Utility-Based Decision Support for Investment and Acquisition

Paper presented at the 30th International Symposium On Military Operational Research (30 ISMOR), Royal Holloway (University of London), 29th July - 2nd August 2013

John Moore
QINETIQ/13/01580
Version 1.00 dated 19 July 2013

QinetiQ Ltd
Cody Technology Park
Farnborough
Hampshire
GU14 0LX
Administration page

<table>
<thead>
<tr>
<th>Principal author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Room 1026, A5 Building, Cody Technology Park, Farnborough, GU14 0LX</td>
</tr>
<tr>
<td>UC: <a href="mailto:jmoore3@qinetiq.com">jmoore3@qinetiq.com</a></td>
</tr>
<tr>
<td>R: <a href="mailto:jmoore3@qinetiq.r.mil.uk">jmoore3@qinetiq.r.mil.uk</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Release Authority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
</tr>
<tr>
<td>Post</td>
</tr>
<tr>
<td>Date of issue</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Record of changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue</td>
</tr>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>1.0</td>
</tr>
</tbody>
</table>
Abstract

This paper presents the current status of an investigation, conducted by QinetiQ Ltd, into ways of ensuring that the evaluation processes for defence equipment and services are demonstrably consistent with the legal requirements on transparency arising from DSPCR 2011 and recent case law. It derives from a recent study, conducted by QinetiQ for the Policy and Capability section of the UK Defence Science and Technology Laboratory (Dstl PCS), supplemented by further development work conducted by QinetiQ under internal funding.

The paper proposes multivariate utility theory as a method for ensuring fairness, transparency and rigour in developing and communicating the decision criteria for defence acquisitions. It identifies the issues which the decision-making process must address, and develops a solution, in which the academic formulation of utility theory is developed into a workable, repeatable, and auditable process. The utility-based approach is compared with the weighted-sum methods recommended in DSPCR 2011, and their respective strengths and weaknesses are identified.

The method first generates a joint benefit utility function, which evaluates tradeoffs between benefit criteria and assigns an "overall benefit" score to any combination of outcomes against the individual benefit criteria. It then generates a cost-benefit tradeoff evaluation function, which likewise assigns an overall score to any combination of outcomes against cost and overall benefit. The cost-benefit function can be expressed in the form of a two-dimensional "heat map": a chart of benefit versus cost in which the chart background is colour-coded to indicate visually the overall score that a purchasing authority would assign to any offering, based on its tendered cost and its overall benefit score. The joint utility functions and heat map can be generated and published in advance of the submission of offerings.

It is shown that, because of its greater adaptability to elicited stakeholder preference, the method can be applied to a significantly greater range of problems than weighted-sum methods.
# List of contents

<table>
<thead>
<tr>
<th>Administration page</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>3</td>
</tr>
<tr>
<td>List of contents</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td><strong>Background</strong></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td><strong>Decision Criteria for Defence Procurement</strong></td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>8</td>
</tr>
<tr>
<td><strong>Statutory Requirements for Public Contract Award</strong></td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>9</td>
</tr>
<tr>
<td><strong>Decision criteria for Public Contract Award</strong></td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>9</td>
</tr>
<tr>
<td><strong>Ministry of Defence Guidance</strong></td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>10</td>
</tr>
<tr>
<td><strong>Limitations in current Guidance</strong></td>
<td></td>
</tr>
<tr>
<td>3.5</td>
<td>11</td>
</tr>
<tr>
<td><strong>COEIA Process and Outputs</strong></td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>12</td>
</tr>
<tr>
<td><strong>Time, compliance and risk</strong></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td><strong>Problem Formulation</strong></td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>14</td>
</tr>
<tr>
<td><strong>Requirement</strong></td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>15</td>
</tr>
<tr>
<td><strong>Formalisation</strong></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td><strong>The Joint Utility Method</strong></td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>17</td>
</tr>
<tr>
<td><strong>Overview of Method</strong></td>
<td></td>
</tr>
<tr>
<td>5.2</td>
<td>17</td>
</tr>
<tr>
<td><strong>Example Problem</strong></td>
<td></td>
</tr>
<tr>
<td>5.3</td>
<td>19</td>
</tr>
<tr>
<td><strong>Aggregation of Cost Metrics</strong></td>
<td></td>
</tr>
<tr>
<td>5.4</td>
<td>20</td>
</tr>
<tr>
<td><strong>Developing Utility Functions for Benefit Criteria</strong></td>
<td></td>
</tr>
<tr>
<td>5.5</td>
<td>25</td>
</tr>
<tr>
<td><strong>Worked Example: Cost and Benefit Utility Functions</strong></td>
<td></td>
</tr>
<tr>
<td>5.6</td>
<td>26</td>
</tr>
<tr>
<td><strong>Aggregating Benefit Criterion Utilities</strong></td>
<td></td>
</tr>
<tr>
<td>5.7</td>
<td>28</td>
</tr>
<tr>
<td><strong>Elicitation of the Overall Benefit Function</strong></td>
<td></td>
</tr>
<tr>
<td>5.8</td>
<td>33</td>
</tr>
<tr>
<td><strong>Cost-Benefit Tradeoffs</strong></td>
<td></td>
</tr>
<tr>
<td>5.9</td>
<td>40</td>
</tr>
<tr>
<td><strong>Option Evaluation</strong></td>
<td></td>
</tr>
<tr>
<td>5.10</td>
<td>46</td>
</tr>
<tr>
<td><strong>Comparison of Weighted-Sum and Joint Utility Benefit Functions</strong></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>48</td>
</tr>
<tr>
<td><strong>Comparison of Joint Utility Method with DSPCR Guidance</strong></td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>48</td>
</tr>
<tr>
<td><strong>Example of DSPCR tender evaluation process</strong></td>
<td></td>
</tr>
<tr>
<td>6.2</td>
<td>49</td>
</tr>
<tr>
<td><strong>Issues with the DSPCR tender evaluation process</strong></td>
<td></td>
</tr>
<tr>
<td>6.3</td>
<td>50</td>
</tr>
<tr>
<td><strong>Utility-Based formulation</strong></td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>52</td>
</tr>
<tr>
<td><strong>Capturing the Purchaser’s Cost-Benefit Tradeoff Preferences</strong></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>55</td>
</tr>
<tr>
<td><strong>Summary and Conclusions</strong></td>
<td></td>
</tr>
<tr>
<td>7.1</td>
<td>55</td>
</tr>
<tr>
<td><strong>Summary of Proposed Method</strong></td>
<td></td>
</tr>
<tr>
<td>7.2</td>
<td>55</td>
</tr>
<tr>
<td><strong>Assessment</strong></td>
<td></td>
</tr>
<tr>
<td>7.3</td>
<td>57</td>
</tr>
<tr>
<td><strong>Conclusions and Observations</strong></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>60</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td></td>
</tr>
<tr>
<td>Annex A: The Multivariate Joint Utility Functions</td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>61</td>
</tr>
<tr>
<td><strong>Formulation</strong></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>62</td>
</tr>
<tr>
<td><strong>Derivation: Joint Impact Utility Function</strong></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>63</td>
</tr>
<tr>
<td><strong>Joint Criticality Utility Function</strong></td>
<td></td>
</tr>
<tr>
<td>Document reporting page</td>
<td>65</td>
</tr>
</tbody>
</table>
1 Introduction

This paper reviews the potential of utility-based Decision Theory techniques as a means of formulating and expressing the award criteria for public acquisition decisions, with particular reference to defence acquisition. It identifies the issues which the decision-making process must address, and identifies a possible solution, in which the academic formulations of these techniques have been developed into a workable, repeatable and auditable process.

This document consists of the following sections:

1. Introduction.

2. Background: This section describes the developments in the regulatory structure for defence and security acquisition that created the need for a new formulation of acquisition award criteria.

3. Decision Criteria for Defence Procurement. This section describes the key criteria for defence procurement, how they are currently addressed, and the limitations in the current processes which this study seeks to address.

4. Problem Formulation: This section gives a formal specification of public acquisition as a problem in Decision Theory

5. The Joint Utility Method: This is the principal section of the paper. It develops a generic solution to the problem set out in Section 4, drawing on the techniques of utility theory and Multiple-Criteria Decision Analysis (MCDA). The exposition is based around a hypothetical real-world example.

6. Comparison of Joint Utility Method with DSPCR Guidance. This section compares the process with the Tender Evaluation methods recommended in the Defence and Security Public Contract Regulations (DSPCR), and demonstrates that they offer a viable solution to a number of serious deficiencies in the recommended methods.

7. Summary and Conclusions.

Annex A: This annex derives the recommended form for a joint multivariate utility function for use in the proposed method.
2 Background

The process described in this paper was developed in the course of the COEIA Transparency Pilot Study, which was commissioned from QinetiQ by the UK Defence Science and Technology Laboratory (Dstl) in December 2012 and reported in April 2013 (Refs 1, 2, 3). The study report (Reference 3) included a trial application, using data from an actual COEIA conducted by Dstl. The data used in this trial were subject to security and commercial caveats which mean that they cannot be reproduced here: a hypothetical example is presented instead.

The study addressed the need to ensure that the approvals process for the procurement of defence equipment and services (currently SMART Approvals, Reference 5) is consistent with the legal requirements on transparency arising from recent legislation, in particular the Defence and Security Public Contract Regulations 2011 (DSPCR 2011) (Reference 6), and case law. It was tasked to investigate the potential of a number of “textbook” decision analysis techniques to strengthen the combined operational effectiveness and investment appraisal (COEIA) process, so as to minimise the risk that future decisions based on COEIA outputs could be held to contravene these statutory requirements. The investigation covered both the processes used to generate the outputs of the COEIA and the way in which these outputs feed into the award process. Three specific techniques were selected for assessment: MCDA, utility theory and indifference curves.

The principal outcome of the study that Qinetiq conducted for Dstl was the articulation of the utility-based approach which is described in this paper, along with guidance as to the type of problem to which the approach was best suited, and recommendations for further development, testing and evaluation. Since the completion of the study, QinetiQ has undertaken a limited programme of development with a view to more general application. Since this is work in progress, this document is to be understood as a progress report on the development of the utility-based approach rather than a formal specification.

---

1 This study was subject to DEFCON 503, which grants exploitation rights to MoD while the intellectual property rights (IPR) for the process remain with QinetiQ.
3 Decision Criteria for Defence Procurement

3.1 Statutory Requirements for Public Contract Award

The statutory requirements for public contracts have changed significantly in recent years: these changes have arisen partly from the introduction of new legislation and statutory regulation, most notably the Public Contract Regulations 2006 (PCR 2006) and the Defence and Security Public Contract Regulations 2011 (DSPCR 2011) (Reference 6), and partly from case law such as the 2008 Newham ruling ([Letting International Ltd v London Borough of Newham, Reference 7]). A consistent feature of the changes is the demand for “total” transparency in the award criteria against which tenders for public sectors are awarded.

This is articulated in the Newham ruling as follows:

- “The contracting authority must state the award criteria that it intends to apply in the contract documents or contract notice. Further, potential tenderers should be aware of all the elements to be taken into account by the contracting authority … when they are preparing their tenders.”

- “A contracting authority cannot apply weightings or sub-criteria in respect of award criteria which it has not previously brought to the tenderers’ attention; [and] the criteria and conditions governing each contract must be adequately publicised by the contracting authorities.”

DSPCR 2011 likewise stipulates that:

- “You must award the contract either on the basis of the “lowest price” or ‘most economically advantageous tender’.

- Award criteria must be “objective and linked to the subject matter of the contract in order to … determine which is most economically advantageous.”

- “You must disclose the evaluation criteria [and their] relative weightings [and] the [assessment] methods you will use.”

- All criteria used by a contracting authority, and their relative importance, must be disclosed in advance to tenderers; “weightings or rankings applied to the award criteria [must be published] in the contract documents”. (Reference 6, 13).

- “Recent case law has indicated a need for total transparency in respect of evaluation criteria.”

It can be assumed that the requirement for pre-disclosure applies whether or not a formal MCDA assessment is to be used.²

---

² The cheapest compliant bid solution can be characterised as a special case of MCDA, in which cost is assigned an overwhelming weighting.
3.2 Decision criteria for Public Contract Award

DSPCR 2011 states that the criteria for public contract award “may include, but are not limited to:”

- Quality
- Price
- Technical merit
- Functional characteristics
- Environmental characteristics
- Running costs
- Life cycle costs
- Cost-effectiveness
- After-sales support and technical assistance
- Delivery date and delivery period or period of completion
- Security of supply
- Interoperability and operational characteristics

This is a generic list, and criteria such as “Quality” and “Price” have to be understood in a sense that is appropriate to the nature of the proposed acquisition. The Ministry of Defence provides specific guidance for defence acquisitions, which is summarised in the following sections.

3.3 Ministry of Defence Guidance

The primary source for UK MoD guidance on the appraisal of defence acquisition of options is Joint Service Publication JSP 507: MoD Guide to Investment Appraisal and Evaluation (Reference 9). This identifies three main forms of appraisal for potential investments:

a) **Cost Benefit Analysis** (CBA), in which “all of the costs and benefits of an activity are quantified and valued in monetary terms.” The results of a CBA can be used not only to say which option is best, but also to indicate whether this option is worthwhile, i.e. does it provide a benefit exceeding its cost.” However, “as few activities within MoD produce benefits that can be valued in monetary terms, the use of full-blown CBA by MoD is extremely limited.”

b) **Cost Effectiveness Analysis** (CEA), “which estimates the net present cost of alternative ways of achieving the same requirement. When there are differences in the extent to which the requirement is achieved, these will be noted, and as far as possible quantified, using measures which may be judgemental.”

---

3 In this paper, outside the specific context of CBA, we define a benefit as any “good” sought by the stakeholders to an decision, irrespective of whether its desirability can be expressed in monetary terms. In particular, the effectiveness outputs of a COEIA are characterised as benefits.

4 This definition strictly speaking excludes any consideration of trade-offs between cost and effectiveness. However, in this paper, we apply the more general usage in which the term “Cost-Effectiveness Study” is equally applicable to a study which seeks to identify such trade-offs.
c) **Combined Operational Effectiveness and Investment Appraisal** (COEIA), in which “the total through-life costs of the options to meet a particular requirement are estimated in the Investment Appraisal. The individual parameters contributing to overall performance are identified, and each option assessed against each of these parameters in the Operational Effectiveness Assessment. The two separate assessments are then combined to identify the overall cost effectiveness of each option.”

COEIA is characterised as “a highly formalised type of CEA … used for appraisal of new military equipment, and for other appraisals where the options to meet a requirement offer different levels of military or business effectiveness.”

The purpose of a COEIA is to identify, as far as possible, the “real world” impact of the options being considered. A COEIA may well show that a solution which does not satisfy all the stated requirements is nevertheless capable of delivering the desired real-world outcomes at a much lower cost than a fully-compliant solution; or that requirement is impossible to achieve at an affordable cost. JSP 507 states:

- “Once an appraisal has been completed, both it and the recommendation for action arising from it must be independently scrutinised … to ensure that value for money is likely to be achieved… Scrutiny should not be limited to the recommendation for action, but should encompass the requirement itself. Requirement scrutiny is designed to ensure that only fully justified requirements, which demonstrably contribute to the organisation’s agreed outputs, are approved.”

This provision emphasises the need for COEIAs to cover the broadest possible scope of options and levels of investment, including options which fall short of or “over-achieve” against some of the requirements.

### 3.4 Limitations in current Guidance

The current version of JSP 507 dates from April 2011 and provides no explicit guidance on the following issues, which are implicit in the transparency requirements listed above:

a) Formalisation of the means for evaluating cost-benefit tradeoffs: in particular, the mechanism by which the selection of an option which is more expensive than a minimally-compliant option would or could be justified.

b) The treatment of award criteria other than cost and predicted operational effectiveness: these may include risk, timeliness and compliance with broader public policy (see Section 4 below).

c) The need for the award criteria to be independent of the characteristics of the specific solutions to which they will be applied.

Since JSP 507 is subject to periodic revision, there is a reasonable expectation that these issues will be addressed in subsequent versions. Indeed, the COEIA Pilot Study was envisaged as a potential input to this process.
3.5 COEIA Process and Outputs

The COEIA formulation generates a 2-D plot of “cost” (derived from investment appraisal (IA)) and “operational effectiveness”. The cost metric is usually whole-life cost (WLC) expressed as net present value (NPV).

JSP 507 states

- “Quantitative techniques [for evaluating effectiveness] are the preferred route at all times”;
- Where quantitative techniques are not feasible or sufficient, “recourse must be made” to qualitative assessment: in this case, the outputs must be designated as Measures of Benefit (MOB) rather than Measures of Effectiveness (MOE).”
- “Cost should not be included as a contribution to effectiveness or benefit along the y-axis [effectiveness axis]” (emphasis in original).
- Cost-effectiveness is a relative concept: “it can only be affirmed that one investment is more cost-effective than another, not that it possesses some intrinsic, absolute ‘cost-effectiveness’”.

It is implicit in this guidance that tradeoffs between cost and benefit lie outside the scope of a COEIA; the COEIA’s job is to provide robust and validated data from which such tradeoffs can be assessed.

Examples of possible MOEs for use in COEIAs include:

- Probability of victory or mission success
- Time required to complete mission
- Expected casualties
- Exchange ratio (i.e. expected number of kill per loss)
- Throughput (for a logistics system; typically measured in tons or litres per day)

Figure 1: The efficient set from a hypothetical COEIA assessment (Reference 11, adapted)
Figure 1 shows a hypothetical COEIA output. Each option is subject to uncertainty in both the cost and the effectiveness domains: this is represented in Figure 1 by the ellipses which enclose the possible range of cost and capability outcomes. In this example, Option 2 is shown to be both more expensive and less effective than Option 1: we say that Option 2 dominates Option 1. Likewise, Option 3 dominates Option 4. Options 2 and 4 can therefore be eliminated. The remaining set of non-dominated options forms the “efficient set”. Note that Option 5 is retained, even though its most likely outcome is costlier and less effective than that of Option 3, because the uncertainty analysis shows that Option 5 has a chance of ultimately proving to be the more effective.

This approach leaves the down-selection to be made on policy grounds. Historically, the criteria used by policy-makers have rarely been published, scrutinised or exposed to debate. However, the requirement now is for these criteria to be formally defined in advance of the option evaluation. This implies the development of a formal process for evaluating cost-effectiveness tradeoffs in policy terms.

Note that such a process may bring back into contention options which would be rejected in a pure cost-effectiveness comparison. For example, Option 4 in Figure 1 may cease to be dominated by Option 3 if Option 3 has a seriously adverse environmental impact, or is dependent on an overseas supplier who may be subject to adverse political influence.

3.6 Time, compliance and risk

In general, the evaluation of public investment proposals is not limited to cost and predicted benefit, but also includes other considerations. These may include:

- **Time**: How long it will take for the proposed investment to begin generating the benefits which represent the return on the investment.

- **Compliance**: Does the proposed investment promote, or at least conform to, public policy in areas other than those into which the benefits are targeted? In the defence area, investment (or lack of it) may have major impact on policy areas such as foreign policy, sustainability, industrial policy, health and safety, treaty obligations and international law.

- **Risk and robustness**: Investment options may vary greatly in the degree of certainty with which their outcomes can be predicted. For example, an option which relies on an innovative technology may offer the possibility of very high cost-effectiveness but an appreciable risk of failure or non-compliance. It is legitimate to accept a premium in higher cost or reduced effectiveness in

---

5 Depending on the Concept of Analysis adopted, the uncertainty results will normally be displayed as an ellipse or polygon which encloses between 80% and 90% of the cost-benefit outcomes which are consistent with the COEIA modelling and data assumptions. In some COEIAs, the spread of displayed outcomes represents the spread of outcomes across different scenarios or sensitivity cases, as well as or instead of uncertainty in prediction.

6 In decision theory, option A dominates option B if A is at least as good as B against every decision criterion, and better than B against at least one criterion. An option which is not dominated by any other option is said to be admissible or efficient.
exchange for a more robust solution; however, this implies a tradeoff process which must also be transparent and pre-disclosed.

Of these criteria, only risk is identified as an area for tradeoff in the traditional COEIA process. This is done by defining uncertainty bounds around the cost and effectiveness estimates produced by the assessment, as in Figure 1. As with cost-benefit tradeoffs, JSP 507 does not currently prescribe a formal process for evaluating risk tradeoffs as part of the COEIA process.

Timeliness is usually dealt with in a COEIA by assuming that all the options considered meet the in-service date (ISD) or full operational capability (FOC) target date specified in the user requirement or single statement of user need (SSUN). In general, there is no mechanism for ascribing additional capability to options which can be delivered at an earlier date. COEIA effectiveness assessments are usually based on “snapshots” of operations typically 10 and 20 years ahead. Thus a COEIA assessment based in, say, 2020 will be indifferent between a concept with FOC in 2014 and one with FOC in 2019.

Compliance with broader policy requirements is normally excluded from the COEIA process altogether, except in so far as it is considered under the heading of “other contributory factors (OCF)”. In general, non-compliant options are removed from the COEIA process as soon as they are identified, since effort expended on evaluating them would be regarded as nugatory. Consideration under OCF then has little practical significance, since OCF assessment is generally stand-alone and does not modify in any way the cost-effectiveness output.

In the broader approvals process, compliance tends to be treated as a binary quantity; either the option is compliant or it is not. However, this approach is not consistent with best practice in decision theory. Treating decision criteria as binary gives them effectively 100% weighting in the decision process, and implies discontinuity in the response properties of the assessment to changes in assumptions with regard to compliance requirements. Binary decision criteria accordingly violate the theoretical requirements of Pareto efficiency and continuity (Refs 9, 10). Moreover, successive governments have in practice routinely made tradeoffs between the imperatives of defence and other policy areas. It is accordingly in the interests of Defence that MoD should be well-equipped to advise central government on the most efficient range of tradeoffs.

7 Reference 9, paragraph 3.2.4

8 A decision process is Pareto-efficient if it rejects all dominated options (Reference 10: see also Wikipedia: Pareto efficiency). The term derives from the economist Vilfredo Pareto (1848-1923).
4 Problem Formulation

This section formulates, as a problem in decision theory, the requirements for a process for options or tender assessment that meets both the transparency requirements of public procurement in general, and defence procurement in particular.

4.1 Requirement

It follows directly from DSPCR 2011 and JSP 507 that the process must fully satisfy the following criteria:

1. It must be possible to fully specify the criteria against which a tender is assessed, the rules used for scoring tenders against these criteria, and the rules used for evaluating tradeoffs between criteria, without any knowledge of or assumptions about the tenders that will be submitted, and to supply this information to potential bidders as part of the Invitation to Tender.

2. The scoring criteria must be objective and stable: that is:
   - the scoring of any one tender must depend solely on its own characteristics;
   - the comparison between any two tenders must not be influenced by the presence or absence of a third tender.

3. The process must be formal and communicable: for example, in the form of a practitioners’ guide.

4. The process must be auditable: there must be no ambiguity as to whether the process is being followed in a particular application; if the process offers a choice of procedures, it must always be possible to establish which was used, and the evidence on which that decision was made.

5. The process must be practicable in terms of timescale, affordability and information demand.

6. The process must be theoretically sound and conform to the basic requirements of decision theory with regard to rationality and consistency.

We further assume that the transparency requirement applies with particular force to the methods used for trading off costs and revenues with non-monetary benefits and constraints. This is because compliance with the two criteria which are admissible in DSPCR 2011 for contract award, “lowest price [compliant] tender” or “most economically advantageous tender,” cannot be verified unless these tradeoffs are transparent.

The simplest way to achieve this transparency is to calculate costs and non-monetary outcomes separately and then define a stand-alone procedure for trading them off. This procedure is particularly desirable since it lends itself to a clear form of visual presentation, the “indifference map” or “heat map” (see Section 5.8)). We therefore specify a further process requirement:
a) Overall metrics for monetary and non-monetary decision criteria are calculated separately, and only traded off against each other at the final stage of the process.

4.2 Formalisation

We formalise the problem as follows.

1. The problem domain consists of \( m \) cost criteria, indexed \( i = 1 \) to \( m \), each of which is characterised by a metric \( c_i \); and \( n \) benefit criteria, indexed \( j = 1 \) to \( n \), each of which is characterised by a metric \( x_j \). These metrics are numerical and can be either discrete or continuous. The set of possible values for metric \( c_i \) is designated \( C_i \), and the set of possible values for metric \( x_j \) is designated \( X_j \).

2. A metric is said to be **positive** if, for all \( u, v \) in \( C_i \) or \( X_j \), \( u < v \) if and only if the outcome represented by \( v \) for criterion \( i \) or \( j \) is more favourable than the outcome represented by \( u \). A metric is said to be **negative** if, for all \( u, v \) in \( C_i \) or \( X_j \), \( u < v \) if and only if the outcome represented by \( v \) for criterion \( i \) or \( j \) is is less favourable than the outcome represented by \( u \). All metrics are either positive or negative.

3. Each cost metric has a critical value \( C_{CRIT} \) and a compliant value \( C_{COMP} \). \( C_{CRIT} \) represents the “worst viable” outcome against criterion \( i \): outcomes worse than \( C_{CRIT} \) against any criterion represent a serious threat to the viability of a solution, meaning that significant sacrifices against other criteria would be justified in order to improve a sub-critical outcome. \( C_{COMP} \) represents a satisfactory outcome against criterion \( i \), such that stakeholder groups would be reluctant to make appreciable sacrifices against other criteria in order to further improve an already-compliant outcome. For all positive metrics, \( C_{CRIT} < C_{COMP} \); likewise, for all negative metrics, \( C_{COMP} < C_{CRIT} \).

4. Each benefit metric has a critical value \( X_{CRIT} \) and a compliant value \( X_{COMP} \). These correspond in significance to the critical and compliant values of the cost metrics. For all positive metrics, \( X_{CRIT} < X_{COMP} \); likewise, for all negative metrics, \( X_{COMP} < X_{CRIT} \).

We require:

1. An overall cost function, \( C[c_1 \ldots c_m] \) which maps the domain \{ \( C_1 \ldots C_n \) \} onto the set of real numbers, such that:
   
   a) \( C[C_{CRIT_1}, \ldots, C_{CRIT_n}] = 0 \) \hspace{1cm} Equation 1
   
   b) \( C[C_{COMP_1}, \ldots, C_{COMP_n}] = 100 \) \hspace{1cm} Equation 2
   
   c) If cost metric \( i \) is positive, then for all \( u, v \) in \( C_i \), and, over \( r \neq j \), for all \( c_r \) in \( C_r \), \( C[c_1, \ldots, u, \ldots, c_i] < F[c_1, \ldots, v, \ldots, c_n] \) only if \( v < u \) \hspace{1cm} Equation 3

---

\(^9\) Sets are indicated by **bold** type, and vectors by *underlining*. 

QINETIQ/13/01580 

QinetiQ Proprietary
2. An overall benefit function, \( B[x_1 \ldots x_n] \) which maps the domain \( \{X_1 \ldots X_n\} \) onto the set of real numbers, such that:
   
   a) \( B[X_{CRIT_1}, \ldots X_{CRIT_n}] = 0 \) 
   
   b) \( B[X_{COMP_1}, \ldots X_{COMP_n}] = 100 \) 
   
   c) If metric \( j \) is positive, then for all \( u, v \) in \( X_j \), and, over \( i \neq j \), for all \( x_i \) in \( X_i \), 
   
   \[ B[x_1, \ldots u, \ldots x_n] < F[x_1, \ldots v, \ldots x_n] \onlyif u < v \] 
   
   Equation 7
   
   d) If metric \( j \) is negative, then for all \( u, v \) in \( X_j \), and, over \( i \neq j \), for all \( x_i \) in \( X_i \), 
   
   \[ B[x_1, \ldots u, \ldots x_n] < F[x_1, \ldots v, \ldots x_n] \onlyif v < u \] 
   
   Equation 8

3. A cost-benefit function, \( F[c, b] \) which maps the real x-y plane onto the set of real numbers, such that:
   
   a) \( F[0, 0] = 0 \) 
   
   b) \( F[100, 100] = 100 \) 
   
   c) For all \( u, v, b \), \( F[u, b] < F[v, b] \) only if \( u < v \) 
   
   d) For all \( u, v, c \), \( F[c, u] < F[c, v] \) only if \( u < v \) 
   
   Equation 9

   Equation 10

   Equation 11

   Equation 12

For all three metrics, conditions (c) and (d) ensure a basic rationality criterion: if all other criteria are unchanged, an improvement against a single criterion can never produce a reduction in the objective function. In conjunction with criteria (a) and (b), this in turn ensures that, if each metric \( j \) has a value in the range \([CCRIT_j, \ldots CCOMP_j]\) or \([XCRIT_j, \ldots XCOMP_j]\), then the values of \( C, B \) and \( F \) will all lie in the range \([0, 100]\). 

---

10 Ideally, a unilateral change against any one criterion should produce a non-zero change (however small) in the objective function. The joint utility method developed in this paper ensures this whenever a metric lies in the range \([XCRIT, XCOMP]\), but not for extreme values: for each metric, there is a limit beyond which further changes cease to impact on the objective function. Work to date indicates that the practical significance of this limitation is slight, since options which have extremely poor outcomes against significant criteria are likely to be decisively rejected in any case, while the value of extreme levels of overachievement is deliberately discounted even before these constraints come in play.

11 Theoretical treatments of utility theory such as Keeny and Raiffa [Reference 14] normally define an objective function on the range \([0,1]\). We use a 0 to 100 scale for two reasons: intuitive comparisons between outcomes are easier to make on a 0-100 scale; and a 0 to 100 scale allows us to approximate the continuous objective functions by functions returning integer values, without significant loss of resolution.
5 The Joint Utility Method

5.1 Overview of Method

The proposed solution can be summarised as follows.

1. The assessment criteria for the required acquisition are identified.

2. A unitary cost metric is developed, using the established processes of Investment Appraisal.

3. Metrics are assigned to each benefit criterion other than cost: these may explicit measures of real-world quantities (such as probability of survival), or scores based on qualitative assessment (e.g. 0 = “unacceptable”, 1 = “poor”, 2 = “moderate”, etc.)

4. Criterion utility functions are elicited for each metric.

5. An aggregated “Measure of Benefit” function is developed as an analytic function of the utility scores obtained for the individual benefit criteria other than cost.

6. A bivariate joint utility function is developed for the cost and aggregate benefit.

7. A “heat map” for the cost/benefit tradeoff space is generated from the joint utility function and plotted.

It will be shown that, supplemented by formal specifications of the criterion and joint utility functions, the heat map provides a simple means of communicating to potential tenderers the precise criteria by which the tender which offers the most favourable cost-benefit trade-off will be selected.

These elements are developed in this section through a worked example. For clarity of exposition, and to demonstrate the broader applicability of the proposed method, a simple example from the non-defence domain is used.

5.2 Example Problem

We consider the following problem:

Clara is the proprietor of a small independent clothing shop. Her nightwear supplier has unexpectedly gone into liquidation. She needs to choose a new supplier quickly, before her remaining stock is exhausted. Possible suppliers range from high-volume SE Asian manufacturers to small local companies; she aim to sell at least 150 garments per week and to pay about £6 per garment.

Other major considerations include:

- Product quality and range
- Ethically sourced
- Reliability and assurance of supply
After consultation with the other stakeholders to this decision, Clara secures a consensus on the following set of decision criteria:

1. Cost per garment
2. Capacity of supplier (garments per week)
3. Reliability of supply
4. Product quality
5. Product range
6. Ethical standards
7. Supplier risk
8. Lead time for contract placement

For some of these criteria, it is relatively easy for Clara to define a numerical metric. For others, the assessment is essentially qualitative. The preferred method for creating qualitative metrics is to define a standardised scale of assessment, which is common to all the qualitative criteria, and then assign a numerical score to each of the ratings on her qualitative scale. A suitable scale of assessment for this purpose is that shown in Table 1.

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unacceptable</td>
<td>0</td>
</tr>
<tr>
<td>Very Poor</td>
<td>1</td>
</tr>
<tr>
<td>Poor</td>
<td>2</td>
</tr>
<tr>
<td>Moderate</td>
<td>3</td>
</tr>
<tr>
<td>Good</td>
<td>4</td>
</tr>
<tr>
<td>Excellent</td>
<td>5</td>
</tr>
</tbody>
</table>

*Table 1: Numerical Scoring of Qualitative Assessments*

The next step is to define the scale of measurement and, relative to this scale, the critical and compliant values for each metrics. For our example, we assume that Clara settles on the values specified in Table 2.

---

12 To simplify the exposition, we assume without comment that a similar consensus, where required, is secured for the remainder of the problem formulation.
### Table 2: Critical and compliant values for the Nightwear Supplier Decision Criteria

<table>
<thead>
<tr>
<th>Decision Criterion</th>
<th>Metric</th>
<th>Critical Value</th>
<th>Compliant Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>£/Garment</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Capacity</td>
<td>Garments/wk</td>
<td>150</td>
<td>400</td>
</tr>
<tr>
<td>Reliability</td>
<td>% of timely deliveries</td>
<td>90%</td>
<td>98%</td>
</tr>
<tr>
<td>Product Quality</td>
<td>0-5</td>
<td>2 (Poor)</td>
<td>4 (Good)</td>
</tr>
<tr>
<td>Product Range</td>
<td>0-5</td>
<td>2 (Poor)</td>
<td>4 (Good)</td>
</tr>
<tr>
<td>Ethical Standards</td>
<td>0-5</td>
<td>2 (Poor)</td>
<td>4 (Good)</td>
</tr>
<tr>
<td>Supplier Risk</td>
<td>0-5</td>
<td>2 (Poor)</td>
<td>4 (Good)</td>
</tr>
<tr>
<td>Lead Time</td>
<td>Weeks</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

5.3 Aggregation of Cost Metrics

In the example, the cost criteria are unproblematic. In general, the costing element of a public acquisition is much more complex and onerous. However, as a rule, cost tradeoffs can be resolved using well-established conventions to generate a standard cost metric such as Equivalent Annual Cost or NPV. Such conventions include the use of an agreed discount rate to trade-off immediate with future costs, and the use of actual or notional interest rates to calculate cost of capital and opportunity cost. In contrast, tradeoffs between non-cost metrics are far more problematic. There are no conventions for combining for example, Quality, Environmental Characteristics and Security of Supply into a single measure that command anything like the degree of acceptance and statutory endorsement that is given to the conventions which combine Purchase Price, Running Costs and Lifecycle Costs into an NPV.13

This paper will therefore concentrate on the issues of benefit aggregation and cost-benefit tradeoff. Specifically, it will assume that either:

a) There is only one cost metric (that is, \( m = 1 \) in the formulation of Section 4.2), or

b) The overall cost function \( C[c_1 \ldots c_m] \) can be derived analytically from the \( m \) cost metrics, without the need to elicit any information from the stakeholders with regard to the relative importance or significance of these metrics.

These two assumptions effectively amount to the same thing, since we can turn case (b) into case (a) by simply replacing the \( m \) cost metrics with the unified metric. Accordingly, from this point we will simply refer to a unified cost metric \( c \) with domain \( C \) and critical and compliant values \( C_{CRIT} \) and \( C_{COMP} \). Note that there is no real loss of generality in making this assumption, since the process, described in the next section, which is used to develop the overall benefit function can also be applied to generate an overall cost function if necessary.

---

13 The criteria used as examples here are taken from the DSPCR guidance quoted in Section 3.2
5.4 Developing Utility Functions for Benefit Criteria

The treatment of benefit metrics is very different from that of cost. Instead of deriving the overall benefit directly from the individual metrics, we proceed as follows:

a) The metrics \( x_1 \ldots x_n \) are transformed into measures called **criterion utility functions**, \( u_1[x_1] \ldots u_n[x_n] \), which are “normalised” so that, for each metric, XCRIT is assigned a utility of 0 and XCOMP a utility of 100.

b) Judgements are elicited from the stakeholder community as to the relative importance or significance of these metrics. The form of these judgements depends on whether we are following an MCDA approach or a joint utility approach.

c) The elicited judgements are used to parameterise the overall benefit function \( B[x_1 \ldots x_n] \). This will be linear if we are following an MCDA approach, and (in general) non-linear if we are following a joint utility approach.

A utility function is an **interval scale of strength of preference** (Reference 12). An interval scale is one in which the differences between two pairs of numbers can be meaningfully compared: for example, if \( u[x] \) is a utility function such that \( u[a] = 0 \), \( u[b] = 25 \), and \( u[c] = 100 \), then we can meaningfully assert both that our preference for \( c \) over \( a \) is four times greater than our preference for \( b \) over \( a \), and that our preference for \( c \) over \( b \) is three times greater than our preference for \( b \) over \( a \).

Since the strength of preference between different values of criteria such as Capacity and Ethical Standards is not obviously inherent in the criterion itself, either we must make some assumptions regarding the utility functions, and elicit approval from the stakeholders for the preferences that are implied by the proposed function; or we must build up the utility function from directly-elicited stakeholder preferences.

5.4.1 Elicitation of Utility Functions

Decision theory offers a number of approaches to the elicitation of a utility function for a decision criterion, from a single “decision maker” or a panel of stakeholders. One of the most practical is the **certainty-equivalence** or “50/50” method,\(^{14}\) which can be summarised as follows.

1. Our first step to build the curve is to elicit the value \( x \) of the MOE such that the stakeholders consider an improvement from XCRIT to \( x \) to have half the significance of an improvement from XCRIT to XCOMP. One way to do this is to try different values of \( x \) and ask whether the improvement from XCRIT to \( x \) is equally, less or more significant than a further improvement from \( x \) to XCOMP; adjusting \( x \) until the two improvements are considered equally significant. Another approach, pioneered by von Neumann and Morgenstern,\(^{15}\) is to try

---

\(^{14}\) Goodwin & Wright (Reference 13), p.116

different values of x and ask whether stakeholders would be indifferent between an assured outcome of x and an investment which had a 50/50 chance of yielding either XCRIT or XCOMP (if there is uncertainty as to a precise value of x, we derive the interval outside which the stakeholders would have a clear preference between the certainty and the gamble, and take x to be the midpoint of that interval). Since the expected utility of the gamble is 0.5*u[XCRIT]+0.5*u[XCOMP]) (50 in our example), this must also be the value of u[x]. We call this value of x X_{50}.

2. Repeat step 1 on the intervals (XCRIT, X_{50}) and (X_{50}, XCOMP) to generate the points (X_{25}, 25) and (X_{75}, 100).

3. The same method can be used to extend the curve beyond XCOMP. For example, we can find X_{150} as the value x such that an improvement from X_{50} to XCOMP has equal significance with an improvement from X_{100} to x.

4. Likewise, we can extend the curve beyond XCRIT; for example, X_{-50} is the value of x such that an improvement from x to XCRIT has equal significance with an improvement from XCRIT to X_{50}.

5. Fill out the remainder of the curve between the elicited points using a piecewise interpolation function. For this purpose, a piecewise polynomial function is recommended, since this allows us to have linear sections of curve. A cubic interpolation function can be parameterised so as to remain monotonic throughout each section of curve, while maintaining a smooth join between adjacent sections.\textsuperscript{16}

6. Add tails to the curve to extend it beyond the range of elicited utilities. In general, it is preferable that the tails should decrease without limit beyond XCRIT and be bounded above beyond XCOMP. For this purpose, exponential functions are recommended, since these can easily be parameterised to exhibit this behaviour and also to join smoothly with the elicited sections of the curve.

\textsuperscript{16} A “smooth join” here means that the derivative of the interpolated utility function is continuous across the join.
Figure 2 shows an example of the outcome of this process. The flattening out of the curve for XCOMP < x is typical, and reflects the perception that, if a compliant outcome can be assured, decision-makers will be reluctant to take risk against this outcome, or to make appreciable sacrifices against other characteristics, in the hope of achieving further improvements.

5.4.2 Default Utility Functions

The elicitation procedures described in the previous section are quite difficult to execute successfully. Authors such as Goodwin and Wright (Reference 13) point out that the von Neumann – Morgenstern formulation is primarily focussed on problems where the main challenge is uncertainty of outcome rather than multiplicity of objectives, and question its relevance to a multiple-criterion decision. The same authors observe that the outcome of a decision analysis is rarely sensitive to the precise shape of the utility curve and that “linear utility functions are extremely robust solutions”.

These considerations lead to an alternative approach to the definition of criterion utilities: the use of a standardised “default” form of the utility function, which represent a credible set of preferences between different outcomes against a criterion, subject to stakeholder review. We postulate the following as criteria for credibility in a criterion utility function u[x]:

1. u[x] is monotonic and satisfies \( u[\text{XCRIT}] = 0, \ u[\text{XCOMP}] = 100 \)
2. u[x] is a concave function of x: that is, any straight line joining two points on the curve will lie either on or below the curve. Concave utility functions represent a perception of diminishing returns: that is, the greater the level of achievement against a metric, the less sacrifice stakeholders are willing to make for the sake of further improvements.
3. u[x] does not increase rapidly for outcomes better than XCOMP: this is inherent in the assumption that XCOMP represents “a satisfactory outcome, such that stakeholder groups would be reluctant to make appreciable sacrifices against
other criteria in order to further improve an already-compliant outcome" (Section 4.2).

4. \( u[x] \) decreases increasingly rapidly for outcomes worse that \( X_{CRIT} \): this is inherent in the assumption that outcomes worse that \( X_{CRIT} \) represent "a serious threat to the viability of a solution, meaning that significant sacrifices against other criteria would be justified in order to improve a sub-critical outcome" (ibid).

Assigning a default utility function to a criterion is normally a two-stage process. The first stage is to normalise the criterion metric by defining the transformation

\[ v_n(x_n) = \frac{100*(x_n - X_{CRIT,n})}{(X_{COMP,n} - X_{CRIT,n})}. \]

Equation 13

This transformation eliminates the effect of differences in the scale of measurement, and also eliminates the difference between positive and negative metrics. The function \( v_n \) has the property that outcome \( x \) if preferred to \( y \) if and only if \( v_n(x) > v_n(y) \); in decision theory, a function with this property is termed a value function.\(^\text{17}\)

The second stage is to use \( v_n \) as the argument for a generic or "off-the-shelf" utility function. This function is defined by a small set of parameters which are independent of \( X_{CRIT} \) and \( X_{COMP} \).

An example of a generic utility function is:

\[
UG[v] = \begin{cases} 
UMAX*(1-\exp[-v/UMAX]) & \text{if } v < 0 \\
0 & \text{if } 0 \leq v < 100 \\
UMAX - (UMAX-100)*\exp[-(v-100)/UMAX-100]] & \text{otherwise}
\end{cases}
\]

Equation 14

where \( U_{MAX} \) is an upper bound which is approached asymptotically as \( v \) increases. This function embodies the above-mentioned suggestion of Goodwin and Wright that utility can be treated as linear or near-linear within a trade-off space bounded below by the critical outcomes, and above by the compliant ones.

---

\(^{17}\) Goodwin & Wright (Reference 13), p.37
Figure 3: Default utility function (linear between XCRIT and XCOMP)

Figure 3 shows the outcome of applying this function, with XMAX = 150, to a criterion with XCRIT = 2, XCOMP = 4. Between XCRIT and XCOMP, the curve is linear, but it steepens below XCRIT and flattens above XCOMP. This ensures that all the credibility criteria listed above are satisfied.

An alternative form of default utility function can be used when stakeholders are unhappy with the assumption of a distinct region in which preferences can be treated as linear. The function is specified by the formula

$$U_G(v) = U_{MAX} \times (1-(1-100/U_{MAX})^{v/100})$$

Equation 15

Figure 4: Default utility function (concave between XCRIT and XCOMP)
Figure 4 shows the outcome of applying this function, with values for UMAX of 101, 150 and 300, to the same criterion as in Figure 3. This family of utility functions does not change its response properties at XCOMP and XCRIT, and so is more suitable for criteria where stakeholders are unable to define a clear breakpoint between critical and acceptable outcomes. By varying UMAX, it is possible to represent both near-linear and strongly non-linear tradeoff preferences in the region between XCOMP and XCRIT.

5.5 Worked Example: Cost and Benefit Utility Functions

In our example, we adopt the default utility function shown in Figure 3 with UMAX = 150. Figure 5 shows the cost utility function derived from this function, while Figure 6 shows, by way of example, the utility function for the Capacity benefit. Note that these curves are identical in shape to that of Figure 3 (except for left/right reversal with a negative metric such as cost). Note also the effect of imposing an upper bound on utility: for example, the Capacity metric shows little improvement in utility for increases in capacity beyond 600. This means that a supplier who is capable of delivering in much larger quantities will not be preferred on that account to a supplier who is capable of delivering 400-500 garments per week and is markedly superior in other respects.

![Figure 5: Cost Utility Function for the Nightwear Supplier Decision](image1)

![Figure 6: Capacity Utility Function for the Nightwear Supplier Decision](image2)
5.6 Aggregating Benefit Criterion Utilities

5.6.1 Weighted Sum Approach

The weighted sum approach to generating an aggregate benefit score is a simple adaption of the MCDA methodology. The main difference between the approach summarised below and the classic MCDA formulation is that, in MCDA, the actual options are used to define the scoring scale for the options against each criterion.\(^{18}\) In our formulation, the scoring scales are defined in advance of the options being identified; we therefore define the scoring scales relative to the values XCRIT and XCOMP.

The process can be summarised as follows:

1. Starting with all metrics at their critical values, decide which of the criteria creates the greatest increase in benefit when \(x_j\) is changed from its critical to its compliant value, the other metric remaining unchanged. For the sake of example, assume this to be criterion 1.

2. Assign criterion 1 an initial weighting of 100.

3. For each of the other criteria, starting with all other metrics at their critical values, assess the significance of changing \(x_j\) from its critical to its compliant value, leaving the other metrics unchanged, relative to that of changing \(x_1\) to its compliant value. This assessment is typically expressed in percentage terms: for example, “increasing \(x_2\) from XCRIT\(_2\) to XCOMP\(_2\) has 67% of the significance of increasing \(x_1\) from XCRIT\(_1\) to XCOMP\(_1\)”.

4. Assign each criterion an initial weighting based on this assessment. In this example, the initial weighting assigned to criterion 2 will be 67.

5. Normalise the weighting so that they sum to 1. In the example, if there were no other criteria to consider, we would have \(w_1 = 100/(100+67) = 0.6\) and \(w_2 = 67/(100+67) = 0.4\).

6. Create a set of scoring functions, one for each benefit metric, which are **standardised** so as to meet the following requirements:
   - Each criterion is assessed on the same scale (for example, 0-100 or 1 to 10);
   - For each metric, more favourable values are assigned higher scores
   - The scores assigned to XCRIT\(_j\) and XCOMP\(_j\) are the same for every criterion \(j\).

---

\(^{18}\) Goodwin & Wright [Reference 13], p41
Let $a_1$ to $a_N$ be the options under consideration, and, for $d = 1$ to $N$, let $u_j[a_d]$ be the criterion utility score of option $a_d$ against criterion $j$. Then the aggregate benefit score for option $d$ is

$$BS[a_m] = w_1^*u_1[a_d] + \ldots + w_n^*u_n[a_d],$$

*Equation 16*

where $w_1$ … $w_n$ are the normalised weights summing to 1.

### 5.6.2 Multivariate joint utility

In the joint utility approach, we derive an analytic function which directly generates a joint utility score for any feasible combination of criterion utilities.

To generate this function we must, as in the weighted sum approach, elicit judgements from the stakeholders as to the relative significance of each criterion. These judgements may be in either of two forms, as follows.

We first define the **critical option** to be the hypothetical option such that $x_j = \text{XCRIT}_j$ for all criteria; and the **compliant option** to be the hypothetical option such that $x_j = \text{XCOMP}_j$ for all criteria. Then:

1. Taking the critical option as a starting point, the **impact** of criterion $j$, $M_j$, is the significance of a “swing” in criterion $j$ from XCRIT up to XCOMP, given that a swing from XCRIT to XCOMP for all criteria is credited with a significance of 100.

2. Taking the compliant option as a starting point, the **criticality** of criterion $j$, $C_j$, is the significance of a “swing” in criterion $j$ from XCOMP down to XCRIT, given that a swing from XCOMP to XCRIT for all criteria is credited with a significance of 100.

We define the **impact reference option** for criterion $j$, $\text{XM}_j$, to be the hypothetical option with benefit scores ($\text{XCRIT}_1$, … $\text{XCRIT}_{j-1}$, $\text{XCOMP}_j$, $\text{XCRIT}_{j+1}$, … $\text{XCRIT}_n$); and the **criticality reference option** for criterion $j$, $\text{XC}_j$, to be the hypothetical option with benefit scores ($\text{XCOMP}_1$, … $\text{XCOMP}_{j-1}$, $\text{XCRIT}_j$, $\text{XCOMP}_{j+1}$, … $\text{XCOMP}_n$). Then the joint utility of $\text{XM}_j$ must be $M_j$, and the joint utility of $\text{XC}_j$ must be $C_j$.

We can now define the **joint impact utility function** and **joint criticality utility function** as follows.

**Joint impact utility function**

If $M_1 + \ldots + M_n = 100$ then the joint utility function is linear:

$$UJI[x] = M_1^*u_1[x_1] + \ldots + M_n^*u_n[x_n].$$

*Equation 17*

Otherwise, there exists a unique non-zero real number $k$ such that

$$1+k = \prod_{j=1 \text{ to } n} [1+k* M_j /100].$$

*Equation 18*

This $k$ can be determined by iteration, or, in Excel, by the Solver or Goal Seek utilities.

We then define, for $j = 1$ to $n$,

$$UIMIN_j = \begin{cases} -9999 & k \leq 0 \\ -10000/k*M_j & \text{otherwise} \end{cases};$$

*Equation 19*

$$UIMAX_j = \begin{cases} -10000/k*M_j & k < 0 \\ -9999 & \text{otherwise} \end{cases}.$$
and the truncated criterion utility function
\[ UTI_j[x_j] = \max\{UIMIN_j, \min\{UIMAX_j, u_j[x_n]\}\}. \]  
\( Equation \ 20 \)

Then the joint utility function is
\[ UJI[x_1 \ldots x_n] = (100/k) \times \left( \prod_{j=1}^{n} \left[ 1 + k \times M_n \times UTI_j[x_j] / 10000 \right] - 1 \right). \]  
\( Equation \ 21 \)

Joint criticality utility function

If \( C_1 + \ldots + C_n = 100 *(n-1) \) then the joint utility function is linear:
\[ UJ[x_1 \ldots x_n] = (1-C_1/100)*U_1[x_1] + \ldots + (1-C_n/100)*U_n[x_n]. \]  
\( Equation \ 22 \)

Otherwise, there exists a unique non-zero real number \( h \) such that
\[ 1 + h = \Pi_{j=1}^{n} [1 + h \times C_n / 100]. \]  
\( Equation \ 23 \)

We then define, for \( j = 1 \) to \( n \),
\[ UCMIN_j = \begin{cases} 100 *(1 + 100/(h \times C_i)) & h \leq 0 \\ -9999 & \text{otherwise} \end{cases} \]  
\( Equation \ 24 \)

\[ UCMAX_j = \begin{cases} 9999 & h < 0 \\ 100 *(1 + 100/(h \times C_i)) & \text{otherwise} \end{cases} \]  
\( Equation \ 25 \)

and the truncated criterion utility function
\[ UTC_j[x_j] = \max\{UCMIN_j, \min\{UCMAX_j, u_j[x_n]\}\}. \]  
\( Equation \ 26 \)

Then the joint utility function is
\[ UJC[x_1 \ldots x_n] = (100/h) \times \left( 1 + h - \Pi_{j=1}^{n} \left[ 1 + h \times C_n / (1 + UTC_j[x_j]/100) \right] \right). \]  
\( Equation \ 27 \)

It is shown in Annex A that:

- \( UJI[XCRIT_1 \ldots XCRIT_n] = UJc[XCRIT_1 \ldots XCRIT_n] = 0 \) \( Equation \ 28 \)
- \( UJI[XCOMP_1 \ldots XCOMP_n] = UJc[XCOMP_1 \ldots XCOMP_n] = 100 \) \( Equation \ 29 \)
- For all \( j \), \( UJI[XM_j] = M_j \) \( Equation \ 30 \)
- For all \( j \), \( UJC[XC_j] = C_j \) \( Equation \ 31 \)
- \( UJI \) and \( UJC \) are monotonic in all \( u_j \) for all feasible combinations of \( u_j \), and strictly monotonic whenever all \( u_j \) are between 0 and 100.

### 5.7 Elicitation of the Overall Benefit Function

In order to use either of the benefit aggregation methods described in the previous section, it is necessary to elicit the parameters which are used in their respective aggregation formulas: criterion weightings for the weighted-sum method, and impacts or criticalities for the joint utility method. In our example we will pursue both methods, so that their respective merits can be compared.

---

\textsuperscript{19} This truncation, and the similar truncation in the joint criticality utility function, are necessary to ensure that \( UJ \) is monotonic throughout the tradeoff space. See Annex A
5.7.1 Weighted-sum Method: Elicitation of Criterion Weights

Elicitation of criterion weights is normally conducted in a workshop at which the key stakeholders are represented. Taking the critical option as a starting point, we seek a consensus on the significance of a "swing" from the critical to the compliant value for each criterion. The procedure given below is quoted from the official Department for Communities and Local Government handbook (Reference 8):

"First, the ... criterion with the biggest swing [between the critical and compliant outcomes] is identified. If the MCDA model includes only a few criteria, then the biggest swing can usually be found quickly with agreement from participants. With many criteria, it may be necessary to use a paired-comparison process: compare criteria two at a time for their preference swings, always retaining the one with the bigger swing to be compared to a new criterion. The one criterion emerging from this process as showing the largest swing in preference is assigned a weight of 100; it becomes the standard to which all the others are compared in a four-step process.

- First, any other criterion is chosen and all participants are asked to write down, without discussion, a weight that reflects their judgement of its swing in preference compared to the standard. If the criterion is judged to represent half the swing in value as the standard, for example, then it should be assigned a weight of 50.

- Second, participants reveal their judged weights to the group (by a show of hands, for example, against ranges of weights: 100, 90s, 80s, 70s, etc.) and the results are recorded on a flip chart as a frequency distribution.

- Third, participants who gave extreme weights, high and low, are asked to explain their reasons, and a general group discussion follows.

- Fourth, having heard the discussion, a subset of participants makes the final determination of the weight for the criterion.

... If there is not a consensus, then it might be best to take two or more sets of weights forward in parallel, for agreement on choice of options can sometimes be agreed even without agreement on weights. Even if this does not lead easily to agreement, explicit awareness of the different weight sets and their consequences can facilitate the further search for acceptable compromise."

In our example, if Product Quality emerges as the most important criterion, we assign a weight to Capacity (say) by asking: "If improving Product Quality from Poor to Good is worth 100, how much is it worth to improve Capacity from 150 garments a week to 400?" These comparisons require the stakeholders to evaluate a strength of preference between two outcomes rather than a simple statement as to which is preferred. It is implicit in this guidance that stakeholders are able to answer questions of this form, and the handbook cites a number of case studies to justify this.

---

20 It is implicit throughout the process description that either consensus is achieved on all the key stakeholder inputs, or that divergent opinions can be provided for by analysis of the sensitivity of the outputs to the spread of opinions expressed.

21 For a textbook description, see Goodwin & Wright (Reference 13).
Once the weights relative to the most important criterion have been elicited, all that remains is to normalise them by dividing each weight by the sum of all weights. In our example, we assume the following outcome (Table 3):

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Metric</th>
<th>Cost per garment</th>
<th>Metric (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(£) (Garments/wk)</td>
<td>(% of timely deliveries)</td>
</tr>
<tr>
<td>Metric ranking</td>
<td>2</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Metric initial weighting</td>
<td>70</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>Metric normalised weighting</td>
<td>18.5%</td>
<td>5.4%</td>
<td>27.0%</td>
</tr>
</tbody>
</table>

**Table 3: Criterion weights for the Nightwear Supplier Decision**

### 5.7.2 Joint Utility Method: Elicitation of Impacts or Criticalities

#### Direct Elicitation

In order to develop a joint utility function, it is first necessary to decide whether to use the joint criticality utility function, in which case we have to elicit the criticality of each criterion; or the joint impact utility function, in which case we have to elicit the impact of each criterion.

To do this, we first ask the question:

- *Are any criteria likely to have a criticality greater than 50?* That is, taking the critical option (with $x_j = X_{CRIT}$ for all criteria) as a starting point, are there any criterion such that a swing from $X_{CRIT}$ to $X_{COMP}$ for that criterion alone is likely to have more than 50% of the significance of a swing from $X_{CRIT}$ to $X_{COMP}$ for all criteria.

Three cases arise:

1. Two or more criteria have a criticality greater than 50. In this case we elicit the criticality of every other criterion, calculate the parameter $h$, as defined in Section 5.6.2, and so derive the joint criticality utility function.

2. One criterion has a criticality greater than 50. In this case, we elicit the impact of this criterion. If this is greater than the criticality of the same criterion, then we elicit the impact of every other criterion, calculate the parameter $k$, as defined in Section 5.6.2, and so derive the joint impact utility function. Otherwise, we derive the joint criticality utility function as in 1.

3. No criterion has a criticality greater than 50. In this case, we look for the criterion with the highest impact. If the highest elicited impact exceeds the highest elicited criticality, we elicit the impact of every criterion, calculate $k$, and derive the joint impact utility function. Otherwise, we derive the joint criticality utility function as in 1.
The rationale for this process is that if $M_i$ is the impact of criterion $i$, $C_j$ is the criticality of criterion $j$, and $i \neq j$, then:

$$M_i + C_j \leq 100. \quad (Equation \ 33)$$

It follows that:

- If any criterion $j$ has a criticality of $N$, no criterion other than $j$ can have an impact of more than $100-N$.
- If any criterion $j$ has an impact of $N$, no criterion other than $j$ can have a criticality of more than $100-N$.
- If two or more criteria have a criticality of $N$ or more, no criterion can have an impact of more than $100-N$.
- If two or more criteria have an impact of $N$ or more, no criterion can have a criticality of more than $100-N$.

In general, the very characterisation of a problem as “multiple criterion” militates against it having high-impact criteria. If a criterion has impact near 100, this means that a near-compliant solution ($UJ \approx 100$) can be obtained simply by achieving a compliant outcome against that criterion ($u_j[x_j]=100$), even if the option has critical shortfalls against all other criteria; hence this is the only criterion which actually has to be satisfied.

For this reason, the default assumption for the joint utility approach is that the joint criticality utility function will be used in preference to the joint impact function. While high-impact criteria are unusual in genuinely multiple-criterion problems, it is quite typical of such problems that they should have one or more high-criticality criteria, all of which must be at least partially satisfied in order to avoid a sub-critical outcome. A simple example of such a problem would be the acquisition of an armoured vehicle with decision criteria of (inter alia) tactical mobility, combat effectiveness and reliability. It is reasonable to expect all these criteria to have relatively low impact but high criticality: for example, it is largely irrelevant whether a vehicle is reliable and/or combat-effective if it cannot reach the places where the decisive combats are being fought in time to have an impact.

It is important to note that there is no obstacle to all criteria being high-criticality. In fact it can be shown that, if $C_1 = C_2 = \ldots = C_n = 100$, then the joint criticality utility function simplifies to the product of the criterion utility functions:

$$UJC[x_1 \ldots x_n] = 100 \times \Pi_{j=1}^{n} \frac{UTC_j[x_j]}{100}. \quad (Equation \ 34)$$

On the other hand, if $C_1 + C_2 + \ldots + C_n = 100*(n-1)$, the joint criticality utility function is a linear sum of the criterion scores. These two cases illustrate that the joint criticality utility function automatically adapts to the criticality inputs which are fed into it, so as to provide the appropriate response properties across the whole of the

---

22 This is most easily shown by an example. Suppose we have 3 criteria. Then we have $M_1 = UJ[XCOMP_1, XCRIT_2, XCRIT_3]$, and $C_2 = 100 - UJ[XCOMP_1, XCOMP_2, XCRIT_3]$. Since $UJ$ is monotonic in the criterion utilities, we must have $UJ[XCOMP_1, XCRIT_2, XCRIT_3] \leq UJ[XCOMP_1, XCOMP_2, XCRIT_3]$; substituting yields $M_1 \leq 100 - C_3$, so $M_1 + C_3 \leq 100$, as required.

23 The parameter $h$ in Equation 28 tends to infinity as all criticalities converge to 100. Hence Equation 34 is, formally, the limiting case of Equation 28 for $C_1 = C_2 = \ldots = C_n = 100$. 

QinetiQ Proprietary
tradeoff space. This can be regarded as a major advantage of the joint utility approach, as compared to the weighted-sum approach.

**Elicitation of Criticalities from the Cost-Benefit Function ("Buy-Back")**

In order to elicit criticalities directly, we need stakeholders to be able to answer questions like:

"If option A scores 100 against all benefits, option B scores 0 against Product Quality and 100 against all other benefits, and option C scores 0 against all benefits, how significant is the swing from A to B, given that the swing from A to C has a significance of 100?"

Such questions are not easy for stakeholders to answer directly: because the endpoints of the two “swings”, B and C, are so very different in their implications, it is difficult for the stakeholders to envisage a common scale against which they can be compared. However, if we have already generated the tradeoff function for overall benefit against cost, then we can elicit the criticality of the benefit criteria indirectly by asking how much stakeholders would be willing to pay in order to “buy back” a swing from a compliant to a critical outcome for each criterion. This “buy-back” cost can then serve as a basis for comparison between these outcomes.

The questions required by the buy-back approach are of the following form:

"If option A cost £6 per garment and scores 0 against Product Quality and 100 against all other benefits, and option B scores 100 against all benefits, what would the price of option B have to be in order for you to be indifferent between A and B?"

Suppose that the answer is 8. Let:

- \( x_{100} \) be the outcome (XCOMP\(_1\), … XCOMP\(_n\));
- \( xc \) be the outcome obtained by replacing XCOMP\(_j\) with XCRIT\(_j\)in \( x_{100} \);
- \( UB[x] \) be the joint benefit function (currently unspecified) for the benefit outcome defined by the vector \( x = (x_1 \ldots x_j) \);
- \( UCB[c, b] \) be the joint cost-benefit function for an option with cost c and overall benefit score b.

Then the cost-benefit outcome for option A must have the same cost-benefit score as option B at £8 per garment: that is,

\[
UCB[6 \ , \ UB[xc]] = UCB[8 \ , \ UB[x_{100}]] \quad \text{Equation 34}
\]

At this stage, the function \( UB[x] \) has not yet been specified. However, we have already stipulated that \( UB[x_{100}] = 100 \); and that, by definition, \( UB[xc] = 100 - C_j \). Hence this equation simplifies to

\[
UCB[6 \ , \ 100 - C_j] = UCB[8 \ , \ 100]. \quad \text{Equation 35}
\]

Hence, if we have already determined the function \( UCB[c, b] \), we can solve this equation to find \( C_j \).

In general, we define the buy-back cost for benefit criterion \( j \), \( CBB_j \), to be the cost increase relative to CCOMP such that stakeholders are indifferent between the cost-benefit outcomes (CCOMP, \( xc \)) and (CCOMP + \( CBB_j \, x_{100} \)). Given \( CBB_j \), the criticality of criterion \( j \) can be determined by solving the equation:
Example Values

In our example, we assume the outcomes shown in Table 4. An example of how these values could be elicited by "buy-back" is given in Section 5.8.6 below.

Note that there is no question of normalising the elicited criticalities. In the example, the criticalities sum to 220, which means that we would expect significant non-linearity in the response properties of the joint utility function.

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Cost per garment (£)</th>
<th>Capacity (Garments/ wk)</th>
<th>Reliability (% of timely deliveries)</th>
<th>Product Quality (0-5)</th>
<th>Product Range (0-5)</th>
<th>Ethical standards (0-5)</th>
<th>Supplier risk (0-5)</th>
<th>Lead time (weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical outcome</td>
<td>9</td>
<td>150</td>
<td>90%</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Compliant outcome</td>
<td>6</td>
<td>400</td>
<td>98%</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Criticality</td>
<td>40</td>
<td>10</td>
<td>60</td>
<td>30</td>
<td>35</td>
<td>25</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Criterion criticalities for the Nightwear Supplier Decision

5.8 Cost-Benefit Tradeoffs

5.8.1 Indifference Curves and Utility Contours

We articulate the criteria for evaluating cost-benefit tradeoffs by an "indifference map". A simple example of such a map is illustrated in Figure 7. This figure shows cost and benefit plotted on the x and y axes respectively. The black curves are "indifference curves", which have the following properties:

1. If two cost-benefit tradeoffs lie on the same indifference curve, they are equally favourable or "indifferent". 24
2. All tradeoffs that lie above or to the left of an indifference curve are preferred to all tradeoffs on or below the curve;
3. All tradeoffs that lie below or to the right of an indifference curve are dispreferred to all tradeoffs on or above the curve.

In this example, four utility curves are plotted, and the regions between the curves are colour-coded from red to green to indicate increasingly favourable cost-benefit combinations. It can be seen that tradeoffs B and D are preferred to tradeoffs A, C and E, since B and D are above an indifference curve which A, C and E are below. It can also be inferred, from the shape of the four curves plotted, that the indifference curve through B would pass below D and that D is therefore the most favourable of the five options in cost-benefit terms. Note that B and D are preferred to both the cheapest option (C) and the option delivering the highest level of benefit (A).

---

24 In decision theory, two options are said to be indifferent if, according to the decision criteria in use, neither of them can be shown to be preferable to the other.
The publication of cost-benefit indifference curves such as Figure 7, in advance of a COEIA assessment, would be a major advance in transparency for the acquisition process (see Section 6.3 below). To achieve this, we derive an objective function, as specified in Section 4.2, which takes the overall cost and benefit scores as arguments. With this formulation, the indifference curves become “utility contours”: that is, lines joining all points in the cost-benefit tradeoff space which have the same joint utility, and the indifference map is expressed as a “heat map” in which each point in the tradeoff space is colour-coded according to its value.

5.8.2 The generic form of the Cost-Benefit Function

The cost-benefit objective function is a special case of the joint utility function derived in Section 5.6.2. We take cost as criterion 1 and overall benefit as criterion 2. For n = 2 the joint impact and joint criticality functions for the cost and benefit criteria reduce to the formula:

$$UCB(c, b) = \frac{(M_c \cdot UC + M_b \cdot UB)}{100} + \frac{(100 - M_c - M_b) \cdot UC \cdot UB}{10000}$$

$$= \frac{(1 - C_b/100) \cdot UC + (1 - C_c/100) \cdot UB + (C_c + C_b - 100) \cdot UC \cdot UB}{10000}$$

Equation 37

where:
- UC = overall cost utility score
- UB = overall benefit utility score
- $M_c$ = Impact of cost criterion
- $M_b$ = Impact of overall benefit criterion
- $C_c$ = Criticality of cost criterion
- $C_b$ = Criticality of overall benefit criterion

To populate this formula, we must elicit from stakeholders either the two impacts or the two criticalities. Methods for doing this are discussed in Section 5.8.4 below.
5.8.3 **Representation of the Cost-Benefit Function by a Heat Map**

Figure 8 shows the output of Equation 37 with $\text{CCRIT}=9$, $\text{CCOMP}=6$, $C_c=70$, $C_b=80$, expressed as a “heat map”. In this form of presentation, a two-dimensional array is created in which the horizontal and vertical directions correspond to increasing values of the cost and benefit metrics respectively. Each cell represents the intersection of the benefit score associated with a particular row and the cost associated with a particular column, and the value taken by the cell is the joint utility of the row benefit score and the column cost score. Each cell is colour-coded according to its value, creating a set of regions which represent different scoring bands for joint utility. As in Figure 7, the boundaries between different-coloured regions are utility contours; for example, the boundary between the “$U > 100$” region and the “$70 < U < 100$” is the $u = 100$ contour. The four small grey squares mark the reference points $(\text{CCOMP}, 0)$, $(\text{CCOMP}, 100)$, $(\text{CCRIT}, 0)$ and $(\text{CCRIT}, 100)$ that form the vertices of the core tradeoff space. Note that $u[\text{CCRIT}, 0] = 0$ and $u[\text{CCOMP}, 100] = 100$; this will always be the case when this formulation is used.

![Heat Map Diagram](image)

**Figure 8: Representation of the cost-benefit joint utility function as a Heat Map**

5.8.4 **Elicitation of the Criticality Parameters for the Cost-Benefit Function**

We noted in Section 5.7.2 the difficulty of eliciting the criticalities of the benefit criteria directly from the stakeholders. Similar difficulties arise with the elicitation of $C_c$ and $C_b$. For example, to elicit $C_c$ using the method outlined in Section 5.7.2, we must elicit the significance of a “swing” in overall benefit from 100 down to 0, given that a swing from $(\text{CCOMP}, 100)$ to $(\text{CCRIT}, 0)$ is credited with a significance of 100.

In our example, we would therefore pose the question:
"If option A costs £6 per garment and scores 100 against all benefits, option B cost £9 per garment and scores 100 against all benefits, and option C cost £9 per garment and scores 0 against all benefits, how significant is the swing from A to B, given that the swing from A to C has a significance of 100?"

As noted in Section 5.7.2, endpoints of the two “swings”, B and C, are very different in their implications, which makes it is difficult for stakeholders to envisage a common scale against which they can be compared.

For this reason, it may be preferable to conduct the above process in reverse: instead of eliciting $C_c$ and $C_b$ directly and simply accepting the utility contours that emerge, we elicit from stakeholders what they would expect the utility contours to look like, and then iterate over different values of $C_c$ and $C_b$ until we find a combination that produces the desired map; or at least the best possible approximation to it.

Effectively the decisions to be made are:

1. How rectilinear we want the utility contours to be: specifically, the $UCB=0$ contour. This determines the value of $C_c + C_b$; when the sum equals 100, the $UCB=0$ contour is generally diagonal, while when the sum is close to 200, the $UCB=0$ contour becomes rectilinear (Figure 9; see also Figure 8, which is the same map with $C_c = 70$, $C_b = 80$). The more rectilinear the contours, the more strongly “balanced” outcomes (with roughly equal cost and benefit utility scores) are favoured relative to “cheap and nasty” and “good but pricey” options.

2. Subject to (1), the relative significance of cost and benefit. This determines the slope of the utility contours: the more critical cost is, the steeper the slope.

![Figure 9: Cost-benefit heat map for $C_c=45$, $C_b = 55$ (left) and $C_c = C_b = 99$ (right)](image)

The advantage of this procedure is that it does not require stakeholders to quantify their preferences between different cost/benefit tradeoffs: it is sufficient for them to be able to state a preference – or lack of one – in a small number of cases. For example, suppose stakeholders can state that they are indifferent between the outcomes $(4, 5)$ and $(8, 50)$ – that is, they have no clear preference between these two outcomes. Figure 9 indicates that, for both $C_c = 45$, $C_b = 55$ and...
C_c = C_b = 99, the point (8, 50) is much closer than (4, 5) to the UBC = 30 contour (the boundary between the yellow and orange regions). However, Figure 8 shows the points (8, 50) and (4, 5) are very close to the UBC = 30 contour. Hence this elicited judgement alone is sufficient to indicate that the values of C_c and C_b which most closely reflect stakeholder preferences are likely to have a sum close to 150. The elicitation of a small number of similar cases of indifference between outcomes will enable the contours to be fitted at closely as possible to the elicited preferences. 25

5.8.5 Troubleshooting the Cost-Benefit Function

A further problem will arise if it is impossible to find any values of C_c and C_b which give a reasonable approximation to the elicited preferences across the tradeoff space. In this case we first review the elicited preferences to validate their internal coherence and logical consistency. If this review reaffirms the validity of these preferences, the next step is to consider changing the cost utility function. For this purpose a good starting point is to use the alternative form of default function derived in Section 5.4.2 above:

$$UC[c] = \text{UMAX} \times (1 - (1 - (100/\text{UMAX}))^{v[c]/100}),$$

where

$$v[c] = 100 \times (c - \text{CCRIT})/(\text{CCOMP} - \text{CCRIT}).$$

Equation 38

The concavity of this function can be changed by varying UMAX; this will provide a second degree of freedom to fit the elicited preferences. We can also experiment with varying CCRIT and CCOMP. If we can get a good fit with this function in the core tradeoff space, we can then look at modifying the tails of the cost utility function to improve the fit outside this region.

In addition, we can consider modifying the benefit function. The benefit input to the cost-benefit function is already a utility; either the weighted sum derived in Section 5.6.1 or the joint utility derived in Section 5.6.2. Hence the standard practice is to use this input unmodified. However, it is open to us to transform it before inputting it into the cost-benefit function. For example, we can introduce more concavity by using the transformation of Equation 38, in order to improve the fit in the core tradeoff region; if this works we can also modify the tails to improve the fit outside this region.

At this stage in the development of the process, it is impossible to say whether there are likely to be a significant number of cases in which these measures will be insufficient to generate a cost-benefit heat map which is acceptably close to the validated stakeholder preferences. If this is the case, it will generally indicate there is a strong degree of utility dependence between the cost and benefit utilities. Utility dependence means that the strength of preference between different cost outcomes, as expressed by the shape of the cost utility function, depends on the level of benefit being delivered; or vice-versa. Preference dependence would be manifest if, for example, the cost utility curve was strongly concave when the benefit score is high and linear when the benefit score is low. 26

---

25 In this context, we take the term “preferences” to include assertions of indifference (i.e. of “no preference”)

26 Keeney and Raiffa, p 224
All the weighted sum and joint utility functions derived above assume utility independence.\textsuperscript{27} Note that this does not mean that there is no interaction between the metrics. On the contrary, the joint utility function causes the metrics to interact very powerfully: for example, in Figure 8 we see that an increase in cost per garment from £6 to £9 reduces overall utility from 100 to 30 when the benefit score is 100, but only from 20 to 0 when the benefit score is 0. The point is the assumption that the shapes of the utility functions do not change.

There is substantial reason to assume that utility dependence is unlikely to be a serious obstacle to the generation of either joint benefit utility functions or of the joint cost-benefit function. The argument can be summarised as follows:

1. As already noted (Section 5.4.2), some theorists doubt that the exact shape of the criterion utility functions is likely to have a major impact on the outcome of a decision analysis, and that there will therefore be little loss of fidelity if practical applications of the method use default rather than elicited utilities.

2. It is doubtful whether it will be possible to elicit criterion utilities, criticalities and tradeoff preferences from the stakeholders to a defence acquisition with the level of precision which would be necessary for utility dependence to be identifiable. Uncertainties arising from this issue are likely to be insignificant compared to those arising from the spread, imprecision and volatility of stakeholder preference.

3. High precision in discriminating between options is only essential in that region of the tradeoff space in which the credible contenders are expected to lie. Hence it will usually suffice to define each criterion utility function on the assumption that the other criteria lie within this region. While this may result in a progressive loss in discrimination as options decline in credibility, this will impact neither our ability to distinguish between credible and non-credible options nor our ability to rank the credible options. If the purchasing authority in a competitive tender believes that there is a significant risk that no credible contenders will be submitted, it can protect itself by retaining the option to re-run the competition if none of the submissions lie within the region for which the cost-benefit evaluation function has been optimised.

4. The small number of acquisitions for which none of considerations 1-3 apply can be treated as special cases. For example, if we used $UC[c] = UMAX*(1-(1-100/UMAX)\times c/100^2)$ as the cost utility function, we could introduce utility dependence by making UMAX a function of the benefit score, which would allow us to vary the convexity of the function as the benefit score changes. However, it is envisaged that such solutions would be developed on a case-by-case basis.

For these reasons, we do not propose at present to develop a general treatment of utility dependence in the context of public acquisition decisions.

\textbf{5.8.6 Elicitation of Benefit Criterion Criticality by “Buy-Back”}

In Section 5.7.2 we developed the idea of using “buy back” to elicit the criticalities of the benefit criteria. It was shown that, if we have already defined the cost benefit function, we can define the buy-back cost for criterion $j$, $CBB_j$, such that

\textsuperscript{27} This can be verified by observing that, if the values of every metric except $j$ are held at fixed values, then the weighted sum formula and the joint utility formulas all become linear transformations of the utility function for metric $j$.\hfill\vspace{1cm}
stakeholders are indifferent between the cost-benefit outcomes (CCOMP, $x_c$) and (CCOMP + CBB, $x_{100}$), and then determine the criticality of criterion $j$ by solving the equation:

$$UCB[CCOMP, 100 − C_j] = UCB[CCOMP + CBB, 100]. \quad \text{Equation 39}$$

For this approach to work, we must be willing to define the cost-benefit function without a full knowledge of the joint benefit function. This is feasible, since we have shown in Section 5.8.4 that, in order to define the cost-benefit function, we only need to know the joint benefit scores for the critical and compliant options, which are always 0 and 100.

For example, in Section 5.7.2 we supposed that £8 was the answer to the question:

“If option A cost £6 per garment and scores 0 against Product Quality and 100 against all other benefits, and option B scores 100 against all benefits, what would the price of option B have to be in order for us to be indifferent between A and B?”

In other words, stakeholders would be willing to pay an extra £2 per garment in order to “buy back” a compliant outcome for Product Quality. Then it was shown that the criticality $C_j$ of Product Quality must satisfy:

$$UCB[6, 100 − C_j] = UCB[8, 100]. \quad \text{Equation 40}$$

The cost-benefit function shown in Figure 8 has CCRIT = 9, CCOMP=6, $C_c = 70$, $C_b = 80$. Substituting in the generic formula yields

$$UCB[c, b] = 0.2*UC + 0.3*UB + 0.005*UB*UC. \quad \text{Equation 41}$$

So $C_j$ satisfies

$$0.2*UC[6] + 0.3*UB[100 − C_j] + 0.005* UB[100 − C_j] *UC[6] = 0.2*UC[8] + 0.3*UB[100] + 0.005* UB[100] *UC[8]. \quad \text{Equation 42}$$

We have $UC[6] = 100$, $UC[8] = 33.3$ and $UB[100] = 100$. Solving yields $UB[100 − C_j] = 41.7$.

Since joint benefit is already expressed as a utility, we have $UB[b] = b$; so $C_j = 58.3$. Given the imprecision of the inputs, we would normally round this off to 58 or 60.
5.9 Option Evaluation

Table 5 summarises the benefit metrics, the critical and compliant values for each metric, and the criterion weights and criticalities elicited in Sections 5.7.1 and 5.7.2 (Tables 3 and 4) respectively.

<table>
<thead>
<tr>
<th>Decision Criterion</th>
<th>Metric</th>
<th>Values for Critical Option</th>
<th>Values for Compliant Option</th>
<th>Criterion Weighting</th>
<th>Criticality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity</td>
<td>Garments/wk</td>
<td>150</td>
<td>400</td>
<td>18.9%</td>
<td>40</td>
</tr>
<tr>
<td>Reliability of Supply</td>
<td>% of timely deliveries</td>
<td>90%</td>
<td>98%</td>
<td>5.4%</td>
<td>10</td>
</tr>
<tr>
<td>Product Quality</td>
<td>0-5</td>
<td>2 (Poor)</td>
<td>4 (Good)</td>
<td>27%</td>
<td>60</td>
</tr>
<tr>
<td>Product Range</td>
<td>0-5</td>
<td>2 (Poor)</td>
<td>4 (Good)</td>
<td>13.5%</td>
<td>30</td>
</tr>
<tr>
<td>Ethical Standards</td>
<td>0-5</td>
<td>2 (Poor)</td>
<td>4 (Good)</td>
<td>16.2%</td>
<td>35</td>
</tr>
<tr>
<td>Supplier Risk</td>
<td>0-5</td>
<td>2 (Poor)</td>
<td>4 (Good)</td>
<td>10.8%</td>
<td>25</td>
</tr>
<tr>
<td>Lead Time</td>
<td>Weeks</td>
<td>5</td>
<td>1</td>
<td>8.1%</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 5: Criterion Weights and Criticalities for the Nightwear Supplier Decision Criteria

Table 6 lists the options submitted for consideration for the garment supply. The table shows the decision criteria from Table 2, with the assessed level of delivery of each option against each criterion. This information, along with the elicited data in Table 5, the default criterion utility function shown in Figure 3, and the elicited cost-capability criticalities, are sufficient to evaluate these options in both overall benefit and cost-benefit terms.

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Cost per garment (£)</th>
<th>Capacity (Garments/wk)</th>
<th>Reliability (%)</th>
<th>Product Quality (0-5)</th>
<th>Product Range (0-5)</th>
<th>Ethical Standards (0-5)</th>
<th>Supplier Risk (0-5)</th>
<th>Lead time (weeks)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Nites</td>
<td>7</td>
<td>200</td>
<td>88%</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>New UK firm with &quot;hi-tech&quot; production techniques.</td>
</tr>
<tr>
<td>Red Admiral</td>
<td>8</td>
<td>800</td>
<td>96%</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>International consortium with multinational supplier base. Supplies major supermarket chain. Secretive.</td>
</tr>
<tr>
<td>Greenware</td>
<td>10</td>
<td>120</td>
<td>91%</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>UK-based, founded 5 years ago. Dilled as eco-friendly.</td>
</tr>
<tr>
<td>Purple Dragon</td>
<td>6</td>
<td>10000</td>
<td>99%</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>SE-Asia based. High volume, large export business.</td>
</tr>
<tr>
<td>Blue Horizon</td>
<td>9</td>
<td>500</td>
<td>99%</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>Long-established local firm. Recently downsized.</td>
</tr>
</tbody>
</table>

Table 6: Options for the Nightwear Supplier Decision
5.9.1 Evaluation of overall benefit for each option

Table 7 shows the normalised benefit scores for these options, obtained from the formula $v_j(x_j) = (x_j - \text{XCRIT}_j)/(\text{XCOMP}_j - \text{XCRIT}_j)$ (see Section 5.4.2 above). Normalised scores of 100 or more are colour-coded green, and normalised scores of 0 or less are colour-coded red. It will be seen that most options over-achieve against some criteria and are non-compliant in others. However, any benefit function based on normalised scores that gives an appreciable weight to capacity will clearly be dominated by Purple Dragon, since it has an overwhelming normalised score of 3940 against this criterion, which will completely eclipse its serious failings against Product Range and Ethical Standards.

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Cost per garment</th>
<th>Utility (0 = critical, 100 = compliant)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Capacity</td>
<td>Reliability</td>
</tr>
<tr>
<td>White Nites</td>
<td>20.0</td>
<td>-25.0</td>
</tr>
<tr>
<td>Red Admiral</td>
<td>260.0</td>
<td>75.0</td>
</tr>
<tr>
<td>Greenwear</td>
<td>-12.0</td>
<td>12.5</td>
</tr>
<tr>
<td>Purple Dragon</td>
<td>3940.0</td>
<td>112.5</td>
</tr>
<tr>
<td>Blue Horizon</td>
<td>140.0</td>
<td>112.5</td>
</tr>
</tbody>
</table>

Table 7: Normalised Scores for the Nightwear Supplier Options

This example therefore illustrates the need for benefit tradeoffs to be based on utility functions, in which the differences in scores are proportional to the differences in stakeholder preference (see Section 5.4). If we apply the default utility function of Figure 3, with UMAX = 150, to these normalised scores, we obtain the results shown in Table 8.

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Cost per garment</th>
<th>Utility (0 = critical, 100 = compliant)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Capacity</td>
<td>Reliability</td>
</tr>
<tr>
<td>White Nites</td>
<td>7</td>
<td>20.0</td>
</tr>
<tr>
<td>Red Admiral</td>
<td>8</td>
<td>148.0</td>
</tr>
<tr>
<td>Greenwear</td>
<td>10</td>
<td>-13.5</td>
</tr>
<tr>
<td>Purple Dragon</td>
<td>6</td>
<td>150.0</td>
</tr>
<tr>
<td>Blue Horizon</td>
<td>9</td>
<td>127.5</td>
</tr>
</tbody>
</table>

Table 8: Criterion Utilities for the Nightwear Supplier Options

Compared to the normalised scores, the utilities are the same for scores between 0 and 100, but penalise shortfall against the critical requirement more heavily and drastically reduce the reward for overachievement against the compliance requirement. All this is in accordance with the credibility requirements postulated in Section 5.4.2.

Figure 10 shows the overall benefit scores derived from the criterion utility scores of Table 8, for both the weighted-sum method, with the elicited criterion weights given in Table 5, and for the joint utility method, with the elicited criticalities also given in Table 5. Both sets of scores are generated on the same 0-100 scale as that used for the individual utilities: hence, the hypothetical critical option (with $x = \text{XCRIT}$ for all criteria) would return zero scores for both methods; and the compliant option (with $x = \text{XCOMP}$ for all criteria) would return scores of 100 for both methods. Note that an option with sub-critical outcomes against one or more criteria could have a negative overall score. Such a score should not be interpreted as implying that the
option is actually destructive of value; it simply means that it is inferior overall to the critical option.

Figure 10: Overall Benefit Scores for the Nightwear Supplier Options

Figure 10 shows that the rankings of the three higher-scoring options by the two metrics are the same. The ordering of the two lowest-ranking options is sensitive to the choice of metric: Purple Dragon is slightly preferred to Greenwear by the weighted sum metric, but clearly dispreferred by the joint utility metric. All options score worse against the joint utility function than against the weighted-sum function. In general whenever the sum of the criticalities is greater than 100, the joint utility will be lower than a normalised weighted sum; and the greater the inconsistency in scoring across the criteria, the greater the discrepancy. In the example, this is very evident in the case of Purple Dragon, which excels in the two most significant criteria, Capacity and Product Quality, but is nevertheless severely compromised, against the joint utility function, by its shortfalls in Product Range and Ethical Standards.
5.9.2 Cost-Benefit Comparison

Figures 11 and 12 (next page) show the cost-benefit trade-offs for the example options, using the option costs in Table 4 and the benefit scores shown in Figure 10, for the weighted-sum and joint utility benefit functions, using the heat map representation of the cost-benefit function derived in Section 5.8.3, with criticalities of 70 for cost and 80 for benefit. For each option, the large diamond represents the calculated overall benefit score, while the dotted line represents the range of scores against the individual criteria, as shown in Table 8. For example, the criterion scores for Greenwear range from -27 (for Lead Time) to 132 (Product Quality and Ethical Standards), while the overall benefit scores are 60 (weighted-sum) and 30 (joint utility). Note that the spread of scores in the same in both charts, and goes off the bottom of the chart for Purple Dragon (due to an unacceptable outcome, scoring -142).

Table 9 summarises the output of the analysis in tabular form. The scores are colour-coded for comparison with Figures 11 and 12.

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Cost (£ per garment)</th>
<th>Cost Score</th>
<th>Benefit Score</th>
<th>Cost-Benefit Score</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Weighted-sum</td>
<td>Joint benefit</td>
<td></td>
</tr>
<tr>
<td>White Nites</td>
<td>7</td>
<td>66.7</td>
<td>67.4</td>
<td>42.8</td>
<td>56.0</td>
</tr>
<tr>
<td>Red Admiral</td>
<td>8</td>
<td>33.3</td>
<td>77.4</td>
<td>51.0</td>
<td>42.8</td>
</tr>
<tr>
<td>Greenwear</td>
<td>10</td>
<td>-37.3</td>
<td>59.8</td>
<td>29.9</td>
<td>-6.7</td>
</tr>
<tr>
<td>Purple Dragon</td>
<td>6</td>
<td>100.0</td>
<td>63.2</td>
<td>13.8</td>
<td>70.5</td>
</tr>
<tr>
<td>Blue Horizon</td>
<td>9</td>
<td>0.0</td>
<td>72.0</td>
<td>47.4</td>
<td>21.5</td>
</tr>
</tbody>
</table>

Table 9: Option cost-benefit: summary of outcomes

Figures 11 and 12 offer a rich picture of the cost-benefit trade-offs offered by the five options. By plotting the cost and benefit scores on the background of the cost-benefit heat map, we express in a single graphic representation:

- the overall desirability of the benefit outcomes for each option, as expressed by the overall benefit score;
- the consistency of benefit delivery across the benefit criteria, as expressed by the range of criterion benefit scores;
- the cost of each option;
- the desirability of the cost-benefit tradeoff offered by each option, expressed by the position of each cost-benefit point relative to the utility contours; and
- for each losing option, what would need to change in order for it to come back into contention. For example, Red Admiral and White Nites lie in a region where the cost-benefit contours are climbing steeply, which means that it is much more feasible for Red Admiral to close the gap by reducing its price than to do so by enhancing its scores against the benefit criteria.
The relatively narrow band between CCOMP and CCRIT makes price a powerful discriminator, which results in Red Admiral and Blue Horizon doing badly in cost-benefit terms despite their solid performance against the benefit criteria. Greenwear performs well against the important criteria of Product Quality and Ethical Standards, but its performance against the other benefit criteria is too mediocre for it to compete against White Nites, Red Admiral and Blue Horizon even in benefit
terms; being inferior on both price and overall benefit, Greenwear is said to be dominated by these three rivals.

The most contentious option is Purple Dragon, whose fate is critically dependent on which method is used to calculate overall benefit. The fact that Purple Dragon is ranked first by the weighted-sum method highlights a well-known weakness of that method, which is its relative insensitivity to "show-stoppers". This is due to the fact that the weights in the weighted sum approach are constrained to add up to 1, which seriously constrains the extent to which variations in the outcome against any one criterion can affect the overall benefit score. This problem becomes worse as the number of criteria increases. For example, if there are 20 benefit criteria, then the average weight of each criterion can only be 5%, and if there are more than two or three criteria with weights of more than 20%, then the impact of the lowest-weighted criteria will be so small as to defeat the purpose of including them in the assessment. This is a serious problem, given that the number of decision criteria in a major defence acquisition can easily run into the hundreds.

The normal approach to this problem in the MCDA methodology is to set minimum acceptable thresholds for each benefit criterion and eliminate from consideration all options which have an unacceptable outcome against any criterion, irrespective of the weight assigned to that criterion. In our example, this would be implemented by eliminating, as a show-stopper, any option which is assessed as "unacceptable" on any of the four qualitative criteria (Product Quality, Product Range, Ethical Standards and Supplier Risk).

The weakness of this approach is its extreme sensitivity to judgements as to when a poor outcome becomes an unacceptable one. This problem is particularly acute where the criterion metric, or its underlying drivers, are continuous, so that a marginal change in the metric or its drivers can suddenly tip an option from outright winner to non-contender. This situation could arise in our example if the Ethical Standards metric is itself a composite assessment, derived from a number of indicators, some of which are continuous (for example, the minimum hourly wage or piece-rate of the supplier's workforce).

A more robust approach solution is for the response properties of the cost-benefit function, and specifically the benefit function, to be continuous, while ensuring that the overall benefit score are highly sensitive to any approach by a criterion to an unacceptable level. This is not feasible with a conventional MCDA function, derived from swing weighting, in which all the scores achieved by an option are constrained to be non-negative. However, the utilities used in our weighted-sum method are negative for outcomes worse than the critical value, and in general have no lower bound; this means that any required degree of sensitivity to under-achievement can be achieved by increasing the magnitude of the corresponding (negative) utilities. This in practice means steepening the “tail” of the utility function - that is, the shape of the curve for outcomes worse than XCRIT.

For example, we could require that any unacceptable outcome for a criterion with a weight of 10% or more would produce an overall utility of 0 or less even if all other criteria were compliant \( (u_x = 100) \). This can be achieved by assigning to unacceptable outcomes utility of -900 or worse. If we set the utility of the Ethical Standards outcome for Purple Dragon to -900, rather than -142 as shown in Table 8, the weighted-sum benefit score falls to -60, taking this option completely out of contention.
The same approach can be applied to the joint utility benefit function: in this case, the desired condition might be that a criterion with a criticality of 10% or more would produce an overall utility of 0 or less even if all other criteria were compliant \((u_x = 100)\). This likewise requires assigning to unacceptable outcomes utility of -900 or worse.

5.10 Comparison of Weighted-Sum and Joint Utility Benefit Functions

A comparison of Figures 11 and 12 shows that the outcome of the comparison of overall benefit, and hence the cost-benefit comparison, can vary depending on whether the overall benefit score is derived from a weighted-sum or a joint utility function. This raises the question as to how to which is the more appropriate method for a given decision analysis.

From a theoretical perspective, the joint utility approach is superior in that it represents the actual preferences of the stakeholders with regard to tradeoffs between decision criteria, as expressed by the elicited criticalities, with much greater fidelity than the weighted-sum approach. To recapitulate:

1. The criticality of criterion \(j\) is the estimate, elicited from the stakeholders, of the change in overall utility when, starting from the compliant outcome \((x = XCOMP\) for all criteria), criterion \(j\) is "swung" from \(XCOMP\) to \(XCRIT\), given that the effect of swinging all criteria from \(XCOMP\) to \(XCRIT\) is to reduce overall utility from 100 to 0.

2. Any combination of criterion criticalities can be captured in a joint utility function.

3. A combination of criterion criticalities can be captured in a weighted-sum formula only if the criticalities happen to sum to 100 or thereabouts. In this case, the weight of each criterion is proportional to its criticality.

4. The deviation of the weighted-sum formula from the behaviour which is implied by the criticalities increases as the sum of the criticalities moves further from 100.

5. As the number of criteria increases, the weight assigned to each criterion by the weighted-sum formula must reduce, since they are constrained to add up to 100. This means that the sensitivity of the formula to each single criterion diminishes. With the joint utility function, the criticality of each criterion is independent of the number of other criteria and their criticalities.

It follows that:

- The fidelity with which the joint utility function reflects elicited stakeholder preferences will always either equal or exceed the fidelity of the weighted-sum function.
- This advantage is likely to increase as the number of benefit criteria increases.

The weighted-sum approach has the following advantages over the joint utility approach.

1. It is an established and widely-recognised technique with an extensive history of use in a wide range of decision-making environments.
2. It is simple, transparent and easy to communicate.

3. It is the technique recommended by DSPCR 2011 for defence and security acquisition, and by the Treasury Green Book (Reference 15) for public-sector acquisition in general.\textsuperscript{28}

The fact that the weighted-sum approach is computationally simpler than the joint utility approach is not significant in itself, since the joint utility approach computations are easy to implement in a spreadsheet or bespoke tool; however, the simplicity of the weighted-sum formula is much more intuitive and appealing than the more abstruse joint utility formula. This creates the possibility that it may, paradoxically, be easier to obtain stakeholder buy-in for a weighted-sum formulation than for a joint utility formulation, even when the latter is more faithful to the stakeholders' actual preferences.

\textsuperscript{28}“The most common technique used to compare both unvalued costs and benefits is … multi-criteria analysis. [This] involves assigning weights to criteria, and then scoring options in terms of how well they perform against those weighted criteria. The weighted … sums can be used to rank options. The weights … incorporate the judgments of stakeholders and decision makers”. (Reference 15, paras 5.78 – 5.79.)
6 Comparison of Joint Utility Method with DSPCR Guidance

In this section we show that the methods for tender evaluation which are offered as examples in DSPCR 2011 are not wholly satisfactory in delivering the requirements for consistency and transparency which are stipulated in those regulations, and assess the extent to which the methods described in Section 5 would address these issues.

6.1 Example of DSPCR tender evaluation process

The following example is adapted from DSPCR 2011, Chapter 15, Annex B: Most Economically Advantageous Tender Assessment, Example 1 (Reference 6). The recommended method is described in this guidance as follows:

“In this example, you establish the most economically advantageous tender by weighting the technical evaluation score against the financial offer score on a 60/40 split… marks out of 60 are allocated to the best technically compliant bid, and marks out of 40 to the best price. The other marks are calculated using a percentage difference method … this approach is illustrated in the European Commission publication Practical Guide to Contract Procedures for EU External Actions dated November 2010 (Updated March 2011).”

In our adapted example, the technical marks and prices of the tenders being compared are as follows:

<table>
<thead>
<tr>
<th>Tender</th>
<th>Technical Mark out of 60</th>
<th>Financial Offer (£M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tender A</td>
<td>38</td>
<td>140</td>
</tr>
<tr>
<td>Tender B</td>
<td>25</td>
<td>100</td>
</tr>
<tr>
<td>Tender C</td>
<td>56</td>
<td>190</td>
</tr>
</tbody>
</table>

Table 10: MEAT: Technical Marks and Price

According to guidance:

- “The tender with the higher technical mark receives 60 points, in line with the 60/40 split.” In this example, Tender C has the highest technical mark, 56.

- “The other tenders are awarded points representing how far they fall short of the best mark, by means of the % difference”. For example, Tender B scores 25 and so receives (25/56) x 60 = 26.8 marks.

The resulting technical scores are shown in Table 11

<table>
<thead>
<tr>
<th>Tender</th>
<th>Technical Mark out of 60</th>
<th>Technical Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tender A</td>
<td>38</td>
<td>40.7</td>
</tr>
<tr>
<td>Tender B</td>
<td>25</td>
<td>26.8</td>
</tr>
<tr>
<td>Tender C</td>
<td>56</td>
<td>60</td>
</tr>
</tbody>
</table>

Table 11: MEAT: Technical Scoring

Likewise:
“The lowest priced tender receives 40 points, in line with the 60/40 split.” In this example, Tender B has the lowest price, £100M.

“The other tenders are awarded points representing how far they fall short of the lowest price, by means of the % difference. First, the bids are calculated as a percentage difference against the lowest price tender”. For example, Tender C bid £190M and so the percentage difference is \((190-100)/100 = 90\%\).

“The bids are then converted into a percentage of 40°. For example, Tender C scores \((100\% - 90\%) \times 40 = 4\).

The resulting technical scores are shown in Table 12.

<table>
<thead>
<tr>
<th>Tender</th>
<th>Financial Offer (£M)</th>
<th>% difference</th>
<th>Financial Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>140</td>
<td>40%</td>
<td>24</td>
</tr>
<tr>
<td>B</td>
<td>100</td>
<td>0%</td>
<td>40</td>
</tr>
<tr>
<td>C</td>
<td>190</td>
<td>90%</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 12: MEAT: Price Evaluation

Finally the results are combined, as shown in Table 13. The winning score, that of Tender B, is highlighted. Note that Tenders A and B are both preferred to C.

<table>
<thead>
<tr>
<th>Tender</th>
<th>Technical Score</th>
<th>Financial Score</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>40.7</td>
<td>24</td>
<td>64.7</td>
</tr>
<tr>
<td>B</td>
<td>26.8</td>
<td>40</td>
<td>66.9</td>
</tr>
<tr>
<td>C</td>
<td>60</td>
<td>4</td>
<td>64</td>
</tr>
</tbody>
</table>

Table 13: MEAT: Combined Technical and Price Evaluation

6.2 Issues with the DSPCR tender evaluation process

We now consider what the outcome would have been in this example if Tender C had not been submitted. Since Tender C was dispreferred to both A and B, it would be expected that its absence would make no difference to the comparison between A and B. However, the working shown in Table 14 shows that this is not the case: in the absence of Tender C, A is preferred to B.
### Table 14: MEAT: Combined Technical and Price Evaluation (Tenders A and B only)

The reason for this anomalous behaviour is the practice of awarding full marks to the best outcome in each criterion, irrespective of whether that outcome is good or bad relative to the contracting authority’s requirements, and irrespective also of whether the tender that offers this outcome is also a credible contender in other respects. In our example, Tender B is technically inferior to A but cheaper: when C is introduced, its high technical score drives down the technical scores of A and B (from 60 and 39.5 to 40.7 and 26.8), thereby reducing the differential between them from 21 to 14, while the financial scores are unchanged, since C is not the cheapest option. As a result, A’s better technical score no longer prevails against B’s lower cost.

### 6.3 Utility-Based formulation

We now consider the outcome of applying the techniques developed in Section 5 to this example. To do this, we must first define critical and compliant values for the Technical Merit and Price criteria. For illustrative purposes, we postulate the values listed in Table 15.

### Table 15: Critical and Compliant Values for the MEAT Assessment

In order to remain faithful to the intentions of the original example, we will assume that tradeoffs will be linear within the core tradeoff region. The joint utility function is therefore simply

\[ UCB[c, b] = 0.4*UC[c] + 0.6*UB[b], \]

where UC and UB are the criterion utility functions shown in Figures 13 and 14 below. These functions are derived from Equations 13 and 14 with UMAX = 150.
Figure 13: Cost utility function for the MEAT assessment

Figure 14: Technical Merit utility function for the MEAT assessment
Table 16 and Figure 15 show the outcome in tabular and graphical form. These show that the cost-benefit tradeoff offered by Tender A is marginally more attractive than that of Tender B, while C is put out of contention by its excessive cost. Note that, although the technical mark is more heavily weighted than the price, this is not quite sufficient to outweigh Tender B’s lower cost.

<table>
<thead>
<tr>
<th></th>
<th>Tender A</th>
<th>Tender B</th>
<th>Tender C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tech mark</td>
<td>38</td>
<td>25</td>
<td>56</td>
</tr>
<tr>
<td>Cost</td>
<td>140</td>
<td>100</td>
<td>190</td>
</tr>
<tr>
<td>Normalised tech score</td>
<td>26.7</td>
<td>-16.7</td>
<td>86.7</td>
</tr>
<tr>
<td>Normalised cost score</td>
<td>20.0</td>
<td>100.0</td>
<td>-80.0</td>
</tr>
<tr>
<td>Tech utility</td>
<td>26.7</td>
<td>-17.6</td>
<td>96.7</td>
</tr>
<tr>
<td>Cost utility</td>
<td>20.0</td>
<td>100.0</td>
<td>-103.7</td>
</tr>
<tr>
<td>CB score</td>
<td>24.0</td>
<td>29.4</td>
<td>9.7</td>
</tr>
</tbody>
</table>

Table 16: Utility-based Combined Technical and Price Evaluation

If Figure 15 (without the actual tenders) was published as part of the invitation to tender, it would allow a bidder to see at a glance what overall score would be assigned to any combination of price and technical score that they may wish to offer. This is impossible with the method described in Section 6.1 since the score assigned to one tender is critically dependent on the unknown properties of rival tenders. It would also ensure that every bidder would, from the range of options that they would be willing to offer, chose one which maximised the overall utility to the contracting authority. The method of Section 6.1 offers no assurance of this, and creates the risk of a “lose-lose” outcome in which a tender could be contracted despite the fact that there was a solution available which would have been preferred by both the winning bidder and the contracting authority.

6.4 Capturing the Purchaser’s Cost-Benefit Tradeoff Preferences

In addition to the advantages described in the previous section, the utility-based method offers to a contracting authority a much greater range of cost-benefit tradeoff preferences than is possible with the tender evaluation scheme described.
in DSPCR 2011. In this example, if the contracting authority decided - in advance of the competition - that the heat map shown in Figure 15 was unduly favourable to potential tenders which were outside the critical boundaries for Technical Score or Price, then this could be remedied using the methods described in Section 5.8.4. For example, if the criticalities assigned to the Technical Score and Price criteria were 70 and 80 respectively, (as in the worked example of Section 5) then the joint utility function would become:

\[
UCB[c, b] = 0.2*UC[c] + 0.3*UB[b] + 0.005*UC[c]*UB[b].
\]

Equation 44

For the tenders of Table 10, this would result in the outcomes shown in Table 17 and Figure 16.

<table>
<thead>
<tr>
<th>Tech mark</th>
<th>Tender A</th>
<th>Tender B</th>
<th>Tender C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>140</td>
<td>100</td>
<td>190</td>
</tr>
<tr>
<td>Normalised tech score</td>
<td>26.7</td>
<td>-16.7</td>
<td>86.7</td>
</tr>
<tr>
<td>Normalised cost score</td>
<td>20.0</td>
<td>100.0</td>
<td>-50.0</td>
</tr>
<tr>
<td>Tech utility</td>
<td>26.7</td>
<td>-17.6</td>
<td>86.7</td>
</tr>
<tr>
<td>Cost utility</td>
<td>20.0</td>
<td>100.0</td>
<td>-105.7</td>
</tr>
<tr>
<td>CS score (JU)</td>
<td>14.7</td>
<td>5.9</td>
<td>-40.0</td>
</tr>
</tbody>
</table>

Table 17: Combined Technical and Price Evaluation with non-linear joint utility.

Comparing Figures 15 and 16, it can be seen that the points representing the three tenders lie in the same positions in the x-y plane, but the underlying heat map shows how differently these outcome would be scored if the non-linear formula of Equation 44 was used instead of Equation 43. As intended, Equation 44 rewards tenders which achieve at least the critical levels against both criteria; this is expressed in Figure 16 by the greater rectilinearity of the utility contours. As a result, tender A, the only tender which costs less than £150M and achieves a
Technical Mark of more than 30, is the preferred tender with this assessment scheme.
7 Summary and Conclusions

7.1 Summary of Proposed Method

For ease of reference, the summary of the proposed process at Section 5.1 is reproduced here, as follows:

1. The assessment criteria for the required acquisition are identified.
2. A unitary cost metric is developed, using the established processes of Investment Appraisal.
3. Metrics are assigned to each benefit criterion other than cost: these may explicit measures of real-world quantities (such as probability of survival), or scores based on qualitative assessment (e.g. 0 = “unacceptable”, 1 = “poor”, 2 = “moderate”, etc.)
4. Criterion utility functions are elicited for each metric.
5. An aggregated “Measure of Benefit” function is developed as an analytic function of the utility scores obtained for the individual benefit criteria other than cost.
6. A bivariate joint utility function is developed for the cost and aggregate benefit.
7. A “heat map” for the cost/benefit tradeoff space is generated from the joint utility function and plotted.

Supplemented by formal specifications of the criterion and joint utility functions, the heat map provides a simple means of communicating to potential tenderers the precise criteria by which the tender which offers the most favourable cost-benefit trade-off will be selected.

7.2 Assessment

Section 4 specifies the requirements for a process for options or tender assessment that meets both the transparency requirements of public procurement in general, and defence procurement in particular, as follows:

1. It must be possible to fully specify the criteria against which a tender is assessed, without any knowledge of or assumptions about the tenders that will be submitted.
2. The scoring criteria must be objective and stable: that is:
   - the scoring of any one tender must depend solely on its own characteristics;
   - the comparison between any two tenders must not be influenced by the presence or absence of a third tender.
3. The process must be formal and communicable: for example, in the form of a practitioners’ guide.
4. The process must be auditable: there must be no ambiguity as to whether the process is being followed in a particular application; if the process offers a choice of procedures, it must always be possible to establish which option has been pursued, and the evidence on which that decision was made.

5. The process must be practicable in terms of timescale, affordability and information demand;

6. The process must be theoretically sound and conforms to the basic requirements of decision theory with regard to rationality and consistency.

7. Overall metrics for monetary and non-monetary decision criteria are calculated separately, and only traded off against each other at the final stage of the process.

At their current stage of development, the utility-based method described in this paper can be assessed against these criteria as follows:

7.2.1 Specification in advance

The method allows the award criteria to be disclosed in advance of options being developed, including the criteria for evaluating cost-benefit trade-offs (Figure 8).

7.2.2 Objective and Stable Scoring Criteria

Once elicited and endorsed, the joint utility functions, and the utility contours derived from them, form an objective basis against which any option can be assessed. When applied to Tender assessment, the process ensures complete consistency and transparency in the scoring of each tender, thereby addressing the deficiencies in these respects of both the methods recommended in DSPCR 2011 (see Section 6) and those version of MCDA which rely on swing weighting for criterion weighting.\(^{29}\)

7.2.3 Formal and Communicable Process

The method is under development and has yet to be specified in sufficient detail to create a practitioners guide or a community of practitioners. Even when matured, it will still be more complex than the traditional MCDA techniques recommended by DSPCR 2011. However, there is no inherent obstacle to the process being formalised.

It is envisaged that a suite of spreadsheet-based tools will be developed to support the practice of the method. Prototype tools were developed in the course of the Dstl tasking (Reference 4), these are also at an early stage of development, but will be developed into configured, verified and validated software products as the process matures.

---

\(^{29}\) Swing weighting has the same drawback as the DSPCR methods described in section 6: the scores assigned to a tender depend on the weights assigned to each criterion, which in turn depends on the characteristics of the rival tenders.
7.2.4 Auditability

The evaluation process is fully specified in advance of the submission of options for assessment, in the form of the benefit metrics, utility functions, joint benefit utility function and cost-benefit function, and so provides complete auditability of any contract award decision based on the benefit and cost criteria to which these functions apply.

As with any decision support process based on elicited stakeholder preferences, the credibility of the outcome depends on the credibility of the stakeholder judgements underpinning the elicited utilities, and on the skill of the practitioner in eliciting, articulating and confirming these judgements. Since there is at present no body of practitioners for this method, it is not possible estimate at present the extent to which the credibility and robustness of the outcome of the process is likely to enhanced or degraded according to the competence of the practitioner.

7.2.5 Practicality

The mechanics of the method were been tested in the Dstl tasking and in the examples described in this paper, and have been shown to be straightforward.

The main challenge in applying the method in practice will be the elicitation of credible and consistent inputs from the stakeholder community for the parameters required to populate the utility functions: in particular, the benefit criterion criticalities and the cost-benefit joint utility function. A number of options are described in Section 5 to approach these problems, but none of them has yet been fully tested.

7.2.6 Theoretical soundness

The method is based on established decision analysis techniques; in particular, the theory of decision making with multiple objectives developed in the late 1960s by Ralph Keeney and Howard Raiffa. The formalisation of the process is well advanced and will present no difficulties in verification. It is demonstrably superior to traditional decision support techniques in consistency, stability and transparency (Sections 6.3, 6.4).

7.2.7 Separate calculation of cost and benefit outcomes

This is the fundamental design feature of the method. Compared to the existing techniques used by public contracting authorities in the UK, the method is designed to surpass them in its ability to evaluate and communicate the value of the tradeoffs available between cost and benefit across the whole of the cost-benefit tradeoff space. This is achieved by the creation of a cost-benefit joint utility function which more faithfully represents stakeholder preferences, and the heat map derived from it (Section 5.10), which more effectively communicates them.

7.3 Conclusions and Observations

1. There is a gap between the statutory requirements for objectivity and transparency in the award criteria for public sector acquisition, as specified in the 2006 Public Contract Regulations (PCR 2006) and the 2011 Defence and

---

30 Reference 14. In particular, the joint utility functions developed in Section 5.6.2 and 5.8.2 are derived from the “multiplicative utility functions” described in pp 288-291 of this text.
Security Public Contract Regulations (DSPCR 2011), and the capabilities of the decision analysis and cost-effectiveness analysis methods which are currently endorsed for public sector acquisition in the UK.

Specifically:

- The Multiple-Criteria Decision Analysis (MCDA) techniques recommended in DSPCR are subject to a dilemma. If the weightings assigned to criteria are fixed in advance, the MCDA objective function will give excessive weight to over-achievement and insufficient weight to potentially critical shortfalls (Section 5.9.1). If, conversely, swing weighting or other forms of adaptive weighting are used, then the process will be deficient in consistency and transparency (Section 6.3); for example, the ranking of two tenders could be reversed, depending on whether a third tender was submitted or withheld.

- The cost-effectiveness and COEIA methods recommended in JSP 507 will not produce outcomes which are sufficient to justify the award or withholding of an acquisition contract. This is because cost and operational effectiveness will not normally be the sole criteria for such an award.

- Cost-benefit analysis (CBA) techniques able to capture a wider range of criteria, by using the device of assigning, to each decision criterion, a financial equivalent to each of its possible outcomes. However, Treasury guidance, as summarised in the “Green Book”, recognises that not all the criteria relevant to public sector acquisition can be “monetarised” in this way (Section 5.10). This is particularly true of defence acquisition: JSP 507 specifically states that “as few activities within MoD produce benefits that can be valued in monetary terms, the use of full-blown CBA by MoD is extremely limited.”

2. The “classical” von Neumann – Morgenstern theory of utility provides, in theory, a complete solution to the problem of decision-making under uncertainty, for single or multiple objectives. Later theorists such as Keeney and Raiffa have worked on the development of practical applications of utility theory, with some success.\(^{31}\) However, the “textbook” form technique is labour-intensive, lacks formal structure or process and difficult to audit. In the context of consultancy to commercial decision-making, where decision-makers have a high degree of autonomy, these attributes are not necessarily serious drawbacks. However, for public sector acquisition, a more formal, auditable and communicable process is required.

3. The methods described in this paper are an attempt to formalise and systematise the application of utility theory to public sector acquisition, thereby creating a decision analysis process which addresses the limitations of the methods currently in use. In comparison to existing MCDA techniques, the

\(^{31}\) For case studies, see Keeney and Raiffa (Reference 14), chapters 7 and 8.
utility-based approach can capture stakeholder preferences between different benefit-benefit or cost-benefit tradeoffs with much greater fidelity, while maintaining absolute consistency in the comparative scoring of options. In comparison with COEIA and CBA techniques, it allows all the criteria governing contract award to be evaluated within a single framework.

4. The method is still at the conceptual stage; there is present no body of practitioners and no unclassified case studies. Hence no firm conclusions can be drawn with regard to its practicality, or the timescales and resources which an assessment using these methods will require. It is envisaged that the use of standardised or “default” forms of the criterion utility functions and cost-benefit joint utility function could reduce the workload to a level similar to that required by a conventional MCDA appraisal, in which case the greater fidelity of the utility-based method will be gained at little or no cost; however, this remains to be verified.

5. The cost-benefit heat map (Figures 8, 11, 12, 15) is a simple and powerful method of articulating and communication the preferences of a purchaser of contracting authority over the full range of possible offerings, without the need for prior knowledge of the offerings actually tendered. It should be noted that the advantages of the heat map for of presentation apply irrespective of how the “benefit” parameter is calculated; for example, it would be easy to express the cost-effectiveness outcomes of a COEIA as a heat map.
8 References

1. QinetiQ proposal for COEIA Pilot Study. QINETIQ/12/02412 Issue No 1.0 dated 15 October 2012
Annex A: The Multivariate Joint Utility Functions

This annex specifies two functions which, given a set of \( n \) decision criteria, each of which has its own utility function of the form specified in the main text, will generate a joint utility for any combination of values of the individual metrics.

A1 Formulation

Let \( x_1 \) to \( x_n \) be the metrics for a multiple-criterion decision problem. For \( j = 1 \) to \( n \), let \( U_j[x_j] \) be the elicited utility function for \( x_j \); let \( X_{Cj} \) be the “critical” value of \( x_j \) such that \( U_j[X_{Cj}] = 0 \); and let \( X_{Sj} \) be the “compliant” or “satisfactory” value of \( x_j \) such that \( U_j[X_{Sj}] = 100 \).

We require a joint utility function \( U_J[x_1 \ldots x_n] \) such that:

1. \( U_J[x_1 \ldots x_n] \) is defined for every feasible combination of metric values \( x_1 \) to \( x_n \)
2. \( U_J \) is monotonic in \( U_1[x_1] \) to \( U_n[x_n] \); that is, for all \( x_1 \) to \( x_n \), an increase in any \( U_i[x_i] \) will never cause a decrease in \( U_J \).
3. \( U_J[X_{C1} \ldots X_{Cn}] = 0 \); \hspace{1cm} \text{Equation A1}
4. \( U_J[X_{S1} \ldots X_{Sn}] = 100 \); \hspace{1cm} \text{Equation A2}

We first define the critical option to be the hypothetical option such that \( x_i = X_{Cj} \) for all criteria; and the compliant option to be the hypothetical option such that \( x_i = X_{Sj} \) for all criteria. Then:

- Taking the critical option as a starting point, the impact of criterion \( j \), \( M_j \), is the significance of a “swing” in criterion \( j \) from \( X_{Cj} \) up to \( X_{Sj} \), given that a swing from \( X_{C} \) to \( X_{S} \) for all criteria is credited with a significance of 100.
- Taking the compliant option as a starting point, the criticality of criterion \( j \), \( C_j \), is the significance of a “swing” in criterion \( j \) from \( X_{Sj} \) down to \( X_{Cj} \), given that a swing from \( X_{S} \) to \( X_{C} \) for all criteria is credited with a significance of 100.

We can then develop two joint utility functions; the joint impact utility function, which is parameterised by the criterion impacts, and the joint criticality utility function, which is parameterised by the criterion criticalities. These are specified in the main text (Section 5.6.2). In this Annex, we derive joint criticality utility function, which is envisaged as being the one which would be more commonly used. The derivation of the joint impact utility function is analogous.

---

32 This is adopted as a more compact notation than that used in the main text, in which these quantities are termed XCRIT and XCOMP respectively.
A2 Derivation: Joint Impact Utility Function

We postulate a joint utility function of the form

\[ U_J[x_1 \ldots x_n] = \alpha + \beta \prod_{j=1}^{n} [1 + \lambda_j^* U_j[x_j]/100]. \]  

*Equation A3*

From Equation A2, we have

\[ 0 = U_J[X_{C1} \ldots X_{Cn}] \]

\[ = \alpha + \beta \prod_{j=1}^{n} [1 + \lambda_j^* 0] \]

\[ = \alpha + \beta; \]

so equation A3 simplifies to

\[ U_J[x_1 \ldots x_n] = \alpha^* (1 - \prod_{j=1}^{n} [1 + \lambda_j^* U_j[x_j]/100]). \]  

*Equation A4*

To determine the remaining parameters, we must elicit from stakeholders the joint utility of each impact reference outcome. For \( j = 1 \) to \( n \), let \( M_j \) be the elicited joint utility of the outcome \( [X_{C1} \ldots, X_{S_j}, \ldots X_{Cn}] \); that is the outcome for which \( x_j = X_{S_j} \) and \( x_r = X_{C_r} \) for all \( r \neq j \). It is assumed that, for all \( j \), \( 0 < M_j < 100 \).

Then for all \( j \), we have

\[ M_j = U_J[X_{C1} \ldots, X_{S_j}, \ldots X_{Cn}] \]

\[ = \alpha^* (1 - (1+\lambda_j)) \]

\[ = -\alpha^* \lambda_j \]

so equation A4 in turn simplifies to

\[ U_J[x_1 \ldots x_n] = \alpha^* (1 - \prod_{j=1}^{n} [1 + \lambda_j^*/\alpha]). \]  

*Equation A5*

Finally, from Equation A1, we have

\[ 100 = U_J[X_{S1} \ldots X_{Sn}] \]

\[ = \alpha^* (1 - \prod_{j=1}^{n} [1 - M_j/\alpha]). \]  

*Equation A6*

Define \( k = -100/\alpha \); then Equations A5 and A6 can be rewritten as

\[ U_J[x_1 \ldots x_n] = (100/k)^*(\prod_{j=1}^{n} [1 + k^* M_j^* U_j[x_j]/10000] - 1), \]  

*Equation A7*

where

\[ I + k = \prod_{j=1}^{n} [1 + k^* M_j^* /100]. \]  

*Equation A8*

It can be shown that Equation A8 has a unique non-zero solution for \( k \) if and only if \( \sum_{j=1}^{n} M_j \neq 100 \), and that the limiting solution as the sum approaches 100 is the simple weighted sum:

\[ U_J[x_1 \ldots x_n] = \sum_{j=1}^{n} M_j^* U_j[x_j]/100 \]  

*Equation A9*

Equation A7 cannot be used as a joint utility function as it stands since it is not monotonic for all possible outcomes. This is corrected as follows. From Equation A7, we have

---

33 This is the general form of what Keeney and Raiffa term the “multiplicative utility function” (Reference 14, p.289). They derive a utility function which is equivalent to that given in Equations A7 and A8.
\[ \frac{\partial U_J}{\partial U_j} = (100/k) \cdot \frac{\partial \Pi_{j=1}^{n} [1 + k \cdot M_j \cdot U_j(x)]}{10000} \]
\[ = (100/k) \cdot (k \cdot M_j \cdot U_j(x)/10000) \]
\[ = (M_j /100) \cdot \Pi_{r \neq j} [1 + k \cdot M_r \cdot U_r(x)/10000] \]
\[
\text{Equation A10}
\]

M_j is positive for all j, so a sufficient condition for \( \frac{\partial U_J}{\partial U_j} \) to be non-negative for all j is that, for all j,
\[
1 + k \cdot M_j \cdot U_j(x)/10000 \geq 0;
\]
\[
\text{Equation A11}
\]

This condition can be assured by replacing the criterion utility functions \( U_j(x) \) in Equation A7 by the truncated utility functions:
\[
U_{TI} (x_j) = \max \{ U_{MIN} (x_j), \min \{ U_{MAX} (x_j), u_j(x) \} \}.
\]
\[
\text{Equation A12}
\]

where
\[
U_{MIN} = -9999 \quad k \leq 0
\]
\[
-10000/k \cdot M_j \quad \text{otherwise};
\]
\[
\text{Equation A13}
\]
\[
U_{MAX} = -9999 \quad k < 0
\]
\[
9999 \quad \text{otherwise}.
\]
\[
\text{Equation A14}
\]

This yields the definitive form of the joint impact utility function, as given in the main text (Equations 17 to 22):

If \( M_1 + \ldots + M_n = 100 \) then
\[
U_J (x) = M_1 \cdot u_1(x_1) + \ldots + M_n \cdot u_n(x_n);
\]
\[
\text{Equation A15}
\]

Otherwise,
\[
U_J (x_1 \ldots x_n) = (100/k) \cdot (\Pi_{j=1}^{n} [1 + k \cdot M_j \cdot U_{TI}(x)/10000] - 1),
\]
\[
\text{Equation A16}
\]

where k is the unique non-zero real number such that
\[
l + k = \Pi_{j=1}^{n} [1 + k \cdot M_j /100];
\]
\[
\text{Equation A17}
\]

and \( U_{TI} (x_j) \) is as defined in Equations A12 to A14.

### A3 Joint Criticality Utility Function

The corresponding joint utility function, derived from elicited criticalities rather than impacts, is specified below. Its derivation is analogous to that of the joint impact utility function.

Two cases arise:

**Case 1: \( C_1 + \ldots + C_n = 100 \cdot (n-1) \)**

In this case the metrics are linearly independent and the joint utility function is
\[
U_J (x_1 \ldots x_n) = (1-C_1/100) \cdot u_1(x_1) + \ldots + (1-C_n/100) \cdot u_n(x_n).
\]
\[
\text{Equation A18}
\]

**Case 2: \( U_X_1 + \ldots + U_X_n \neq 100 \)**

In this case, there exists a unique non-zero real number h such that
\[
1+h = \Pi_{j=1}^{n} [1+h \cdot C_n /100].
\]
\[
\text{Equation A19}
\]

We then define, for \( j = 1 \) to \( n \),
\[
U_{MIN} = -100 \cdot (1+h)/(h \cdot C_j) \quad h \leq 0
\]
\[
-9999 \quad \text{otherwise};
\]
\[
\text{Equation A20}
\]
\[
U_{MAX} = 9999 \quad h < 0
\]
\[ 100 \times (1 + \frac{100}{(h \times C_j)}) \quad \text{otherwise} \] \hspace{1cm} \text{Equation A21}

and the truncated criterion utility function

\[ UTC_j[x_j] = \max\{ UCMIN_i, \min\{ UCMAX_i, u_i[x_n]\} \}. \hspace{1cm} \text{Equation A22} \]

Then the joint utility function is

\[ UJC[x_1 \ldots x_n] = \left( \frac{100}{h} \right) \times \left( 1 + h \times \prod_{j=1}^{n} \left[ 1 + h \times C_n \times \left( 1 - UTC_j[x_j]/100 \right)/100 \right] \right). \hspace{1cm} \text{Equation A23} \]
# Utility-based Decision Support for Investment and Acquisition

## Abstract

This paper is an unclassified summary of a recent study, conducted by QinetiQ for Dstl PCS, into ways of ensuring that the evaluation processes for defence equipment and services are demonstrably consistent with the legal requirements on transparency arising from DSPCR 2011 and recent case law.

The paper proposes multivariate utility theory as a method for ensuring fairness, transparency and rigour in developing and communicating the decision criteria for defence acquisitions. It identifies the issues which the decision-making process must address, and develops a solution, in which the academic formulation of utility theory is developed into a workable, repeatable, and auditable process. The utility-based approach is compared with the weighted-sum methods recommended in DSPCR 2011, and their respective strengths and weaknesses are identified.

The method generates a benefit evaluation function and a cost-benefit indifference map against which the cost-benefit tradeoffs offered by competing acquisition solutions can be plotted and ranked: these can be generated and published in advance of the submission of offerings. These results can be presented in the form of a heat map: a chart of benefit versus cost in which the chart background is colour-coded to indicate the overall score that would be assigned to any cost-benefit outcome.

It is shown that, because of its greater adaptability to elicited stakeholder preference, the method can be applied to a significantly greater range of problems than weighted-sum methods.

## Keywords / Descriptors

Utility, Decision Support, MCDA, Tender Assessment, COEIA