Contractor Evaluation: Using An Evidential Reasoning Approach

162 - 173

Contractor Evaluation in the Aerospace Industry Using the Evidential Reasoning Approach

Gary Graham
University of Salford, Salford, England
Glenn Hardaker
Huddersfield University Business School, England

Abstract: 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. 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 for aero-engine equipment is then examined using this approach. Given the role of small firms as sub-contractors, and the difficulties that they face in managing the pricing part of the marketing mix, which is heightened when tendering is involved, this paper has lessons that extend outside of the particular industry that forms the case study. The case study itself draws from the real life experience of actual defence contract assessors.

1. INTRODUCTION

Tendering is an approach largely associated with the award of project contracts, in the aerospace sector (Cova & Holistius, 1993). 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. Projects (and thus Project Management) aim to deliver beneficial, one-off transformational changes not achieved through improvements in the efficiency of existing operations (Bonaccorsi & Paliwoda, 1994). 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 (Turner, 1993). Despite this, however, there has been no commensurate improvement in the performance of aerospace project management. Instead, there have been extensive delays in planned schedules, cost overruns, and an increased number of claims and litigation.

The standard conditions for awarding contracts in the aerospace industry have remained relatively unchanged for nearly twenty years (GC/Stores/1, 1979; MOD, 1983). Many now believe, that the prime contractor approaches to contractor evaluation, concentrating as it does solely on bid price (Taylor & Hayward, 1989; Cahill, 1993), is one of the major causes of project delivery problems (Hartley & Martin, 1993; Schofield, 1995; Latham, 1995). This leads to suppliers who 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 (MOD, 1983).

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 (Sen & Yang, 1995). There are a number of theoretical drawbacks of this simple weighted approach (Huang & Yoon, 1981), 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 multi-criteria decision analysis methods (see e.g. Perry & Grace, 1997, for a general overview of this well established technique; and Hatush & Skitmore, 1998, 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 (see, for example: Sen and Yang, 1995; Hatush & Skitmore, 1998); and Saaty, 1988). By far, the main difficulty encountered with multi-criteria methods is trying to compare different criteria, which have been measured on different scales. Ellis and Herbsman (1991) 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 (1993) combine 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°/o 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 evidential approach to treat the aforementioned contractor selection problem (Yang & Sen, 1997). 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 (Perry & Grace, 1997; Hatush & Skitmore, 1998). The main advantages of this approach, if compared with other multi-criteria 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 use 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 (Huang, 1981; Evans, 1993). A window-based intelligent decision system (IDS) using the evidential reasoning approach is used to support this research.

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 (Exhibit One).

Exhibit One:
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 all located within the buying organisation,* by firstly ranking criteria for five potential sub-contractor bidders in the order of relative importance using a Likert scale from 1 (unimportant) to 7 (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 in Exhibit One. From this Exhibit 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 (jointly with reputation) went to the costs of the quality system; this reflects that sub-contractors need to meet pre-determined buyer quality accreditation in order to be included on the (approved) supplier list. Therefore any uncertainty concerning quality as an order-winning criteria presented in Exhibit One is reduced to a minimum level. The criteria of financial capacity and reputation is to some extent a reflection of perceived risk, and he 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 in Exhibit One are subsequently broken down into sub-criteria to give a total of nineteen criteria. The relative weights of the sub-criteria are then applied using the same procedure as the main criteria and presented in Exhibit Two. It may be noted that the sum of sub-criteria weights, in respect to the associated upper level criteria is equal to one.

Exhibit Two:
Scores of the five Bidders for the Complete Set of Criteria

Bid amount	**(0.3)**						
Capital bid	(0.75)	Conditions of payment	(0.05)	Variation formulae	(0.1)	Finance proposal	(0.1)
Financial soundness	**(0.2)**						
Financial stability	(0.3)	Credit rating	(0.2)	Bank arrangements	(0.15)	Financial status	(0.35)
Quality	**(0.1)**						
	(0.4)	Appraisal Costs	(0.3)	Internal costs	(0.3)		
Technical ability	**(0.3)**						
Scope of work	(0.05)	Plant/ Equipment	(0.45)	Personnel	(0.3)	Project Experience	(0.2)
Reputation	**(0.1)**						
Past failures	(0.3)	Length of relations	(0.4)	Nature of relations	(0.1)	Quality of communication	(0.2)

4. AN APPLICATION

There is no common method of assessing the nineteen sub-criteria in practice. These may be divided into two types, quantitative and qualitative and this is illustrated in Exhibit Three, although recently attempts have been made to try and use numbers to measure all the criteria (Yang & Sen, 1994; Yang & Singh, 1994). However, most are intangible and involve some degree of subjective assessment. In this case study, 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.

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, indicating only a 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.

Exhibit Three:
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	

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 (Exhibit Four). These evaluations were entered into the IDS for combined assessment and the results of this are presented in Exhibit Five. A graphical distributed assessment was done for each criteria, in order to aid explanation of these results. From Exhibit Six, 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 scoring on this criteria is D, which is primarily due it being scored high on 'below average' and one of only two (the other being, contractor C) to be given an evaluation of 'poor'. The distributed assessment on Bid Amount in Exhibit Seven 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 Exhibit Four, that E is the poorest scoring contractor Exhibit Eight 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 of average. The highest score for technical ability goes to contractor B and the distributed assessment is presented in Exhibit Nine. 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. 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 are the poorer evaluations of their quality, given by some of the assessors with respect to Contractor E.

The overall assessment, score and ranking for every contractor are shown in Exhibit Eleven. 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. Contractors C, D and E are seen to be risky options to be awarded the contract.

Exhibit Four:
The original assessment of the criteria

	Contractor A	Contractor B	Contractor C	Contractor D	Contractor E
BID AMOUNT (COMMERCIAL)					
Capital Bid	(5/5, poor)	(3/5, poor)	(3/5, poor)	(5/5, poor)	(5/5, poor)
		(2/5, below av.)	(2/5, below av.)		
Conditions of Payment	(5/5, excellent)	(5/5, excellent)	(5/5, excellent)	(5/5, excellent)	(5/5, excellent)
Variation Formulae	(2/5, poor)	(2/5, poor)	(2/5, poor)	(2/5, poor)	(5/5, poor)
	(3/5, below av.)	(3/5, below av.)	(3/5, below av.)	(3/5, below av.)	
Finance Proposal	(2/5, poor)	(2/5, poor)	(4/5, poor)	(2/5, poor)	(2/5, poor)
	(3/5, below av.)	(3/5, below av.)	(1/5, below av.)	(3/5, below av.)	(3/5, below av.)
QUALITY (SYSTEM)					
Prevention Costs	(5/5, poor)	(5/5, poor)	(2/5, poor)	(5/5, poor)	(1/5, poor)
			(3/5, below av.)		(4/5, below av.)
Appraisal Costs	(5/5, poor)	(1/5, poor)	(5/5, below av.)	(5/5, below av.)	(1/5, poor)
		(4/5, below av.)			(4/5, below av.)
Internal Failure Costs	(5/5, av.)	(3/5, av.)	(5/5 av.)	(2/5, av.)	(5/5, av.)
		(2/5, good)		(3/5, good)	
FINANCIAL CAPACITY (SOUNDNESS)					
Financial Stability	(1/5, av.)	(3/5, below av.)	(1 /5, poor)	(1/5, poor)	(4/5, av.)
	(4/5, good)	(2/5, av.)	(4/5, below av.)	(4/5, below av.)	(1/5, good)
Credit Rating	(4/5, good)	(5/5, av.)	(1/5, poor)	(1/5, poor)	(3/5, av.)
	(1/5, exc.)		(4/5, below av.)	(4/5, below av.)	(2/5, good)
Bank Arrangements	(2/5, good)	(2/5, below av.)	(1/5, poor)	(1/5, poor)	(1/5, av.)
	(3/5, exc.)	(3/5, av.)	(4/5, below av.)	(4/5, below av.)	(4/5, good)
Financial Status	(5/5, exc.)	(4/5, av.)	(5/5, below av.)	(1/5, poor)	(5/5, good)
		(1/5, good)		(4/5, below av.)	
REPUTATION					
Past Failures	15	17	7	13	15
Nature of Relations	(1/5, av.)	(5/5, good)	(1 /5, below av.)	(1/5, av.)	(3/5, below av.)
	(4/5, good)		(4/5, av.)	(4/5, good)	(2/5, av.)
Length of Relations(hip)	12	13	9	13	11
Quality of Communications	(4/5, av.)	(3/5, av.)	(4/5, av.)	(3/5, av.)	(3/5, below av.)
	(1/5, good)	(2/5, good)	(1/5, good)	(2/5, good)	(2/5, good)
TECHNICAL ABILITY					
Scope of Work	(3/5, good)	(2/5, good)	(4/5, av.)	(1/5, poor)	5/5, below av.)
	(2/5, exc.)	(3/5, exc.)	(1/5, good)	(4/5, below av.)	
Personnel	(3/5, av.)	(3/5, good)	(4/5, good)	(3/5, poor)	(5/5, below av.)
	(2/5, good)	(2/5, exc.)	(1/5, exc.)	(2/5, av.)	
Plant & Equipment	(1/5, av.)	(3/5, good)	(2/5, av.)	(5/5, below av.)	(1/5, poor)
	(4/5, good)	(2/5, exc.)	(3/5, good)		(4/5, below av.)
Project Experience	11	13	16	3	5

Exhibit Five:
The combined assessment of the criteria

		A	B	C	D	E
Bid Amount (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 (System)	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 Capacity (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**

Exhibit Six: Distributed Assessment on Financial Capacity

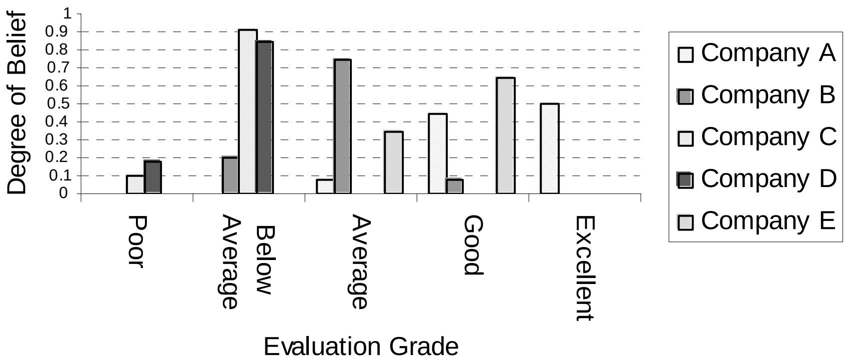


Exhibit Seven: Distributed Assessment on Bid Amount

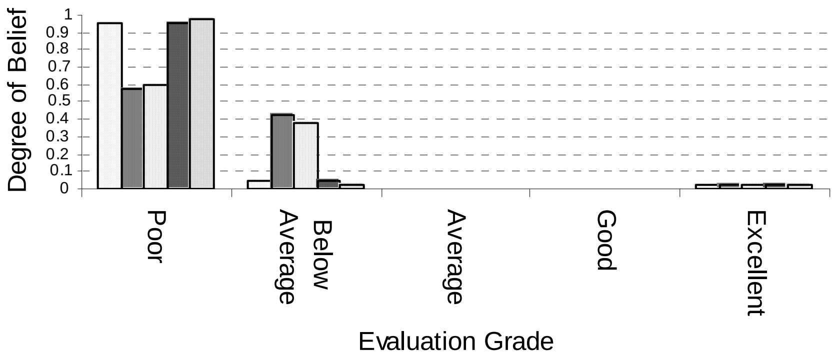


Exhibit Eight: Distributed Assessment on Financial Reputation

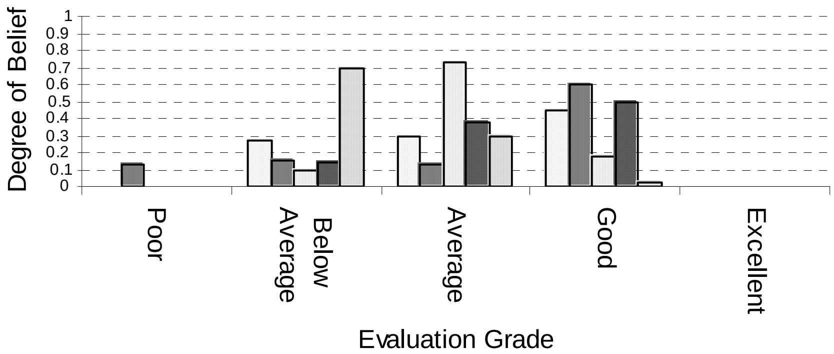


Exhibit Nine: Distributed Assessment on Technical Ability

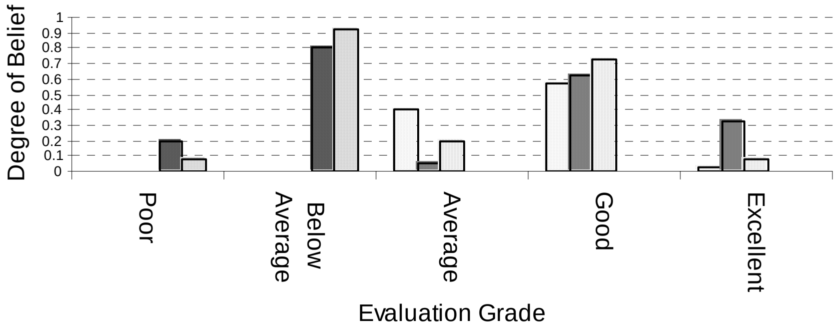


Exhibit Ten: Distributed Assessment on Quality Control

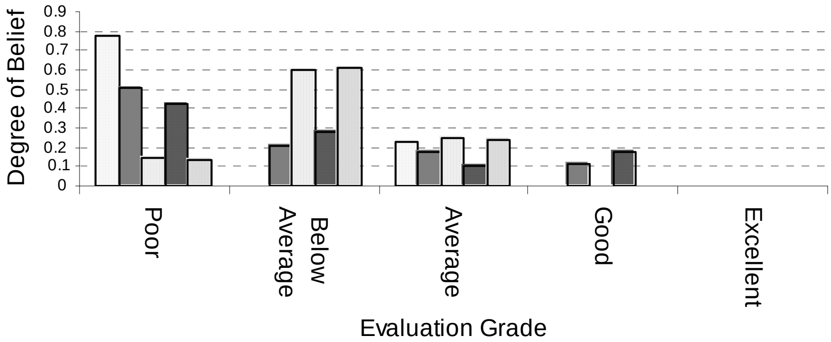
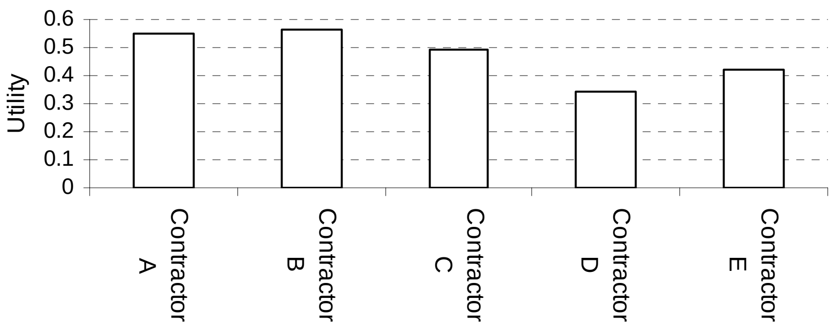


Exhibit Eleven: The overall ranking for each contractor

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

Exhibit Eleven: 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

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

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

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

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

Evans, J.P. (1993), *Techno-economic Analysis of Retro-fit Options*, Research Report, Transmarine Ltd, Wallsend, UK: Wallsend Research Station.

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

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

Hatush, Z. and Skitmore, M. (1998), *Contractor Selection Using Multicriteria Utility Theory: An Additive Model*, Building and Environment, Vol. 33 No. 2 -3, pp. 105-115.

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

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

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. (Editors), 175-191. Netherlands, Dordrecht: *The Future of the Defence Firm: New Challenges, New Directions,* Kluwer Academic Press.

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

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

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

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

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

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

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

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.

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.

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.