More Efficient and Flexible Procurement Assessment with a Multi Criteria Decision Analysis Framework

Robert Hayes, Frazer-Nash Consultancy, r.hayes@fncaustralia.com.au

ABSTRACT

With the increasing complexity of capability requirements, available equipment, and integration considerations, assessing and optimizing vessels has become more onerous and time-consuming to perform using traditional methods. To improve the efficiency, flexibility, and auditability of tender evaluations, a desktop tool based on multi-criteria decision analysis was built for a procurement team to automate and visualize their decision-making process.

This paper examines the underlying framework that supports the tool. It introduces the underlying mathematics, provides an outline of the processes and steps involved, and explains what outputs the framework provides to the end user. The paper discusses the merits of using the framework, along with the pitfalls and challenges of using the tool. Finally, the paper presents suggested circumstances in which the tool may or may not be appropriate.

INTRODUCTION

In a procurement program there are a complex set of interrelated factors to consider, and an end-goal which is to maximise capability, minimise risk and provide value-for-money. In a more practical sense this means balancing the three competing drivers (figure 1):

- Capability – the ability of the platform to achieve the mission system;
- Plausibility – the ability of the platform to be viable from an engineering and physical constraint perspective; and
- Affordability – the ability of the platform to be economically viable.

Therefore, a program needs to seek out the solutions which offer a balanced compromise and support these trade-off decisions with quantifiable data.

In most defence procurement programs there are likely to be several tendered platform designs. Each of these may be offered in different configurations or fit-out states, resulting in a large number of potential procurement options to be assessed. Whilst all options must meet the essential requirements in order to be further considered, it is common to specify desirable requirements – those which the war fighter considers would improve the capability or useability of the system. These desirable characteristics hence form some of the criteria against which the options can be assessed.

Performing trade-off assessments of all these options manually is a very time intensive process, which is prone to error and-or bias. It requires a great deal of discipline and documentation to ensure processes are transparent, traceable and fair. Furthermore, when priorities change mid-program or designs are altered, the updated information needs to pass back through the decision-making process to ensure the best decision is still being made. Finally, with the ultimate decision-making resting with stakeholders not involved in the detail of the assessments, presenting them with information summaries that are clear and defensible is paramount.

The method and framework set out in this paper are drawn from work to develop an Integration Impact Assessment tool for DST & Navy. This tool was developed under the specific direction of DST with funding provided by Defence from a current RAN ship procurement project. We would like to acknowledge the contribution by both the RAN and DST in funding and supporting the development of this tool.

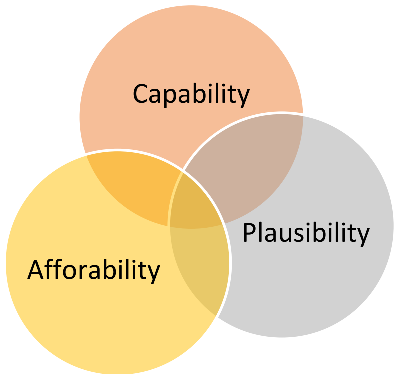

Figure 1: Three competing drivers in a procurement program

THE FRAMEWORK

The Context and Need

This conundrum calls for an efficient and flexible process. Therefore, a procurement program should look to adopt a structured assessment method that is traceable, flexible, scalable, auditable, and repeatable. One such method is to use a decision-making method such as multi-criteria decision analysis (MCDA), backed up by rigorous systems engineering methods. This paper focuses on the framework as a methodology rather than the implementation (e.g. a software tool). Successful implementation depends on the needs and complexities of the procurement program, and any software developed should be bespoke or heavily customised for the end-user to yield maximum benefit.

Developing a framework that can deal with a complex problem in a simple manner, ultimately requires bounds and assumptions to be made. To simplify the problem, we separate the framework into two halves; the first linking capability to the equipment, and the second assessing the impact of integrating that equipment on the platform.

Defining Capability and Relating it to Equipment

Capability of a system can be considered the result of nine fundamental inputs to capability (FICs), namely: personnel, organisation, collective training, major systems, supplies, facilities and training areas, and support. However, when evaluating tendered equipment designs we cannot control any factor but the equipment. Therefore, in terms of maximising our platform's capability, we must assume at the initial stage that equipment is proportional to capability. To determine how our equipment meets our platform's needs, an enterprise architecture such as DODAF should be used (figure 2). The functions of each operational activity can then be mapped to the equipment that enables these functions. This is not a linear mapping and there will be a mixture of one-to-one mapping, one-to-many, and many-to-one of functions to equipment (figure 3).

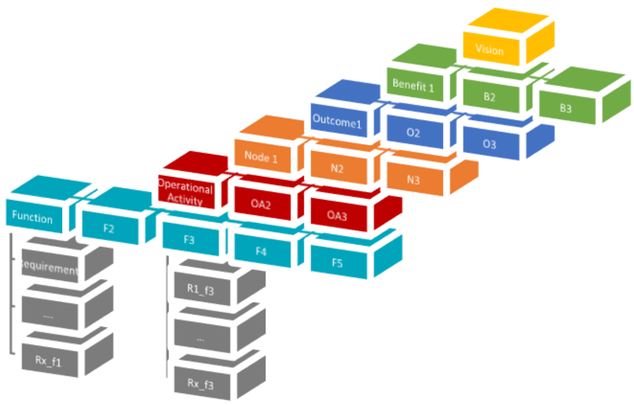

Figure 2: Indicative enterprise architecture

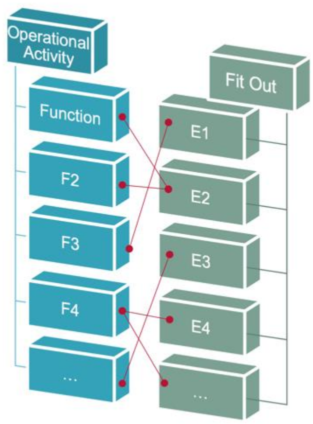

Figure 3: Illustration of mapping equipment to functions

Through these relationships, the functions can be assigned a contribution value (e.g. percentage) towards total platform capability. Since, the equipment is also mapped to these functions, we now have a relationship between equipment and capability. Henceforth, choosing a platform fit-out containing some of this equipment will implicitly indicate which functions have been met and contribution towards capability of that fit-out.

Figure 4: Increasing functions leads to increased capability

To take this further and allow for a multitude of assessments, a procurement program may choose to develop several equipment fit-outs, each representing more equipment. Hence, each fit-out achieves more functions and has an increased capability contribution. For example, you may have three fit-outs, “Fit-out 1, “Fit-out 2”, and “Fit-out 3” that approximately provide 60%, 75%, 90% of desired capability respectively (figure 4). These fit-outs can then be assessed against multiple platforms using MCDA to determine the impact of integration as covered in the next section.

Introduction to Multi-Criteria Decision Analysis

Multi Criteria Decision Analysis (MCDA) is a quantitative tool which helps to aggregate a stakeholder’s preferences against conflicting criteria to find the alternative with the highest overall score when all criteria are considered.

MCDA is comprised of four stages; first is defining the criteria, second is establishing the weighting of each criterion, third is scoring each of a number of discrete options against each criterion, and the fourth stage is calculating the weighted criterion function. The preferred option is provided by the highest value after summation of all weighted criteria functions. This can be automated by a spreadsheet, allowing multiple concurrent calculations to be performed easily.

We use MCDA in this framework to assess the integration difficulty of a number of fit-out options (defined in the previous section) against a number of proposed platforms. These combinations can be set up in a matrix as in table 1.

	Platform Design A	*Platform Design B*	*Platform Design C*
Fit-Out 1	*Option A-1*	*Option B-1*	*Option C-1*
Fit-Out 2	*Option A-2*	*Option B-2*	*Option C-2*
Fit-Out 3	*Option A-3*	*Option B-3*	*Option C-3*

Table 1:Matrix of options for MCDA

In each case we will assess the option against the defined criteria, clustering the criteria in a hierarchy as described below. We will use the product of the weighting and the evaluated score to determine the result for each tier (i.e., the sub-criteria). The results from each tier will be aggregated to provide the evaluated score for the next tier level, until the overall score is determined for all options.

Step 1: Define Criteria

We start by defining the criteria against which the candidate solutions will be assessed. In general, these criteria should consider a plethora of physical and integrated logistics support (ILS) requirements of the platform. Having accurate and representative criteria is crucial to a successful assessment, and input from stakeholders and subject matter experts should be solicited in developing the criteria.

It is recommended that the criteria and sub-criteria are clustered in a structured hierarchy, referred to as a value tree (see Multi-Criteria Analysis: a Manual, 2009, for further details). Using this method, similar criteria are grouped at the same level, and similar groups combined at the subsequent levels up the tree (figure 5). A criterion’s position in the overall value tree can be determined, and competing or duplicate criterion can be identified, and re-defined.

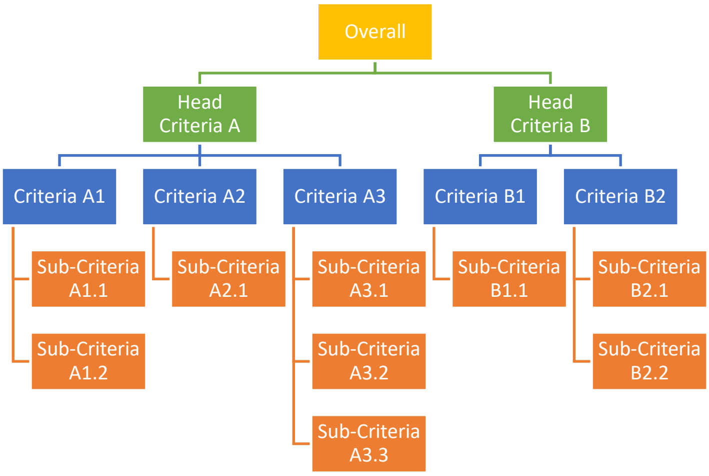

Figure 5: Value tree for determining criteria

Before finalising the choice of criteria, the provisional set needs to be assessed against a range of qualities. If the criteria do not adhere to all of these, it may be necessary to separate the criteria into multiple assessments. These qualities and their justification are:

- Completeness – have all necessary criteria been included? This will be an iterative affair, with multiple passes required to ensure that no major criterion is overlooked, and all the criteria are adequately defined.
- Redundancy – are all criteria necessary? Unimportant or superfluous criteria should be excluded to reduce the complexity of the assessment.
- Operationality – can each option be judged against each criterion? The assessment may be objective (e.g. weight or length) or subjective (e.g. vulnerability) and MCDA can accommodate both simultaneously but the criterion must be defined clearly enough to be assessed.
- Mutual independence of preferences – are scores associated with one criterion able to be assigned without knowledge of the scores of other criteria? This may not always be possible, in which case it may be necessary to combine criterion into a new preference independent criterion.
- Double counting – do any criteria have the same definition? However, just because two criteria appear similar does not necessarily mean there is double counting, likewise, in other instances doubling counting may not be immediately evident. An objective basis cannot be used to determine whether there is double counting, but fortunately checking for independent of preference will generally reveal double counting.
- Size – is there an excessive number of criteria? The structure should not be larger than necessary, and simple evaluation to eliminate criteria which may lead to inconsistencies should be performed early.
- Impacts occurring over time – are criteria affected by future events? MCDA does not have any ability to deal with time-dependent variables. Any criteria that varies with time must be approximated by another method (e.g. a non-linear value function) or modelled externally.

An example of a simple criteria set is in figure 6.

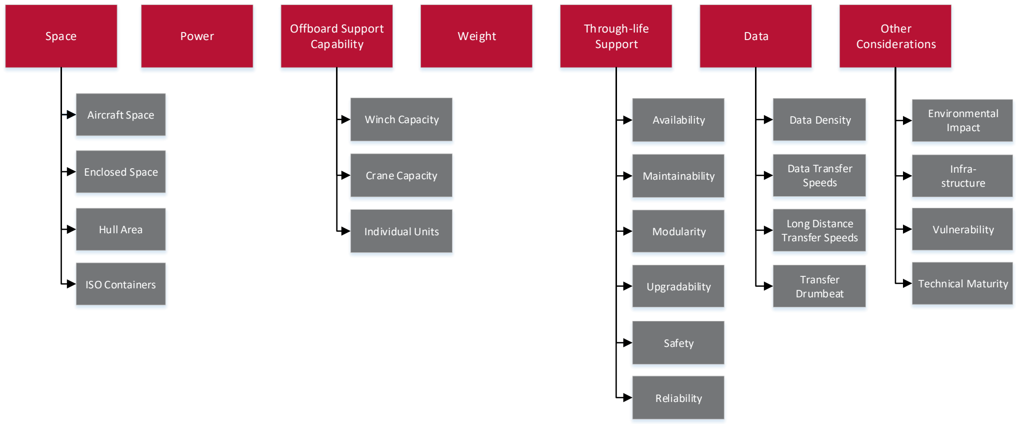

Figure 6: Exemplar criteria for a vessel procurement program

For a particular assessment, we need to populate each sub-criterion with data. In the evaluation (step 4), the data for the platform will be compared against the aggregate data from the fit-out. The criterion and whether it is a physical constraint, a variable value, or a non-physical concept, will drive what mathematical operation is used to ascertain this aggregate value (e.g. summation, averaging, maxima, minima, or otherwise).

For simplicity sake, if we consider a basic physical constraint such as Enclosed Space, we would take the sum of the individual package volume for each piece of equipment in this particular *Fit-Out* and use that define our upper and lower bounds for that criterion. This is because the package volume is the driving requirement for this criterion, that is to say what we as engineers are evaluating when assessing this criterion. Say for example our required value of all equipment is 100m^3, therefore, we can consider the lower bound to be 70% of this value (e.g. 70m^3), as not all the equipment is required to be on the platform at the same time. We can consider the upper bound to 120m^3 as this provides a 20% future upgrade margin. The bounds at this stage can be refined if after an initial pass it is determined they are too onerous or lax. From this example can be seen that having an awareness of what is being assessed is important to ensure the assessment reflects reality,

Step 2: Establishing the Weights of Each Criteria

We need to assign a weighting to each criterion such that when evaluated, the relative merit and preference of this criterion is considered. This also ensures that more important criteria have a stronger influence on the overall assessment than less important criteria.

There are a number of methods possible for eliciting weighting and scoring. This framework utilises a composite approach, with multi-attribute utility functions for scoring, and pair-wise comparisons from the Analytical Hierarchy Process (AHP) to derive weights. AHP is an indirect weighting technique, developed by Thomas R Saaty (Saaty 1987) and uses a series of questions to determine the relative importance of the criteria and hence their weighting. In this framework, the normalisation modifications proposed by Belton and Gear, 1983 and accepted by Saaty, 1994 (Ideal AHP) was implemented to prevent against logical inconsistencies arising from the original method.

It may seem counterintuitive to use pair-wise weighting, as if we are able to non-linearly determine the relationships of the value function (step 3), we may know enough about the function to use a geometrical weighting system. However, this implementation has been used as it enables a first pass to be established with a linear mapping. Further refinement can then be made as more information and understanding is established during the process.

AHP is effective as it does not ask stakeholders to rank criteria, instead asking a series of questions which compel stakeholders to compare two criteria directly to determine which one is more important, and the scale of that importance. This method can be successfully presented to both technical and non-technical stakeholders and if required use fuzzy terms such as strongly/weakly agree/disagree. Furthermore, due its nature of using matrices it is simple to implement and compute using a spreadsheet.

AHP presents a series of questions in turn such as in table 2, and for each question the stakeholder is to indicate their preference one way or the other (or if they are equal) on a scale between 1 and 9, where: 1 = equal; 3 = weakly more important; 5 = strongly more important; 7 = very strongly more important; 9 = absolutely more important.

If there are criteria, there are questions. This practically limits technique to around 20 objectives (190 questions), but more realistically it should be used for between 3 and 8 criteria (i.e. 3 and 28 questions).

Factor	Factor Weighting Score			Factor	
	More Important Than	Equal	Less Important Than		
Criteria 1	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 2	Q1
Criteria 2	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 3	Q2
Criteria 3	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 4	Q3
Criteria 4	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 5	Q4
Criteria 5	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 6	Q5
Criteria 6	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 7	Q6
Criteria 7	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 8	Q7
Criteria 8	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 9	Q8
Criteria 9	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 10	Q9
Criteria 10	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 11	Q10
Criteria 11	9 8 7 6 5 4 3 2	1	2 3 4 5 6 7 8 9	Criteria 12	Q11

Table 2: AHP question format

Step 3: Defining the Value Function for Each Criteria

A value function must be defined for all criteria, this function will include the minimum acceptable value for each criterion, the desirable (or maximum) value for each criterion and a curve approximating the relative value of improving from the minimum value to the desirable/maximum value.

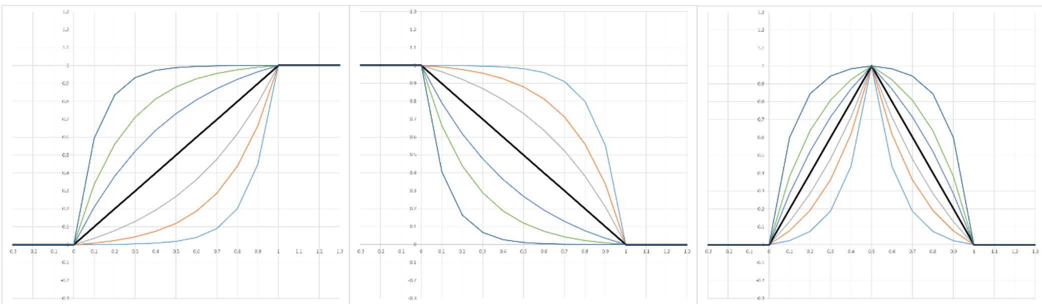

Figure 7: Example curves (from L to R): Increasing, Decreasing, Midpoint with stepwise functions applied

The value curves control how the evaluated value is mapped between the upper and lower limits using an exponential formula controlled by the α-value. For an increasing curve, an α-value of 0 is a linear mapping, a positive value creates a logarithmic curve of increasing steepness, and a negative value creates an exponential curve of increasing steepness. For a decreasing curve, it is the inverse. For midpoint curves, a positive value provides a shallower transition into the midpoint. Furthermore, by applying stepwise functions how values which fall outside the minimum/maximum are treated can be calculated (figure 7 shows such example curves).

These functions are highly flexible but conversely with this flexible, discipline needs to be applied. Experience and knowledge of how the function behaves for the particular criterion is required otherwise the assessment validity may be jeopardised. If in doubt, a basic first pass should always be a linear mapping.

Returning to example from part 1 – Equipment Volume (figure 8); experience may say that adopting an increasing curve with a medium positive α-value (e.g. 5) is most appropriate. This allows the value to rise towards 1.0 quickly, ensure that values that come close to the required value score well, yet leaves there is some headroom for upgradability to be scored.

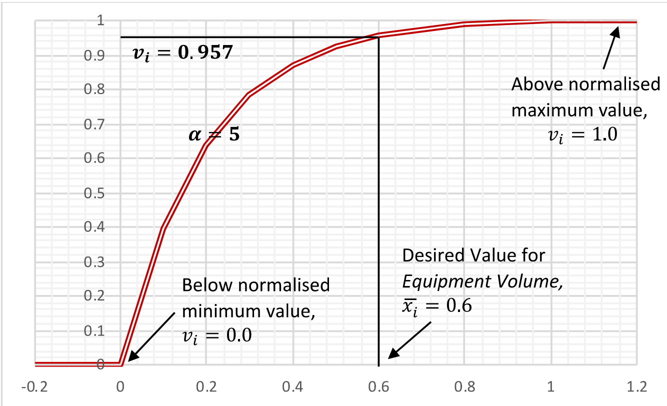

Figure 8: Exemplar value curve for equipment volume example from part 1

An alternative to using the composite approach is to apply REMBRANT (see Lootsma, 1992 & Olsen 1996) or MACBETH (Bana e Costa and Vansnick, 1997 & Bana e Costa et al., 1999) techniques, and more information about these methods can be found at their respective sources or in Multi-Criteria Analysis: a Manual, 2009.

Step 4: Calculate the Criteria Functions

For the last step of MCDA, and penultimate step of the framework we evaluate all the options concurrently. For each option we are determining the impact of integrating of that fit-out on that tendered platform design, against the criteria determined in step 1. The data for this is drawn from two sources; the aggregated equipment data (by means of the fit-outs– touched on in part 1), and the platform data. This forms part of the criteria function, which is for each option is the summation across all assessed criteria, of the product of the value function (from the value curve output) and the weighting.

Looking at table 3, we see the four sub-criteria under the space criteria in figure 6.

Since we are considering impact, we want to minimise this (hence a lower impact). To arrive at this, we invert the MCDA score. The option with the lowest overall impact is best option (e.g. Option A-1 in this example).

Considering the value tree (step 1), by the overall impact in the example above we refer to just one criteria or head criteria. The other criteria are assessed simultaneously, with their own weightings and data, and all these assessments can all be aggregated and used to generate the integration impact for the overall assessment. In this instance, the *total score* generated at the previous level becomes the *value function* at the next level up, and the weightings are calculated for that level. Essentially, this principle of stacking has no limit, but realistically two or three layers would be sufficient for most applications (e.g., sub criteria, criteria, aggregated assessments).

Criteria	Desired Value	Lower Bound	Upper Bound	Alpha	Weights
Aircraft Space	25	15 (60%)	50 (200%)	5	0.412
Enclosed Space	30	20 (67%)	41 (135%)	3	0.049
Hull Area	130	98 (75%)	150 (75%)	2	0.248
ISO Containers	5	5 (100%)	10 (200%)	0	0.29

	Option A-1			Option B-1		
Criteria	**Tendered Value**	**Value Function**	**Wt Value Funct**	**Tendered Value**	**Value Function**	**Wt Value Funct**
Aircraft Space	20	0.514	0.212	25	0.766	0.162
Enclosed Space	30	0.809	0.040	20	0.003	0.000
Hull Area	145	0.970	0.241	100	0.106	0.026
ISO Containers	8	0.600	0.174	10	1.000	0.174
Total Score			0.666			0.362
Impact			33%			64%

Table 3: Exemplar results table from MCDA

This concludes the MCDA part of the framework. The final step of the tool is bridging the capability-to-equipment relationship and the equipment-to-platform integration impact assessment. Therefore, making the equipment the independent variable and capability and integration impact the dependent variables.

Bringing It All Together

The tool that was developed allowed the team to generate their assessments, and for each option have the raw impact assessment values and capability contribution values. However, the tool did not explicitly define how the information and results were summarised. This was because the team did not know at that stage how they wanted to present the information, and wanted to focus on the assessments first and foremost.

In terms of taking this framework forward, we can propose how the information can be presented. The options are either to use charts or tables. The example below gives an indication of a table format. In this example we can see the matrix of options presented, together with their cost in terms of the baseline (e.g. lowest cost option).

	Platform Design A *2.5x Baseline Cost*	*Platform Design B* *1.3x Baseline Cost*	*Platform Design C* *Lowest Cost (Baseline)*
Fit-Out 1 *3.0x Baseline Cost*	*Plausibility 80%* *5.5x Baseline Cost* *90% Capability*	*Plausibility 50%* *4.3x Baseline Cost* *90% Capability*	*Plausibility 30%* *3.0x Baseline* *90% Capability*
Fit-Out 2 *1.4x Baseline Cost*	*Plausibility 90%* *2.9x Baseline Cost* *75% Capability*	*Plausibility 70%* *2.7x Baseline Cost* *75% Capability*	*Plausibility 50%* *1.4x Baseline* *75% Capability*
Fit-Out 3 *Lowest Cost (Baseline)*	*Plausibility 100%* *2.5x Baseline Cost* *60% Capability*	*Plausibility 90%* *1.3x Baseline Cost* *60% Capability*	*Plausibility 70%* *Baseline Cost* *60% Capability*

Table 4: Indicative summary output from framework

By colour coding this matrix correspondingly, we can quickly visualise which options are the best, which are acceptable (with compromise) and which should be discounted.

Together with a knowledge of the cost of the platform and fit-out option leads to setting up a decision-making matrix such as table 4 (figures provided do not relate to previous worked example).

This table, although simple, provides an easily comparable and digestible overview of the options, which can be used by non-technical stakeholders to present and defend decisions.

DISCUSSION

Merits of the Frameworks

MCDA is an agreement process. It allows people to make decisions, informed by analysis. The power of the MCDA method in particular is it models a natural way of thinking, and then if the outcome is counterintuitive or subject to disagreement between stakeholders, sensitivity analysis can explore why and assist in reaching consensus.

In terms of the context of a procurement team it allows the team to compare several options to understand what capability is being achieved, the impact on the platform of achieving that capability, and the cost of that capability, and present this information cohesively and logically, increasing efficiency. Through sensitivity analysis and iteration, the team can understand and quantify trade-offs being made and make an informed decision on whether they are acceptable.

The framework encourages transparency and traceability by documenting criteria, value curves, and AHP judgement scores for each iteration and assessment. Finally, it provides flexibility by allowing a multitude of parameters to be adjusted to suit different assessments and aspects being considered.

Pitfalls of The Framework

It is important to consider that while the framework helps automating decision making, it does not replace a skilled team familiar with what they are procuring. The most important aspect is being mindful and realistic in the scope and limits of the framework.

The experience of procurement program team is key to driving the framework, as the framework requires well defined and accurate criteria and data to be inputted. If used by an inexperienced team there is a risk that the flexibility leads to mistakes being made in the data definition or over complication of scenarios. These in turn will lead to an unusable output and loss of time.

The final risk of the framework is either by intention or not can lead to a team trying to design by numbers rather than design by reality. It must be appreciated that while the tool may encourage combining the best elements, real-world constraints may prohibit this.

Appropriateness of The Framework

The underlying mathematics powering the decision-making MCDA, AHP, is well understood, but subject to significant criticism and debate. It relies on a well-founded and logical scenario to be set up in the assessment. Therefore, it is recommended that the framework be used by a user that is familiar with both the topic being assessed and MCDA. Furthermore, setting up scenarios is a time-consuming process with a steep learning process, so a commitment to using the framework should be made before embarking on the journey. If the assessment is relatively simple, other less complex methods may be more appropriate to get to an answer quicker.

The framework is dependent on the quality of the data input. Therefore, a review of the results should be undertaken by the user to ensure they are reasonable and plausible. The framework should be used to support the decision-making process rather than being relied upon as providing a definitive answer.

Ranking and weighting of criteria is crucial to a good assessment. Before undertaking any complex weightings, a first pass of weighting is recommended to be conducted separately (outside the tool) using another ranking method (e.g. Rank Order Centroid).

Any weighting that is used should consider stakeholder input and experience, and be verified and agreed amongst stakeholders to ensure overall happiness of ranking and order.

The weighting should be respective of the analysis being undertaken (i.e. the priorities should change depending on the characteristics of the platform or system). Additionally, the sensitivities of the input data should be considered, and ideally sensitivity analysis done outside the framework first.

CONCLUSION

In summary we have seen a framework presented that brings together well proven and understood concepts in a matter that is tailored to assessing options in a procurement program. To that end it serves three primary objectives:

- Compare an equipment fit-out against multiple platforms;
- Compare multiple equipment fits-outs against a platform;
- Map equipment to capability – determine what equipment is required.

The tool serves two further secondary purposes:

- Understand the outcome of changing the contribution of functions to capability; and
- Understand the sensitivity of a design to change.

By adopting this framework into a bespoke software tool, the procurement program can increase their efficiency and effectiveness of tender evaluation. This leads to a final outcome where trade-offs are clear, well presented, and platform limitations have been fully considered along the way. Ultimately leading to a successful procurement with no surprises.

ACKNOWLEDGEMENTS

We would like to acknowledge the technical contribution and direction of Karl Slater, Brett Morris, and Dylan Dwyer of DST in the development of this framework (and the resulting Integration Impact Assessment Tool). We would like to further acknowledge Defence in providing funding for the development of the framework and tool from a current RAN ship procurement project.

BIBLIOGRAPHY

1. Multiattribute Value Analysis (MVA), Dr Quoc Do and Prof Stephen Cook, University of South Australia.
2. The Engineering Design of Systems: Models and Methods, Dennis M. Buede, Wiley Publishing, 2000
3. The Analytical Hierarchy Process – What it is and how it is used, R.W. Saaty, Math Modelling, Vol 9, No 3-5, pp 161-176, 1987 https://core.ac.uk/download/pdf/82000104.pdf
4. Multi-criteria analysis: a manual, Department for Communities and Local Government: London, 2009
5. Lootsma, F.A. (1992) The REMBRANDT System for Multi-criteria Decision Analysis via Pairwise Comparisons or Direct Rating, Report 92–05, Faculty of Technical Mathematics and Informatics, Delft University of Technology, The Netherlands
6. Olson, D. (1996) Decision Aids for Selection Problems, Springer Verlag, New York
7. Belton, V,. and Gear, T. (1983) 'On a short-coming in Saaty's method of analytic hierarchies', Omega, 11, pp.228–30
8. Saaty, T.L. (1994). Fundamentals of Decision Making and Priority Theory with the AHP. RWS Publications, Pittsburgh, PA, U.S.A.