An Evidential Reasoning Approach to Contractor Evaluation in the Military Aircraft Sector

Authors

Dr G Graham[1], Business School, University of Salford, Salford, M5 4WT.
Tel: 0161 295 5336
E-mail: g.graham@business.salford.ac.uk

Dr J B Yang and Mr M Sonmez, Manchester School of Management, UMIST, Manchester, M60 1QD
Tel: 0161 200 3427
E-mail: jian-bo.yang@umist.ac.uk

Overview

This paper reports the application of an evidential reasoning (ER) approach, to deal with the evaluation of a contractor, from among different fabrication options of aero-engine equipment on a military tanker-transport aircraft. This is followed by a review of current evaluation practices, which begins by highlighting their limitations and then goes on to justify the use of an evidential reasoning approach. The evaluation model is discussed and techniques for articulating the original evaluation data are also explored. A hypothetical selection problem involving the evaluation of different fabrication options to aero-engine equipment are then examined using this approach.

[1] Please address all of your correspondence to Dr Graham.

An Evidential Reasoning Approach to Contractor Evaluation in the Military Aircraft Sector

1. Introduction

Tendering is an approach largely associated with the award of project contracts, in the aerospace sector [1]. A project is a highly complex transaction, involving products and services integrated through a work to deliver facility or an enhanced organisational capability, of some description. They aim to deliver beneficial, one-off transformational changes not achieved through improvements in the efficiency of existing operations [2]. Their purpose is to enable a customer to obtain certain business benefits within the constraints of time, cost, and quality in order to justify the investment [3]. Despite this, however, there has been no commensurate improvement in the performance of aerospace project management. Instead, there has been extensive delays in the planned schedule, cost overruns, and an increased number of claims and litigation.

The standard conditions for awarding contracts in the defence industry have remained relatively unchanged for nearly twenty years [4, 5]. Many now believe, that the prime contractor approaches to contractor evaluation, concentrating as it does solely on bid price [6, 7], is one of the major causes of project delivery problems [8, 9, 10]. This leads to suppliers that are facing a shortage of work, to be more likely to enter unrealistically low bids, simply to stay in business in the short term. The aim is to somehow raise additional income on project commencement through 'claims' or cutting costs to compensate. The main result of this, is a growing interest in looking for techniques of contractor evaluation, which utilise information, concerning client objectives and supplier capabilities as a means of achieving the best value for money [5].

2. Contractor Evaluation Models: Background and Previous Work

The evaluation of a contractor is fairly complex because one must cater for many performance criteria, which might be either quantitative or qualitative in nature. A simple and currently adopted approach is that of criteria aggregation; it uses certain scoring mechanisms to add up weighted scores [11]. There are a number of theoretical drawbacks of this simple weighted approach [12], and it is quite difficult to define appropriate measures and collect relevant assessment information. The subjective nature and possible absence of evidence means that such scoring is hardly free from vagueness and imprecision. It is very difficult and demanding to pinpoint subjective assessments to a single number in a reliable and consistent way.

The evaluation of contractor selection options for supplying complex products, often, formally involves the use of multicriteria decision analysis methods (see e.g. [13] for a general overview of this well-established technique and [14] for a more comprehensive account). These methods tend to be appropriate when there are many conflicting objectives and sensitivity testing is undertaken with several stakeholders, respectively (e.g. [11, 14, 15]. By far, the main difficulty encountered with multicriteria methods is trying to compare different criteria, which have been measured on different scales. Ellis and Herbsman [16] propose a time and cost approach, to determine the bid winner. In converting the contract time to a cost to the client, there is a straightforward comparison on a single criterion. Holt et al [17] combines what they term P2 scores (representing the scores of the information collected) and their P3 scores (representing the bid price) into a simple index. The simple index is determined by assigning a 40% weighting to P2 scores and a 60% weighting to the P3 scores. This approach is deemed satisfactory for those decision problems where there is a single clear dominant objective, which dominates all others and this is normally profit. Recently, it seems, that, whilst most decision-makers may want to maximise profit in the defence industry, they might have other objectives that are of equal importance to them. It is evident from previous work that while buyers may want to minimise the likely cost of a project, they might also want contractors to maintain schedules, as well as achieving stringent quality targets.

This paper is devoted to applying an ER approach to treat the aforementioned contractor selection problem [18]. This approach has been developed to deal with multiple attribute decision making problems with both qualitative and quantitative attributes where each attribute can have its own sub-criteria, which then could be assessed using subjective judgements with complete or incomplete uncertainty [13, 14]. The main advantage of this approach, if compared with other multicriteria methods, lies in the ability to deal with incomplete uncertainty in a more natural yet rational way. Since certainty could be viewed as a special case of uncertainty, the application of this approach to contractor evaluation is based on the transformation of the original data into equivalent subjective statements with complete uncertainty. If the precise numerical values are not available, it is then more natural to articulate subjective judgements with uncertainty as original evaluation data. This approach can be further substantiated, in that it has been used, as an alternative tool to deal with real-world complex decision analysis problems in engineering design and manufacture [12, 21]. A window-based intelligent decision system (IDS) using the evidential reasoning approach is used to support this reaearch.

3. Sources of Evaluation Criteria and Weightings

The evidential reasoning approach has been applied to different fabrication options for compliance with buyer requirements of A4 SAT (a Supplier Assessment Tool). The options considered to meet the evaluation requirements are bid price, financial capacity, the quality system, technical ability and reputation (Table 1).

Table 1 Weights of the main criteria

Criteria	Bid amount	Financial capacity	Quality system	Technical ability	Reputation
Weight	0.3	0.2	0.1	0.3	0.1

To accommodate the needs of the client and the project, relative weights needed to be assigned to the main criteria. This was done by a team of four professional assessors at the buyer first ranking the criteria in order of relative importance on a likert scale from 1 to 7 (where 1 is unimportant and 7 is extremely important). The ranking scores were subsequently normalised, to a scale of 0 – 1 and applied in such a way that the weights added together up to unity (Table 1). From Table 1 it can be seen that the buyer perceived the bid amount and technical ability to be the most important criterion for evaluating the fabrication options. The lowest weighting (with reputation) went to the costs of the quality system; this reflects the need to meet buyer quality accreditation for inclusion on the supplier’s list. Therefore any uncertainty concerning quality as an order-winning criteria presented in Table 1 is reduced a to minimum level. The criteria of financial capacity and reputation is to some extent a reflection of perceived risk, and the weighting of both is probably due to the mature nature of the fabrication market. All the bidders are quite established and are well known to the buyers. There is little risk attached to the bidders, respectively in the absence of any new entrants. Each of these five main criteria is subsequently broken down into sub-criteria to give a total of 19 criteria (Table 1). The relative weights of the sub-criteria are then applied using the same procedure as the main criteria and presented in Table 2. It may be noted that the sum of sub-criteria weights, in respect to the associated upper level criteria is equal to one.

Table 2 Scores of the five bidders for the complete set of criteria

[1] Bid amount (0.3)

Capital bid	Conditions of payment	Variation formulae	Finance proposal
(0.75)	(0.05)	(0.1)	(0.1)

[2] Financial soundness (0.2)

Financial stability	Credit rating	Bank arrangements	Financial status
(0.3)	(0.2)	(0.15)	(0.35)

[3] Quality (0.1)

Prevention Costs	Appraisal Costs	Internal costs
(0.4)	(0.3)	(0.3)

[4] Technical ability (0.3)

Scope of work	Plant/ Equipment	Personnel	Project Experience
(0.05)	(0.45)	(0.3)	(0.2)

[5] Reputation (0.1)

Past failures	Length of relations	Nature of relations	Quality of communication
(0.3)	(0.4)	(0.1)	(0.2)

4. An Application

There is no common method of assessing the 19 sub-criteria in practice. These may be divided into two types, quantitative and qualitative one's (Table 3), although recently attempts have been made to try and use numbers to measure all the criteria [19, 20]. Most however are intangible and involve some degree of subjective assessment. The qualitatively measured variables of technical ability, financial capacity and reputation were all measured using a point scoring system: 0 – 4 = very poor, 5 – 8 = poor, 9 –12 = average, 13 – 16 = good and 17 – 20 = excellent

Table 3 Classification of sub-criteria

Qualitative criteria	Quantitative criteria
Scope of work	Capital bid
Plant and equipment	Conditions of payment
Personnel	Variation formulae
Project experience	Finance proposal
Past failures	Length of relations
Financial stability	Prevention costs
Credit rating	Appraisal costs
Bank arrangements	Internal failure costs
Financial status	
Nature of relations	
Quality of communication	

Some of the criteria are negatively oriented in terms of desirability. An example of this is the 'past failures' criterion. For ease of comparison and to make the scoring assessment consistent for all criterion, the scores in these criteria were deducted from 20. Let us assume that bidder A has at a score of five, only s few past failures. This implies that the score for bidder A is converted to 15 (20 – 5). Therefore higher scores consistently indicate better bidders for all criteria The only exception to this is the bid price criterion. This is also negatively oriented, as lower bids are more desirable than higher bids, but no change is made to the values submitted by the bidders.

In response to the client’s concerns for maintaining confidentiality, the financial values of each quantitative sub-criteria were normalised to a ratio value of the main criteria. For instance, if one takes the criteria of bid amount, then the ratio scores calculated for each of sub-criteria included the following: 0.01 is poor, 0.1 is below average, 0.2 is average, 0.3 is good and 0.4 is excellent. Likewise, the ratios were used to estimate the costs of quality control and included the following: 0.2 is poor, 0.4 is below average, 0.6 is good and 1 is excellent.

5. Combined Assessment of the Criteria

The first stage involved an original assessment for each contractor on the respective criteria and this is presented in Table 4. These evaluations were entered into the IDS for combined assessment and the results of this are presented in Table 5. A graphical distributed assessment was done for each criteria, in order to aid explanation of these results. In Figure 1, it can be seen that contractor A is the highest scoring, among the options on financial soundness. This reflects a best score assigned to the excellent grade. The lowest

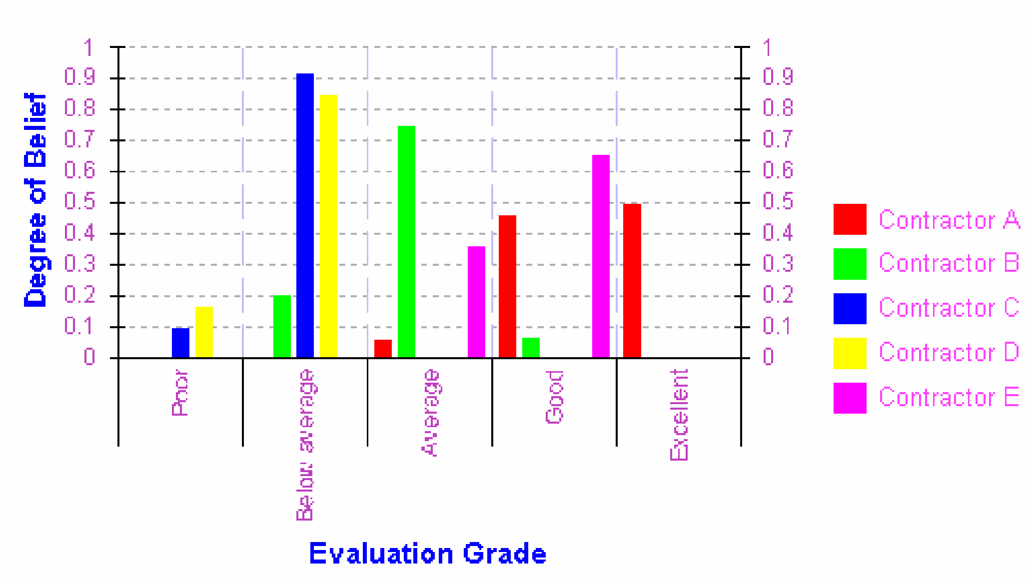

Figure 1 Distributed assessment on financial capacity

scoring on this criteria is D, which is primarily due it being scored high on below average and one of only two (contractor C) to be given an evaluation of poor.

.

Table 4 The original assessment of the criteria

BID AMOUNT					
Capital Bid	(5/5, poor)	(3/5, poor) (2/5, below av.)	(3/5, poor) (2/5, below av.)	(5/5, poor)	(5/5, poor)
Conditions of Payment	(5/5, exc't.)	(5/5, exc't)	(5/5, exc't)	(5/5, exc't)	(5/5, exc't)
Variation Formulae	(2/5, poor) (3/5, below av.)	(2/5, poor) (3/5, below av.)	(2/5, poor) (3/5, below av.)	(2/5, poor) (3/5, below av.)	(5/5, poor)
Finance Proposal	(2/5, poor) (3/5, below av.)	(2/5, poor) (3/5, below av.)	(4/5, poor) (1/5, below av.)	(2/5, poor) (3/5, below av.)	(2/5, poor) (3/5, below av.)
QUALITY SYSTEM					
Prevention Costs	(5/5, poor)	(5/5, poor)	(2/5, poor) (3/5, below av.)	(5/5, poor)	(1/5, poor) (4/5, below av.)
Appraisal Costs	(5/5, poor)	(1/5, poor) (4/5, below av.)	(5/5, below av.)	(5/5, below av.)	(1/5, poor) (4/5, below av.)
Internal Failure Costs	(5/5, av.)	(3/5, av.) (2/5, good)	(5/5 av.)	(2/5, av.) (3/5, good))	(5/5, av.)
FINANCIAL CAPACITY					
Financial Stability	(1/5, av.) (4/5, good)	(3/5, below av.) (2/5, av.)	(1/5, poor) (4/5, below av.)	(1/5, poor) (4/5, below av.)	(4/5, av.) (1/5, good)
Credit Rating	(4/5, good) (1/5, exc.)	(5/5, av.)	(1/5, poor) (4/5, below av.)	(1/5, poor) (4/5, below av.)	(3/5, av.) (2/5, good)
Bank Arrangements	(2/5, good) (3/5, exc.)	(2/5, below av.) (3/5, av.)	(1/5, poor) (4/5, below av.)	(1/5, poor) (4/5, below av.)	(1/5, av.) (4/5, good)
Financial Status	(5/5, exc.)	(4/5, av.) (1/5, good)	(5/5, below av.)	(1/5, poor) (4/5, below av.)	(5/5, good)
REPUTATION					
Past Failures	15	17	7	13	15
Nature of Relations	(1/5, av.) (4/5, good)	(5/5, good)	(1/5, below av.) (4/5, av.)	(1/5, av.) (4/5, good)	(3/5, below av.) (2/5, av.)
Length of Relations	12	13	9	13	11
Quality of Communications	(4/5, av.) (1/5, good)	(3/5, av.) (2/5, good)	(4/5, av.) (1/5, good)	(3/5, av) (2/5, good)	(3/5, below av.) (2/5, good)
TECHNICAL ABILITY					
Scope of Work	(3/5, good) (2/5, exc.)	(2/5, good) (3/5, exc.)	(4/5, av.) (1/5, good)	(1/5, poor) (4/5, below av.)	(5/5, below av.)
Personnel	(3/5, av.) (2/5, good)	(3/5, good) (2/5, exc.)	(4/5, good) (1/5, exc.)	(3/5, poor) (2/5, av.)	(5/5, below av.)
Plant & Equipment	(1/5, av.) (4/5, good)	(3/5, good) (2/5, exc)	(2/5, av.) (3/5, good)	(5/5, below av.)	(1/5, poor) (4/5, below av.)
Project Experience	11	13	16	3	5

Table 5 The combined assessment of the criteria

		A	B	C	D	E
Commercial	Combined Assessment	(0.946, poor) (0.039, below av) (0.014, excellent)	(0.569, poor) (0.416, below av) (0.014, excellent)	(0.601, poor) (0.385, below av) (0.014, excellent)	(0.946, poor) (0.039, below av) (0.015, excellent)	(0.968, poor) (0.018, below av) (0.014, excellent)
	Score	0.0193	0.0532	0.0503	0.0193	0.0171
Quality	Combined Assessment	(0.763, poor) (0.236, av.)	(0.512, poor) (0.216, below av) (0.162, av.) (0.108, good)	(0.157, poor) (0.59, below av) (0.253, av.)	(0.438, poor) (0.281, below av) (0.113, av.) (0.169, good)	(0.134, poor) (0.616, below av) (0.249, av.)
	Score	0.2947	0.3734	0.4191	0.4025	0.4231
Financial soundness	Combined Assessment	(0.053, av.) (0.454, good) (0.493, excellent)	(0.195, below av) (0.743, av.) (0.060, good)	(0.09, poor) (0.91, below av)	(0.159, poor) (0.841, below av)	(0.353, av.) (0.647, good)
	Score	0.8877	0.5731	0.3821	0.3683	0.7294
Reputation	Combined Assessment	(0.275, below av) (0.289, av.) (0.437, good)	(0.108, poor) (0.162, below av) (0.127, av.) (0.603, good)	(0.087, below av) (0.738, av.) (0.175, good)	(0.156, below av) (0.36, av.) (0.484, good)	(0.707, below av) (0.281, av.) (0.012, good)
	Score	0.6324	0.6449	0.6177	0.6657	0.4611
Technical ability	Combined Assessment	(0, below av) (0.403, av.) (0.586, good) (0.011, excellent)	(0.053, av.) (0.63, good) (0.317, excellent)	(0.196, av.) (0.732, good) (0.072, excellent)	(0.204, poor) (0.796, below av)	(0.072, poor) (0.928, below av)
	Score	0.7217	0.8529	0.7751	0.3592	0.3856

The distributed assessment in Figure 2 shows a poor evaluation of the options on bid amount. Contractor B is the highest scoring; this is largely due to it scoring the best among the five on the below average evaluation grade, which is the main grade for distinguishing between them. By comparison, it can be seen from Figure 1, that E is the poorest scoring contractor

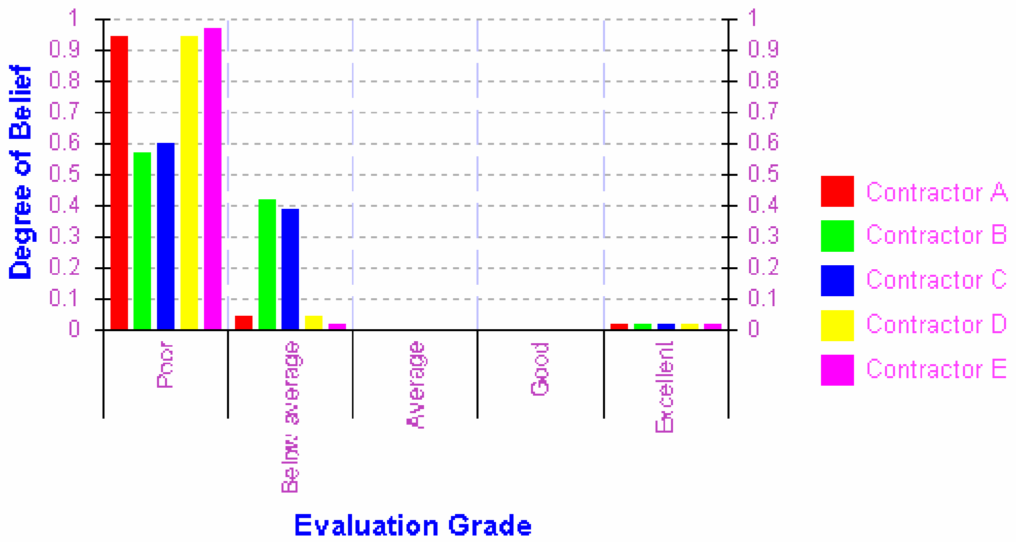

Figure 2 Distributed assessment on bid amount

Figure 3 indicates contractor D, to be the best one for financial reputation on the distributed assessment. The contractor is second best to B, on the evaluation grade of good, but it is B, who scores lower to it on the average grade. B is also evaluated poor, which D is clearly not because of it having the lowest evaluation grading, of average.

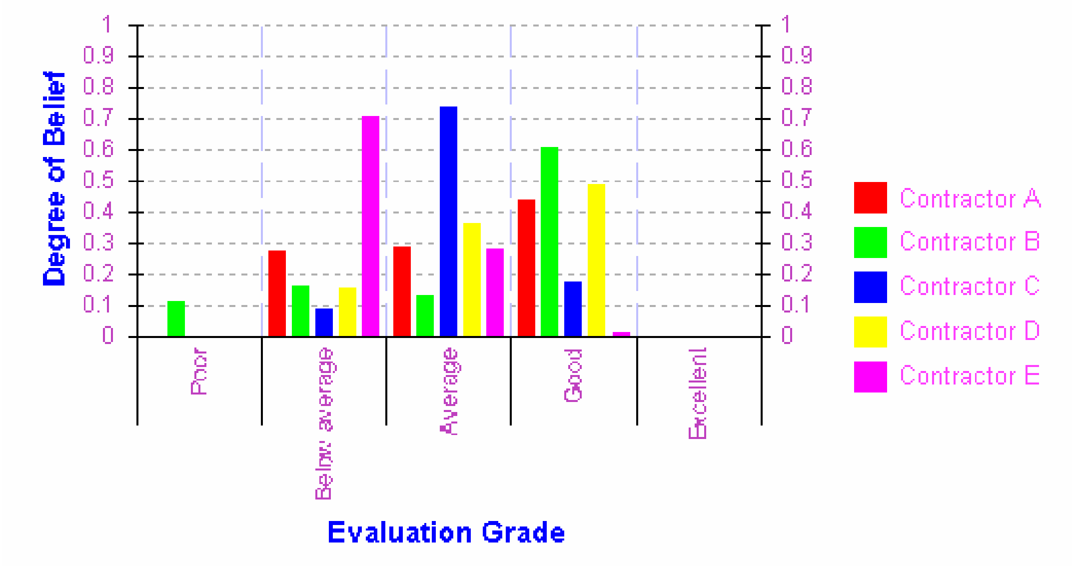

Figure 3 Distributed assessment on financial reputation

The highest score for technical ability goes to contractor B and the distributed assessment is presented in Figure 4. This contractor is assigned the best score on the highest evaluation grade of excellent. The lowest score is assigned to D, who is only either below average or poor on this criteria.

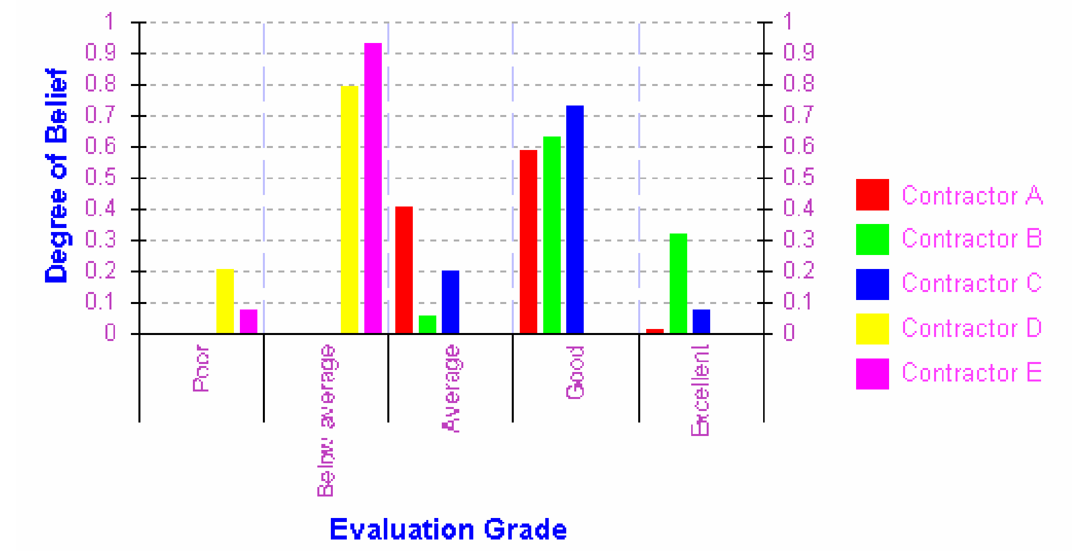

Figure 4 Distributed assessment on technical ability

Contractor E is scored the highest for its quality system. Both B and D have a good evaluation grade, which E does not, but of more significance is the poorer evaluations of there quality, given by some of the assessors with respect to E.

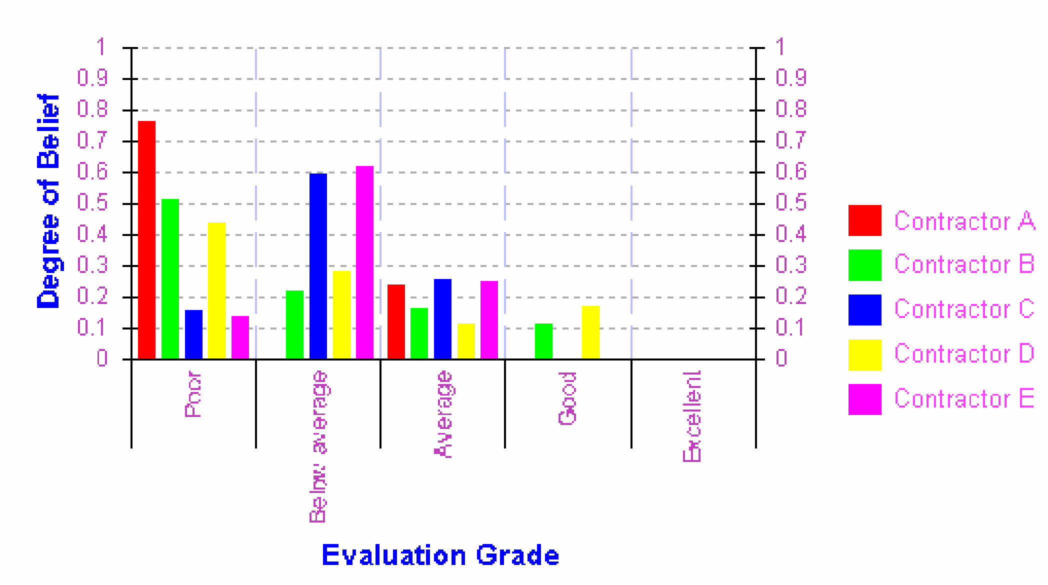

Figure 5 Distributed assessment on quality control

The overall assessment, score and ranking for every contractor are shown in Table 6. In order to illustrate this an overall distributed assessment is presented in Figure 6.

Table 6 The overall ranking for each contractor

	A	B	C	D	E
Overall Assessment	(0.386, poor) (0.033, below av) (0.181, av.) (0.323, good) (0.095, excellent)	(0.235, poor) (0.205, below av) (0.179 , av) (0.279, good) (0.102, excellent)	(0.216, poor) (0.337, below av.) (0.144, av) (0.237, good) (0.026, excellent)	(0.435, poor) (0.478, below av) (0.035, av.) (0.048, good) (0.004, excellent)	(0.337, poor) (0.437 below av.) (0.108, av.) (0.114, good) (0.004, excellent)
Score	0.5488	0.5614	0.4962	0.3415	0.4025
Ranking	2	1	3	5	4

The table indicates and the figure highlights the following: B>A>C>E>D. The main reason for this is that B has the highest scores for the two most important criteria; this is bid amount and technical ability. Contractor A is a reliable contractor with a good reputation, who suffers from being not very competitive on bid amount and technical ability. C, D and E are seen to be risky options to be awarded the contract.

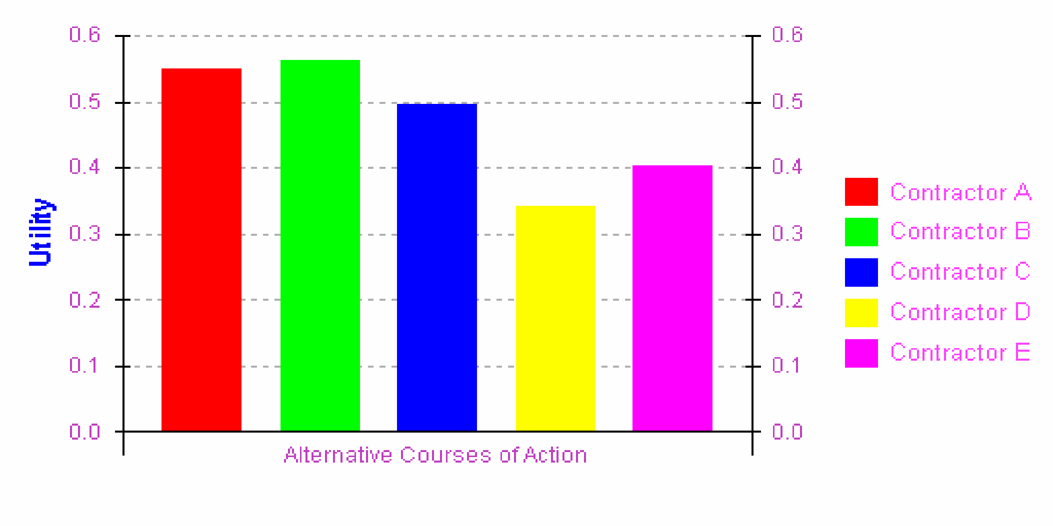

Figure 6 Overall ranking for the contractors

6. Summary

There is a need for a contractor evaluation technique that is capable of considering multiple criteria. Evidential reasoning provides one such approach and is especially useful as it allows the treatment of both quantitative and qualitative criteria and in situations where there are several stake-holders; this is chosen for its simplicity, practicality and appropriateness in risky choice situations. The individual importance of each contractor criterion is specified using a weighting which also incorporates the risk of the decision-maker. A hypothetical case study is described to illustrate the method and includes the results from real interviews with four leading assessors, who are involved in contractor evaluation. The precise assessments of the relative weights was shown to have a crucial bearing on the solution. Evidential reasoning is a technique for use in evaluation decisions where criteria are of different characteristics and it appears to be eminently suited to aerospace contractor evaluation and selection.

REFERENCES

1. Cova, B and Holstius, K (1993) 'How to Create Competitive Advantage in Business', Journal of Marketing Management, 9, 105-121.

2. Bonnacorsi, A and Paliwoda, S (1994) 'Trends in Procurement Strategies within the European Aircraft Industry', Industrial Marketing Management, 23, 235-244.

3. Turner, J.A. (1993) The Handbook of Project-Based Management. McGraw Hill, London.

4. GC/Stores/1 (1979) Standard Conditions of Government Contracts for Stores Purchases, HMSO, London.

5. MOD (1983) Value for Money in Defence Equipment Procurement, 5 Defence Open Government Document 83/01, MOD, London.

6. Taylor, T and Hayward, K (1989) The UK defence Industrial Base: Development and Future Policy Options. Brassey's Defence Publishers, London.

7. Cahill, F (1993) Defence Diversification. Unpublished MBA thesis. Surrey European Management School, University of Surrey, Guildford.

8. Hartley, K and Martin, S (1993) Evaluating Collaborative Programmes, Defence Economics, 4 (2), 19-33.

9. Schofield, S (1995) The Levene Reforms - An Evaluation. Discussion Paper E95/02, School of Business and Economic Studies, University of Leeds.

10 Latham, A (1995) The Structural Transformation of the US Defence Firm: Changes in Manufacturing Technology, Production Process and Principles of Corporate Organisation in Latham, A and Hooper, N The Future of the Defence Firm: New Challenges, New Directions, 175-191. Kluwer Academic Press, Dordrecht, Netherlands.

11. Sen, P and Yang, J (1995) Multiple Criteria Decision Making in Design Selection and Synthesis, Journal of Engineering Design, 6, 207-230.

12. Huang, C L and Yoon K (1981) Multiple Attribute Decision Making Methods and Applications, A State-of-Art Survey (Berlin, Springer-Verlag).

13. Perry, C and Grace, S., (1997) Applying Multiple Criteria Decision Making Techniques to Tender Evaluation in the Australian Electric Industry, with Particular Reference to Austria Electric, Report of an ARC Collaborative Grant Research Project. Faculty of Business, Queensland University of Technology.

14. Hatush, Z and Skitmore, M. (1998) Contractor Selection Using Multicriteria Utillity Theory: An Additive Model. Building and Economics, Vol 33 Nos 2 –3, pp. 105-115.

15. Saaty, T L (1988) The Analytical Hierarchy Process (University of Pittsburgh).

16. Ellis, R and Herbsman, Z (1991) Cost-time Bidding Concept: an Innovative Approach. Transportation Research Record 1282, Washington D.C,., 89-94.

17. Holt, G., Olomolaiye, P and Harris, F., (1993) A Conceptual Alternative to Current Tendering Practice, Building Research and Information, 21(3), 167-172.

18. Yang, J and Sen, P (1997) Multiple Attribute Design Evaluation of Complex Engineering Products Using the Evidential Reasoning Approach. Journal of Engineering Design, Vol 8, No 3, 211-230.

19. Yang, J B and Sen, P (1994) A General Multi-level Evaluation Process for Hybrid MADM with Uncertainty, IEEE Transactions on Systems, Man and Cybernetics, 24, pp. 1-18.

20. Yang, J B and Singh, M G (1994) An Evidential Reasoning Approach for Multiple Attribute Decision Making with Uncertainty, IEEE Transactions on Systems, Man and Cybernetics, 24, pp. 1-18.

21. Evans, J P (1993) Techno-economic Analysis of Retro-fit Options. Research Report, Transmarine Ltd, Wallsend Research Station, Wallsend, Tyne and Wear, UK.