions. Of particular note is a series of papers by Ap Dijksterhuis and his colleagues at the University of Amsterdam, one of which has been published in the prestigious journal *Science* (see Dijksterhuis et al (2006)).

Here are the details of one of the experiments they present in the *Science* paper. Experimental subjects were given information about the attributes of four used cars under different experimental conditions and were asked to pick the best car. In the first set of experiments, subjects were partitioned into groups and given information on 4 attributes for the cars under different experimental conditions. In all there were 16 pieces of information (4 cars $\times$ 4 attributes). The attribute information was such that one of the cars was clearly best. Here are some examples of the attribute information:

For the Kaiwa service is excellent

The Dasuka is new

The Hatsdun has a poor sound system

The Kaiwa has no sunroof

The procedure for each group varied as follows:

Procedure 1 (Conscious): Subjects were given the information on an attribute every 8 seconds via a computer screen. After this information was presented, subjects were given four minutes to determine which car was best.

Procedure 2 (Unconscious): This procedure was the same as Procedure 1 except that, during the four minute period after the information was given, subjects were distracted by having to do anagrams. Subsequent to this period of distraction, subjects were asked to choose the best car.

The difference between these two procedures centers on how subjects made the decision. Clearly Procedure 1 allowed subjects to think about which car was best whereas Procedure 2 did not allow for conscious thought. As it turns out, subjects were able to pick the best car 57% of the time with Procedure 1 which is not really surprising. But it gets more interesting when you increase the number of attributes.

When the number of attributes was raised from 4 to 12, subjects did much better picking the best car under Procedure 2. Almost 60% of Procedure 2 subjects picked the correct car whereas less than 25% of Procedure 1 subjects picked the correct car.

At first blush this is an astonishing result. My immediate reaction was to question the value of a MCDA process. Would we be better off doing anagrams before making an important multi-criteria defence decision? Some researchers think so. Jonah Lehrer (2009) has documented Dijksterhuis's work in his new book "How We Decide," and, therein, quotes Dijksterhuis (page 237):

> "Use your conscious mind to acquire all the information you need for making a decision. But don't try to analyze the information with your conscious mind. Instead, go on holiday while your unconscious mind digests it. Whatever your intuition then tells you is almost certainly going to be the best choice."

Later in the chapter, Lehrer summarizes:

> "Complex problems, on the other hand, require the processing powers of the emotional brain, the supercomputer of the mind. This doesn't mean you can just blink and know what to do – even the unconscious takes a little time to process information – but it does suggest that there's a better way to make difficult decisions. When choosing a couch [furniture], or holding a mysterious set of cards [poker], always listen to your feelings. They know more than you do."

I think Dijksterhuis is wrong. Let's begin by looking at the experimental procedure. In the case where subjects were given 12 attributes on 4 cars, in all 48 pieces of information, they were given a piece of information every 8 seconds. Thereafter, they had no record of this information and were given 4 minutes to think about which car was best before making the decision. Our working memories are clearly not enough to keep track of all this information. To me, it's not surprising that not many subjects were able to get the right car. How almost 60% of subjects were able to get the right car after the distraction period may have been a fluke – after all, there were only 20 subjects. Moreover, to my knowledge, no researchers have been able to replicate Dijksterhuis's results. A good summary of these negative results can be found in Thorsteinson and Withrow (2009).

A better test of how conscious thought compares to unconscious thought would be to allow decision-makers to use a MCDA technique. Such techniques allow decision-makers to supplement their working memories with storage media (a simple spreadsheet would suffice). Once the attribute ratings were captured, decision-makers could use Pugh's Method or some other MCDA technique to determine the best option. In my view, there is some interesting experimental work that could be done to test whether MCDA techniques result in better decisions.

5 Hard OR

5.1 Group Organization for Options Analysis

A significant amount of work in military OR is done in work groups and facilitated work groups. The situation where an OR professional must work with a military work group to arrive at a decision requires quite a bit of skill and know-how. While the literature on how to facilitate work groups generally is large, there is not much on the specific problem of facilitating military work groups. All this to say that military OR should not be just be about modelling and measurement (so-called "Hard OR"). It should also focus on the functioning of the people and organization charged with making the decision ("Soft OR"). Quite frankly, these two aspects of a military OR engagement are improperly named. In my view, the names should be reversed. Coming up with an appropriate model for the problem at hand is "Easy OR"; executing the intermediation that extracts options, criteria, and parameter estimates from the DMs, to my mind, is the "Hard OR."

So what is different about military work groups, particularly as they deal with options analysis problems? I am going to conceive of an *options analysis work group (OAWG)* as two parts: the *Decision-Maker Work Group (DMWG)* and *Decision Support Organization (DSO)*:

1. *The Decision-Maker Working Group (DMWG)*: This is the individual or group of individuals who have been asked to make a decision. Normally it will be a committee of officers with a senior officer as Chair. Sometimes the Chair is responsible for the decision. In other circumstances, DMWG is an ad hoc committee assembled to make a specific recommendation to a commander or more senior DM.

2. *The Decision Support Organization (DSO)*: This group serves in a staff role. They are responsible to the Chair DMWG and take direction from him. Normally this group will comprise operations researchers and decision analysis experts. It is their job to advise the Chair DMWG on appropriate modelling approaches and measurement issues.

A typical options analysis would proceed as follows. Very often they begin with a DM with a problem and neither the time nor expertise to solve it. But he/she does have resources to put together an ad hoc working group to look at the problem. At this point, the following steps would be taken:

1. *OAWG Staffing and Problem Identification*: Presumably the Chair, DMWG would be charged with a responsibility where he felt that an Options Analysis was the appropriate choice. Moreover, in discussion with OR personnel, he has

decided that a numerate approach to the problem is in order. He would them begin to staff both OAO organizations. A crucial part of this organization would be putting together a DSO team and sorting through the Problem Identification stage with DSO staff. There should be preliminary discussions at this point about a possible modelling approach.

2. *Deciding the Modelling Approach*: The Chair, DMO and DSO staff would finalize the modelling approach. Given the number of modelling approaches for these kinds of problems (optimization, voting methods, Multi-Criteria Decision Models, etc.), this is not a straightforward issue. Again, the DSO team must provide sound advice to the Chair, DMWG.

3. *Model Development:* The Chair DMWG and DSO staff would decide the way ahead based on the modelling approach selected. At a minimum, this would include:

   (a) A process to solicit the advice of DMWG officers on possible options and criteria. I believe this is the most important part of the process so care must be taken that it is executed properly.

   (b) A process to do the measurement required by the model once options and criteria have been thoroughly discussed and defined.

4. *Decision and Reporting*: The model will suggest a way forward. There should be a formal process in place where the model output is tempered by the judgement of the full OAO. This will result in final advice that the Chair, DMO will report to his superior.

In my experience working with officer-populated DMWG organizations, I think it is fair to say that DMWG members bring to the table an expertise in defence knowledge that is superior to that of the individual or individuals in the DSO organization. Moreover, by nature, they are inclined to taking a numerate approach to this kind of problems. It is crucial that the DSO approach take into account this expertise and work with the DMWG to extract it. An interaction where the DSO simply consults the DMWG when it needs to populate a model with data – the "doctor-patient" model of OR consultation – is not likely to bear fruit. For this reason, I think a study which looks at best OR practise in this kind of environment is long overdue. DND employs a group of OR scientists, some of whom have significant experience working with DMWG organizations on ill-defined, complex problems. I think it behoves CORA to think about an exercise that would put together a best-practise manual for facilitating DMWGs. In my view, the OR techniques are easy to learn; what is difficult is the interaction with the DMWG that builds and populates a worthwhile model.

For the reader interested in learning more about about facilitation and facilitated work groups, I would recommend Belton and Stewart (2003, chapter 9), Schwarz (1994), Phillips (2006), Phillips and Phillips (1993), and Papamichail (2007) et al (2007).

5.2 Groupthink

The basic rationale for groupwork is that many minds are better than one. But some group decisions have been notoriously bad. Janus (1972) and more recently Suroweicki (2004) have each documented spectacularly poor group decisions. Janus explored "groupthink" – the idea that particular groups, ones characterized by significant cohesion and shared values, can make bad decisions. One could argue that military decision groups might be particularly susceptible to groupthink. Afterall, most military groups are hierarchical (there is a leader and others in the group are subordinate), they are cohesive, and they have shared values in the extreme. Nonetheless, the professional ethos of officers, particularly those in staff positions, requires them to offer the commander their best professional opinion, even if it puts them at odds with the group or the commander. There are two reasons for this. First, most officers feel a tremendous responsibility for the well-being of their soldiers and second, a poor decision can sometimes lead to catastrophic outcomes. A good example is the Dieppe Raid in World War II. History makes it clear that the plan for this operation was at best a fantasy and it cost 907 young Canadian soldiers their lives.

There is a rather large psychology literature on the problem of groupthink. As it applies to military organizations, there are three things a leader can do to lessen its effects:

1. *Put together a group with diverse backgrounds and then listen to them.* In general, the more varied the backgrounds and points of view of group participants, the less chance the group will succumb to groupthink. There are two good examples of this strategy. In a recent *New Yorker* article, Steve Coll detailed the membership of an advisory group that General David Petraeus used in Iraq:

   "Petraeus and Crocker summoned to Baghdad a group of outside advisers who became known as the Joint Strategic Assessment Team. The members included more than a dozen military officers, Iraq specialists from the State Department, and outside academics who could think conceptually about the war and the application of counter-insurgency doctrine. Among them was David Kilcullen, an Australian specialist on guerrilla warfare, whose unconventional thinking had made him an influential figure in the State Department's counterterrorism office. Petraeus also invited Stephen Biddle, a military analyst and a Democrat, whose published work about Iraq had previously made the General "very unhappy," in Biddle's words. The invitation to join the advisory group, Biddle concluded, spoke to “a different way of thinking and working.” Once in Iraq, he found that if Petraeus believed the tenets of the counter-insurgency field manual were impractical on a particular point, he simply disregarded them. “This clearly was not a guy who feels obliged to follow some cookbook, even one he co-wrote,” Biddle said.”

A second example is the cabinet that Abraham Lincoln put together when he was elected President of the United States in 1860. This is detailed in Doris Kearns Goodwin’s (2009) new book “Team of Rivals: The Political Genious of Abraham Lincoln.” Lincoln included three of his rivals – William Seward, Salmon Chase, and Edward Bates – for the Republican nomination and he depended on them for advice throughout his presidency. Seward helped write the Gettysburg Address and all three were a part of the political decision-making process that lead to the Emancipation Proclamation and Union victory in the Civil War.

Both of these examples make it clear that it takes a special leader to appoint and deal with advisors who are likely to dissent.

2. *Do not reveal any information about your initial thoughts to group members.* To the extent that subordinates will incorporate the views of the leader, it is best that the leader keep his thoughts to himself until he has heard from all subordinates. This ensures that the leader will get the best possible advice he can from the group.

3. *If numerate techniques are to be employed, use ones that take dissent into account.* Suppose that a group is divided on the weights of the criteria when a MCDA technique is used in an options analysis. Then the DSO ought to evaluate the options using both sets of criteria weights, one for each side. It may be that both lead to the same option being best. In the case where it doesn’t the leader will have to make a decision. There are other ways to include dissenting opinion. In the case where each group member gives a ranking of the options, Cameron (1998) details a number of standard techniques that can be used to partition group members into like-thinking sub-groups based on pairwise rank correlations.

6 Conclusions

In this paper, I have presented an operational research approach to options analysis in a defence context. Three general techniques were presented:

Optimization

On the resource side of the defence management problem, there are some options analysis problems with a large number of options. One example is Miscellaneous Project selection where choosing from among 150 potential projects is not unusual. The traditional approach to these problems is to rank the projects in order of importance and then choose as many of the highest ranked projects as the budget will allow. The difficulty is that the project cost (investment) is not considered. One way to consider project value and cost simultaneously is to model the problem as a knapsack problem where the value of the projects selected is maximized subject to the cost of these projects being within budget.

I am not the first to suggest this approach to project selection in a defence setting. However I make at least a couple of contributions to this literature. First, I show how to generate second, third, and more generally, k^{th} best solutions. The reason for this is that, in some cases, a decision-maker may have extra-model considerations that cause him to prefer one of these k^{th} best solutions to the model's optimal solution. Second, I show how to come up with a value function for the projects assuming they can be put into rank order. Here I posit a couple of simple functional forms relating project value to rank. It is then straightforward to ask decision-makers to estimate the parameters of this function. Finally, I present an example where the knapsack approach is applied to CF Air Force Miscellaneous Project selection. I show that the knapsack approach provides an increase in value of some 20% over the standard approach.

Median Rank Aggregation and the Ordinal Consensus Ranking Problem

In the case where the number of options is relatively low and a group of decision-makers is trying to put these into rank-order, very often each member of the group is asked to submit a rank-order of the options. There are a number of ways these individual rank-orders can be aggregated into a group rank-order. The problem of determining a group rank-order from the individual rank-orders is termed the ordinal consensus ranking problem.

In my experience, median rank aggregation is rarely considered. Consequently, I first outline how it is calculated and then go on to show that it has some nice properties. One of these is that it is relatively robust to strategic voting in the case where the group consists of decision-makers with competing interests. Finally, I present the results of a Monte Carlo experiment which shows that median rank voting is superior to other aggregations schemes (average rank and truncated average rank) in the presence of strategic voting.

Multi-Criteria Decision Analysis

Multi-Criteria Decision Analysis is by far the most frequently used options analysis technique within defence organizations. For example, forms of it are used for force structure analysis, within the Operational Planning Process, and for the assessment of capability deficiencies. In this report, I detailed the use of a variation of Pugh's Method, a method that is popular in engineering design. In addition, I looked at the Method of Even Swaps, an approach that I think is well suited to defence problems. In my experience, I have not seen either method applied in a defence setting.

Besides techniques, I also discussed a number of MCDA pitfalls. One of these is the problem of rank reversal. There are some MCDA approaches, including the Analytic Hierarchy Process, that have the characteristic that the top two options can switch position if an inferior option is added. Most operations researchers would consider this to be a serious methodological problem. In this report I showed why it happens and how to identify techniques that are subject to it.

In addition to techniques, I also discussed the nature of defence organizational processes to bring the tools of operational research to bear on defence options analysis problems. Very often, the decision-makers are military officers with a very good knowledge of operational problems. Operations research scientists, on the other hand, know a lot about operations research tools and modelling approaches. The key to the success of any OR options analysis study is the interaction of these two groups. For the part of the operations research scientist, there has been little research on what works and what doesn't. That said, there are defence operations researchers with a wealth of experience. In my view, the time has come for CORA to look at an organizational "Lessons Learned" capability. Over time this capability might work nicely into some research which leads to a set of professional principles that govern the operations research scientist's approach to a group of military decision-makers.

References

[1] Baker, S., D. Morton, L. Williams, and R. Rosenthal (2002). "Optimizing Strategic Airlift," *Operations Research 50(4)*, 582-602.

[2] Barzilai, Jonathan (1997). "Deriving Weights from Pairwise Comparison Matrices," *Journal of the Operational Research Society*, 48(12), 1226-1232.

[3] Barzilai, Jonathan (1998). "On the Decomposition of Value Functions," *Operations Research Letters*, 22, 159-170.

[4] Barzilai, Jonathan (2008). "Function Modeling (PFM): The Mathematical Foundations of Decision Theory," Technical Report, Department of Industrial Engineering, Dalhousie University.

[5] Bassett, G.W. and Joseph Persky (1994). "Rating Skating," *Journal of the American Statistical Association,* 89, 1075-79.

[6] Beck, M. P. and B. W. Lin, (1983). "Some Heuristics for the Consensus Ranking Problem," *Computers and Operations Research*, 10(1), 1-7.

[7] Belton, V. and T. Gear (1983). "On a Shortcoming of Saaty's Method of Analytic Hierarchies," *Omega*, 11, 228-230.

[8] Belton, Valerie, and T. J. Stewart (2002). *Multiple Criteria Decision Analysis: An Integrated Approach*, Kluwer Academic Publishers, The Netherlands.

[9] Borda, J. C. (1781). "Memoire sur les elections au scrutin," in *Histoire de l'Academie Royale des Sciences*, Annee MDCCLXXXI, Paris, France.

[10] Brown, G., R. Clemence, W. Teufert, and R. Wood (1991). "An Optimization Model for Modernizing the Army's Helicopter Fleet," *Interfaces, 12(4)*, 39-52.

[11] Brown, G., D. Coulter, and A. Washburn (1994). "Sortie Optimization and Munitions Planning," *Military Operations Research, 1(1)*, 13-18.

[12] Brown, G., R. Dell, H. Holtz, and A. Newman (2003). "How US Air Force Space Command Optimizes Long-Term Investment in Space Systems," *Interfaces, 33(4)*, 1-14.

[13] Brown, G. G., R. F. Dell, and A. M. Newman (2004). "Optimization-Based Military Capital Planning," accessed at:

http://econbus.mines.edu/pdf/newman%20updates/cptech040927.pdf.

[14] Buede, D. M. and T. A. Bresnick (1992). "Applications of Decision Analysis to the Military Systems Acquisition Process," *Interfaces*, 22(6), 110-125.

[15] Cameron, Frederick (1998). "Contrary Schools of Thought Within Military Decision-making Groups," Power Point Presentation available at:

$$http://ismor.cds.cranfield.ac.uk/ISMOR/1998/cameron.pdf$$

[16] Coll, Steve (2008). "The General's Dilemma," *The New Yorker*, September 8, accessed at:

$$http://www.newyorker.com/reporting/2008/09/08/080908fa\_fact\_coll$$

[17] Cook, W. D. and M. Kress (1985). "Ordinal Ranking with Intensity of Preference," *Management Science*, 31(1), 26-32.

[18] Cook, W. D. and L. M. Seiford (1982). "On the Borda-Kendall Consensus Method for Priority Ranking Problems," *Management Science*, 28(6), 621-637.

[19] Cook, W. D., Kress, M., and L. M. Sieford (1997). "A General Framework for Distance-Based Consensus in Ordinal Ranking Models," *European Journal of Operations Research*, 96(2), 392-397.

[20] de Condorcet, M. (1785). *Essai sur l'Application de l'Analyse a la Probabilite des Rendues a la Pluralite des Voix*, Paris.

[21] Degroot, M.H. (1975). *Probability and Statistics*, Addison-Wesley, London.

[22] Dijksterhuis, Ap, Maarten W Bos, Loran F Nordigren, and Rick B van Baaren (2006). "On Making the Right Choice: The Deliberation-Without-Attention Effect," *Science*, 311, 1005-1007.

[23] Dyer, J. S. and R. K. Sarin (1979). "Measurable Multiattribute Value Functions," *Operations Research*, 27, 810-822.

[24] Eilon, S. (1987). "Application of Knapsack Model for Budgeting," *Omega*, 15, 489-494.

[25] Emond, E. J. and D. W. Mason (2003). "A New Rank Correlation Coefficient with Application to the Consensus Ranking Problem," *Journal of Multi-Criteria Decision Analysis*, 11(1), 17-28.

[26] Evans, G. and R. Fairbairn (1989). "Selection and Scheduling of Advanced Missions for NASA Using 0-1 Integer Linear Programming," *Journal of the Operational Research Society, 40(11)*, 971-981.

[27] Ewing, P. L., W. Tarantino, and G. S. Parnell (2006). "Use of Decision Analysis in the Army Base Realignment and Closure (BRAC) 2005 Military Value Analysis," *Decision Analysis*, 3, 33-49.

[28] Hammond, John S., Keeney, Ralph L., and Howard Raiffa, (1998). "Even Swaps: A Rational Method for Making Trade-Offs," *Harvard Business Review*, March-April, 137-149.

[29] Hammond, John S., Keeney, Ralph L., and Howard Raiffa (1999). *Smart Choices: A Practical Guide to Making Better Decisions*, Harvard Business School Press, Boston.

[30] Hurley, W. J. (1998). "An Efficient, Objective Technique for Selecting an All-star Team," *Interfaces*, 28, 51-57.

[31] Hurley, W. J. and Dan Lior (2002). "Combining Expert Judgment: On the Performance of Trimmed Mean Vote Aggregation Procedures in the Presence of Strategic Voting," *European Journal of Operations Research*, 140, 142-147.

[32] Inada, K. (1969). "The Simple Majority Decision Rule," *Econometrica*, 37(3), 490-506.

[33] Janis, Irving L. (1972). *Victims of Groupthink*, Houghton Mifflin, Boston.

[34] Keeney, Ralph L. (1992). *Value Focused Thinking*, Harvard University Press, Cambridge.

[35] Keeney, R. L. and H. Raiffa, (1976). *Decisions With Multiple Objectives: Preferences and Value Tradeoffs*, Wiley, New York.

[36] Kemeny, J. G. and J. L. Snell, (1962). "Preference Ranking: An Axiomatic Approach," in *Mathematical Models in the Social Sciences*, Chapter 2, 9-23, Ginn, Boston.

[37] Kendall, M. (1962). *Rank Correlation Methods*, Hafner, New York.

[38] Kirkwood, C. W. and R. K. Sarin (1980). "Preference Conditions for Multiattribute Value Functions," *Operations Research*, 28, 225-232.

[39] Lehrer, Jonah, (2009). *How We Decide*. Houghton Mifflin Harcourt, Boston.

[40] Loerch, A., R. Koury, and D. Maxwell (1999). "Value Added Analysis for Army Equipment Modernization," *Naval Research Logistics 46(3)*, 233-253.

[41] Martello, S. and P. Toth (1990). *Knapsack Problems*, John Wiley and Sons, New York.

[42] Papamichail, Alves, G., French, S., Yang, J. B. and R. Snowden (2007). "Facilitation Practices in Decision Workshops," *Journal of the Operational Research Society*, 58(5), 614-632.

[43] Phillips, L. D. (2006). "Decision Conferencing," *LSE Working Paper LSEOR 06.85*, London School of Economics.

[44] Phillips, L. D. and M. C. Phillips (1993). "Facilitated Work Groups: Theory and Practise," *Journal of the Operational Research Society*, 44(6), 533-549.

[45] Pugh, S. (1990). *Total Design*, Addison-Wesley, Reading, MA.

[46] Roy, R. L. and Ian Chapman (2006). "An Examination of the Army of Tomorrow's Capability Priorities," DRDC ORD TN 2006-07, Department of National Defence.

[47] Saaty, T.L. (1980). *The Analytic Hierarchy Process*, McGraw-Hill, New York.

[48] Saaty, T.L. (1977). "A Scaling Method for Priorities in Hierarchical Structures," *Journal of Mathematical Psychology*, 15, 234-281.

[49] Saaty, T.L. (1987) "The Analytic Hierarchy Process: What It Is and How It is Used," *Mathematical Modelling*, 9, 161-176.

[50] Saaty, T. L. (1986). "Axiomatic Foundation of the Analytic Hierarchy Process," *Management Science*, 32, 841-855.

[51] Schwarz, R. M. (1994). *The Skilled Facilitator*, Jossey-Bass, San Francisco.

[52] Stillwell, W. G., Seaver, D. A., and W. Edwards (1981). "A Comparison of Weight Approximation Techniques in Multiattribute Utility Decision Making," *Organizational Behavior and Human Performance*, 28, 62-77.

[53] Surowiecki, James (2004). *The Wisdom of Crowds*, Doubleday, New York.

[54] Thorsteinson, T. J. and Scott Withrow, (2009). "Does Unconscious Thought Outperform Conscious Thought on Complex Decisions? A Further Examination," *Judgement and Decision Making*, 4(3)2009, 235-247

[55] Wallenius, Jyrki, Dyer, J. S., Fishburn, P. C., Steuer, R. E. Zionts, S. and K. Deb (2008). "Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead," *Management Science*, 54(7), 1336-1349.

[56] Weingartner, H. (1966). "Capital Budgeting of Interrelated Projects: Survey and Synthesis," *Management Science 12(7)*, 485-516.

[57] Yaniv, Ilan, (1997). "Weighting and Trimming: Heuristics for Aggregating Judgments under Uncertainty," *Organizational Behavior and Human Decision Processes*, 69(3), 237-249.

This page intentionally left blank.

Distribution list

DRDC CORA CR 2009-02

Internal distribution

1 Author (Hard Copy)

1 CORA Library (CD)

1 LFORT (email: Mike.Ormrod@forces.gc.ca)

1 LCDORT (email: Frederick.Cameron2@forces.gc.ca)

1 Section Head Land and Operational Commands OR (email: Dean.Haslip@forces.gc.ca)

1 Section Head Air (email: Denis.Bergeron2@forces.gc.ca)

1 Section Head Maritime (email: Roy.Mitchell@forces.gc.ca)

1 Chief Scientist DRDC CORA (CD)

1 DG DRDC CORA (CD)

1 DDG DRDC CORA (CD)

Total internal copies: 10

External distribution

Department of National Defence

1 DRDKIM (CD)

1 Canadian Forces College Library (CD)

1 Fort Frontenac Library (CD)

International recipients

1 Dr. Richard Deckro (email: Richard.Deckro@afit.edu)
Department of Operational Sciences
Air Force Institute of Technology
AFIT/ENS, Bldg 641
2950 Hobson Way

Wright Patterson AFB OH
45433-7765

1 Dr. Gregory S. Parnell (email: Gregory.Parnell@usma.edu)
Room 432
Mahan Hall
Department of Systems Engineering
United States Military Academy
West Point, NY
10996

1 LTC Lee Ewing (email: plewing@nps.edu)
Glasgow Hall 210
Operations Research Department
Naval Postgraduate School
Monterey, CA
93943

Total external copies: 6

Total copies: 16

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Department of Business Administration Royal Military College of Canada PO Box 17000, Station Forces Kingston, Ontario CANADA K7K 7B4	2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Operational Research Techniques for
Options Analysis in Defence Organizations

4. AUTHORS (Last name, followed by initials – ranks, titles, etc. not to be used.)

Hurley, W.J.

5. DATE OF PUBLICATION (Month and year of publication of document.) November 2009	6a. NO. OF PAGES (Total containing information. Include Annexes, Appendices, etc.) 70	6b. NO. OF REFS (Total cited in document.) 57

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Contract Report

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Defence R&D Canada – CORA
Dept. of National Defence, MGen G.R. Pearkes Bldg., 101 Colonel By Drive, Ottawa, Ontario, Canada K1A 0K2

9a. PROJECT NO. (The applicable research and development project number under which the document was written. Please specify whether project or grant.)	9b. GRANT OR CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC CORA CR 2009-02	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

(X) Unlimited distribution
(   ) Defence departments and defence contractors; further distribution only as approved
(   ) Defence departments and Canadian defence contractors; further distribution only as approved
(   ) Government departments and agencies; further distribution only as approved
(   ) Defence departments; further distribution only as approved
(   ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11)) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

This paper examines approaches to options analysis in a defence setting. These include optimization techniques for problems with many options, a modification of Pugh's Method for multi-criteria decision analysis, and median rank voting. The idea is to present techniques that are not well known to the military operational research community. In addition, I discuss the special organizational characteristics of defence decisions.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Options Analysis
Defence
Multi-Criteria Decision Analysis
Voting
Optimization

DEFENCE DÉFENSE

DRDC CORA

www.drdc-rddc.gc.ca