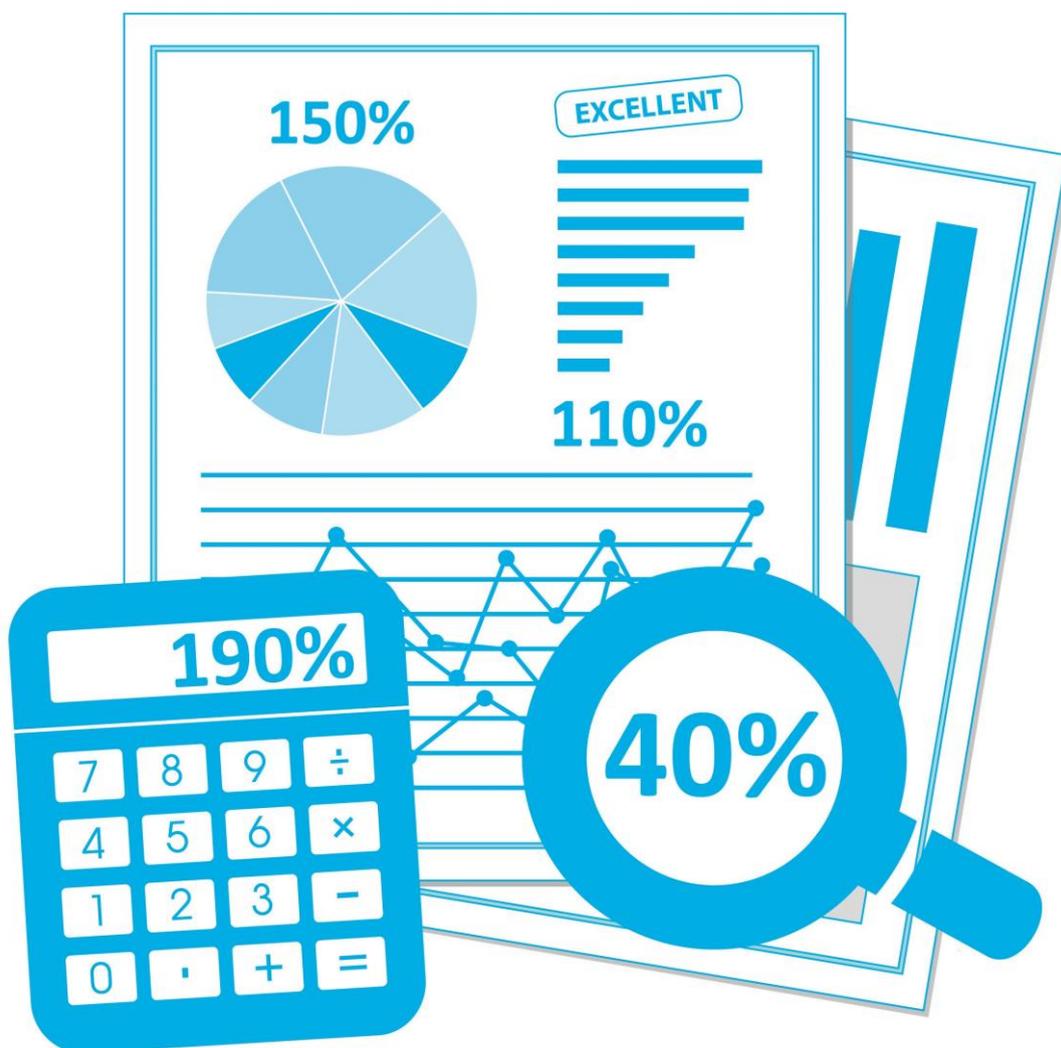


White paper

Weaknesses of traditional supplier evaluation methods - introducing Real Value for Money

Peter Marshall, Professional Services Director



Weaknesses of traditional supplier evaluation methods

Introduction

Commerce Decisions' professional services team has accrued nearly 15 years experience of working with Public Sector procurement teams across all sectors – military hardware, healthcare, education, construction, IT and facilities management to name a few.

As a result, we have had the privilege of becoming unrivalled specialists in the field of supplier selection and proposal evaluation. As our knowledge and experience has deepened, we have become increasingly concerned about the application of many common approaches to procurement and the techniques/methodology used to ensure value for money is achieved.

This white paper seeks to lay out our concerns with these frequently used methods in a structured way. It will also provide some useful pointers on how procurement teams can avoid the pitfalls inherent in these processes.

Commerce Decisions has developed an alternative approach to supplier evaluation that avoids the issues described in this paper; we call this method "Real Value for Money (RVfM)".

Unbalanced perception of risk

The first contributory factor that makes traditional methods problematic is simply that of time. Buyers do not perceive it as important to put sufficient time and effort into understanding and testing their proposed evaluation methods.

Most public-sector procurement teams see a legal challenge as the most significant risk to a successful procurement. This is closely followed by the risk of delays; caused by not releasing Pre-Qualification Questionnaires and Invitations To Tender at the time required to meet the overall procurement timelines. Although these risks are obviously important, with potentially serious consequences, there is another risk that is rarely recognised explicitly – the risk of failing to buy the best solution.

We see a common issue emerging – that of suppliers not being sufficiently engaged prior to running the procurement process. This often leads to solutions being bought which are not as good as they could be and are

bought in a way that is unnecessarily expensive.

In other words, the risks associated with *not* engaging with industry up-front are often not weighed up against the risk of engaging with industry before planning out the procurement. This results in little engagement with industry in order to minimise the only recognised risk – the risk of legal challenge or bidder protest.

Technical differentiation

We regularly hear clients complain of the difficulty in justifying one bid over another on the basis of their greater technical capability, flexibility, innovation or expertise. Clients often find that the best suppliers are not scoring sufficiently highly by comparison with their competitors. All too often many suppliers are achieving full marks or scoring very similar technical scores (within a few percentage points).

Fundamentally, this is a flaw in the assessment scheme and the evaluation criteria being used. Most evaluation criteria we see concentrate on meeting the technical requirements of the solution being sought, giving little consideration to the risks (eg. availability of teams, poorly articulated requirements, changing landscapes) and dependencies (eg. buy-in, other partners, agreements, IPR). In our experience these are the main reasons many strategic projects fail to deliver the project outcomes on time and to budget.

Less often still, and much bemoaned in the press, is a consideration of the overall value proposition of a bid to the Authority or to the wider economy. Often there are bids which will provide far more value than their competitors, the scope of which was never considered at the outset of the procurement (job creation, economies of scale etc).

Building an assessment process which differentiates the suppliers on the basis of their ability to deliver the best value is not simple. Commerce Decisions has spent over 12 years developing our SCD methodology – this is the subject of a separate whitepaper "Structured Criteria Development - getting procurement right". *A copy can be downloaded from the white papers section of our website: cd.QinetiQ.com.*

Buying the best solution

The purpose of engaging with industry up-front is firstly to confirm that the solution being specified is the one that best meets the needs of the buyer but also to ensure that the procurement is being conducted in the most cost-effective way. Having the opportunity to test - with suppliers - the suitability of the specification for the goods or services being bought is useful. But even more important in ensuring value for money is ascertaining that the buyer understands the major cost drivers of the supply base so that they can make an informed decision when specifying the precise outputs or requirements against which the bidders will respond.

It is not difficult to find examples of procurements where suppliers express frustration in not being able to offer better or more cost-effective solutions because of constraints imposed by specific requirements or evaluation criteria.

A recent example experienced by the author was an ITT that asked for details of the scalability of a proposed IT solution (ie. the ability to cope with a large increase in the number of users). The question was worded in such a way as to imply that bidders had to have a solution which was *already* able to cope with a 2000% increase in the number of users over three years. There was no provision for accepting a bid with a solution that could have dealt with the planned user base at contract award along with a robust, resourced plan for making the solution capable of dealing with the increasing user base over three years.

The ITT could have been deliberately worded in this way to mitigate the risk associated with buying a solution that could deal with the increasing user base. However, the author considers it more likely that in this case the buyer simply had not considered the possibility that they may get a more cost effective solution if they opened the competition to providers who were able to meet the immediate need and work with the buyer to meet the evolving need with product enhancements.

The difficulty of analysis

The second contributory factor making traditional methods problematic is the

Weaknesses of traditional supplier evaluation methods

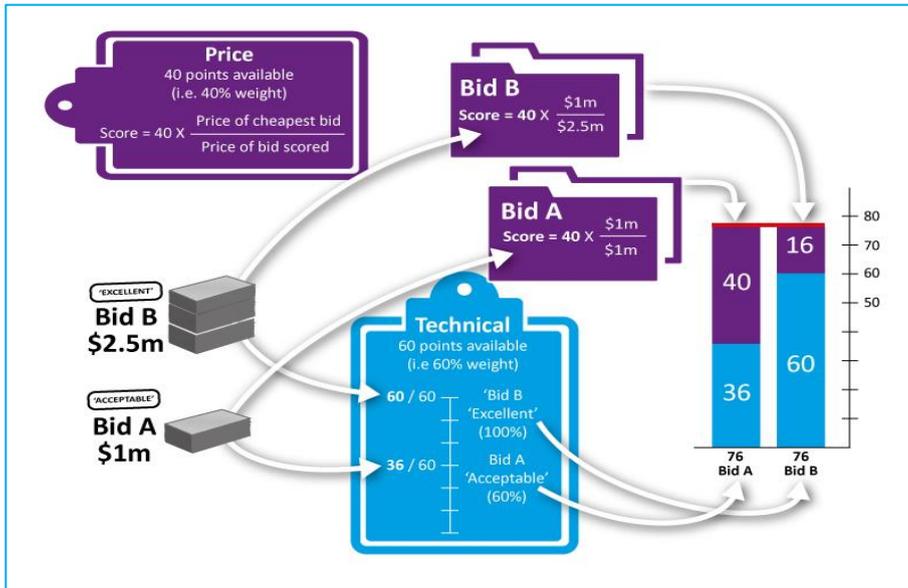


Figure 1: Two very different bids score the same

Figure 2:

$$\text{Score} = \text{Points available} \times \frac{\text{Price of cheapest bid}}{\text{Price of bid being scored}}$$

difficulty of fully understanding the implications of the maths behind the methods.

It is not an exaggeration to say that virtually all of the traditional methods of evaluating suppliers' bids are undertaken by determining a score for price and non-price (technical) criteria, and bringing these scores together (after applying weights) to calculate an overall score.

The overall effect of the weighting of the price and technical sections of a tender is rarely understood by procurement teams. Figure 1 provides an illustration of the effect of a typical weighting and scoring mechanism.

The three main components of the weighting and scoring mechanism illustrated in Figure 1 are chosen as typical examples of common practice. Price is weighted 40%, technical weighted 60%: in our experience this is the most common weighting used. Interestingly, the most common justification for this weighting is "this is what we did last time", rather than the result of analysis of its effect.

The prices are converted into scores by using the formula illustrated in Figure 2; the cheapest bid is given full marks and other bids are given a mark

proportional to how much more expensive they are; bids that are twice as expensive are given half the marks of the cheapest and so on. This is the most common process used for converting prices into scores that we come across.

Finally, the scoring of the technical criteria is done using the scoring scale illustrated where "acceptable" results in a score of 60% of the marks available and all of the marks available are given to bids judged "excellent". Again, in our experience, this is a typical scoring scheme used for the technical section of a bid.

The three main components of the scoring scheme are often defined by different people at different points in time, and without reference to each other. The weight of price is often decided by a limited group of project stakeholders; the formula for converting prices into scores is often determined by the chief commercial officer; and the technical scoring scheme is often defined by the people developing the technical criteria. All too often, the interaction between these components is ignored and therefore not fully understood.

In the example illustrated at Figure 1, it can be seen that an "acceptable" bid (Bid A) costing \$1M would achieve the same overall score as an "excellent" bid (Bid B) costing \$2.5M. In other words, by implementing the weighting and scoring scheme shown, the buyer is effectively saying "we are willing to pay up to two and a half times as much for an 'excellent' bid compared to an 'acceptable' bid."

"We are willing to pay up to two and a half times as much for an 'excellent' bid compared to an 'acceptable' bid."

This defies common sense. In our experience, when buying teams are asked how much more they would consider paying for an "excellent" bid compared to an "acceptable" one, they would typically conclude between 20-50% more.

The point is that buying teams very rarely do the analysis to work out what effect their proposed weighting and scoring scheme will have. So they often do not understand its overall impact.

Indeterminate outcomes

In order to directly combine price and technical scores into an overall score, the prices received from each bidder have first to be converted into individual scores. There are a number of techniques traditionally used to do this, and they all have effects that are rarely understood by their users.

Most methods for converting prices into scores compare each price to either the lowest price, or possibly to the average price, giving a score that represents the difference. For example, possibly the most common method is to use the formula shown in Figure 2 that gives full marks to the cheapest bid and a pro-rata score to other bids. So, a bid that is twice as expensive as the cheapest bid would score half of the points available for price.

The problem with the methods that calculate a score for the price of a tender based on the prices of the other tenders is that this makes the ranking of the received tenders unreliable.

For an example of this see Figure 3 which shows the rankings of two bids changing - dependent solely on whether a third bid is submitted or

Weaknesses of traditional supplier evaluation methods

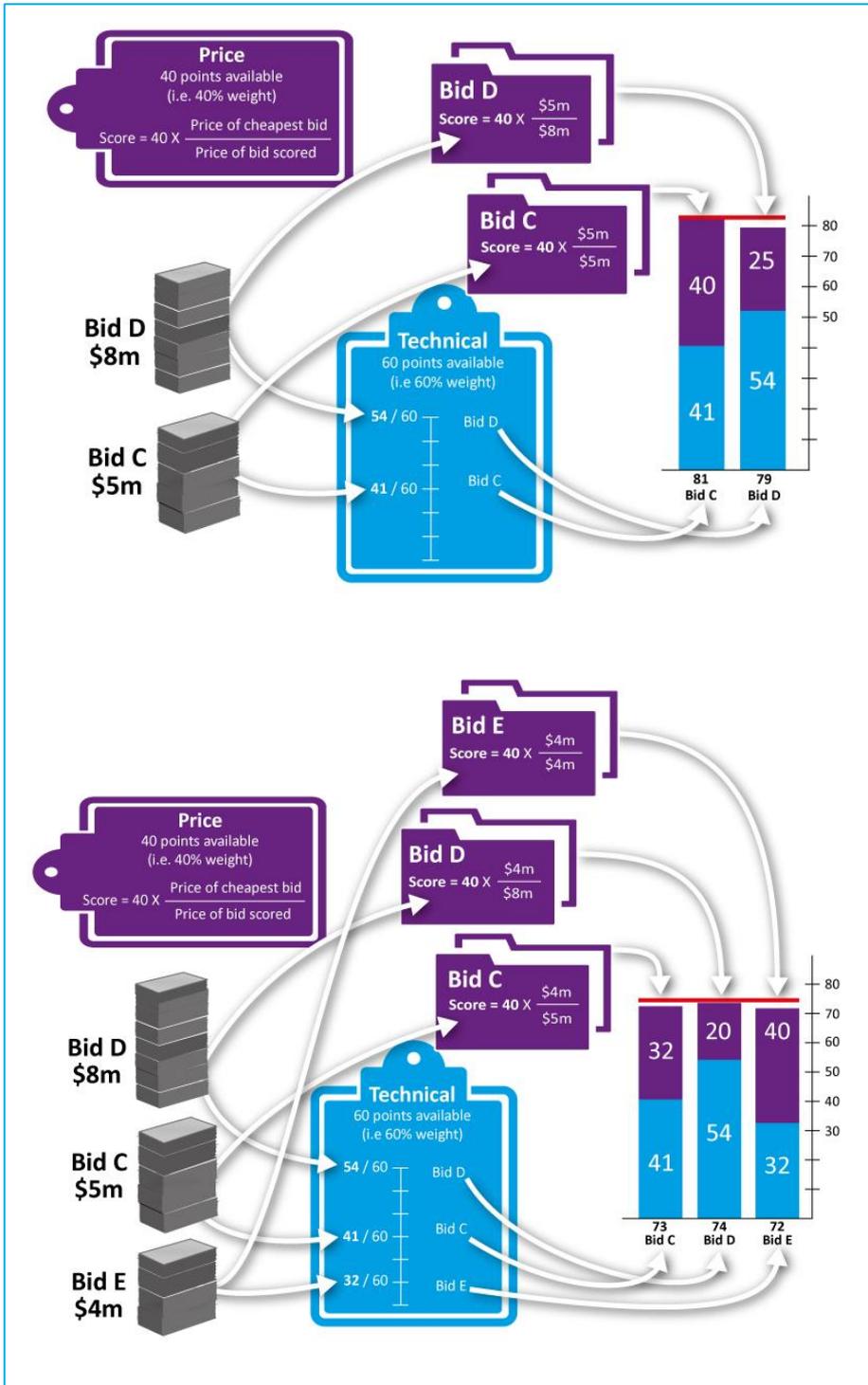


Figure 3: Bid C scores more than Bid D to win the competition. However, the introduction of a third bidder (Bid E) results in Bid D scoring more than Bid C (below).

not. This example makes it difficult to justify that this scoring method does a good job of determining which bid offers better value for money. The method “changes its mind” between which of the two bids offers best value for money without rational reason.

In turn, this gives significant cause for concern regarding the likelihood of protest and how easy it would be to

defend the use of the method should such a challenge be raised. In the example, were Bid C to realise that the scoring scheme meant they would have beaten Bid D if Bid E had never been received, or deemed non-conformant, the author believes that it would be very likely that the legal advice would be to challenge the result, with the consequent delay and risk of financial redress to the buyer.

When will the first challenge come?

At this point it would seem reasonable to ask why we have not seen this type of protest or legal challenge come through the courts, particularly as these scoring schemes have been used over many years with such a fundamental flaw.

The answer lies partly in the lack of information available to each bidder.

It is generally accepted that buyers are required to provide each losing bidder with their scores and the scores of the winning bid, but not necessarily the prices of the other bids. In order to do the full analysis on the effect of the scoring scheme, it is necessary to know all the technical scores and raw prices for the bids. This information is nearly always withheld on the grounds of being considered commercial-in-confidence information.

The answer may also lie partly in the fact that there is little, if any, awareness of these sorts of effects in common scoring schemes.

Buyers and bidders alike, unaware of the theoretical flaws with common scoring schemes, do not think to undertake the analysis required to fully understand them. Moreover, they do not undertake any detailed analysis on individual procurement outcomes to test if the flaws have manifested themselves in practice.

“Buyers who wish to avoid being the defendant on the next case to establish procurement case law should take note”.

However, in our view it is only a matter of time before awareness is raised sufficiently for bidders to start seeking the information they need in order to ascertain that their bid was scored and ranked appropriately. Buyers who wish to avoid being the defendant on the next case to establish procurement case law should take note.

As clear as mud

The good news is that bidders don't fully understand the methods either.

For the reasons discussed above, bidders do not have the information they need to understand how they will be scored in practice. They do not know how their competitors will be pricing their bids and therefore they

Weaknesses of traditional supplier evaluation methods

cannot work out how their own prices will be converted into scores. In turn, this means they cannot be certain how their technical solution and pricing combine to determine an overall score.

Consequently bidders are unable to “game” their responses, and are unlikely to be able to do the analysis necessary to work out how the ranking of their bid is affected by the submission of other bids.

Ignoring ‘Value for Money’

The bad news is that the buyer is very unlikely to get each bidder’s optimal solution.

If suppliers do not understand how their bids will be scored they cannot possibly make informed decisions about how to optimise their solutions. They cannot know if adding something into their bid will result in a higher score – because the additional cost is outweighed by the additional technical benefit – or a lower score because the extra cost outweighs the technical score increase.

Bidders are left trying to come up with the best solution they can offer within the budget they think the buyer has. Sometimes, the buyers will publish the budget, in which case the buyer will get the best solutions that the bidders know the buyer can afford.

Occasionally, a bidder will gamble that by putting in a low-ball bid they stand a good chance of winning by making the more capable competition too expensive.

At no point are bidders making an informed decision about what they include and exclude from their bid based on their robust understanding of what the customer values.

The point is that for bidders this is all guesswork. At no point are bidders making an informed decision about what they include and exclude from their bid based on their robust understanding of what the customer values. Surely the point of articulating the method is to allow suppliers to understand the relative importance of price and technical considerations and hence the basis for the buyer getting value for money?

A way out

Commerce Decisions has an alternative approach which we term **Real Value for Money (RVfM)**. RVfM overcomes the shortfalls highlighted in this paper by starting from a very different place to traditional methods.

All traditional methods combine the technical and price scores into an overall score by converting the prices into scores and then combining these with the technical scores using a weighting. As we have seen, the decisions made about how the prices will be converted and the weights will be applied have potentially unexpected effects that can only be understood after detailed analysis. In short, until we do this analysis, we don’t know how much we could end up paying for what.

RVfM starts by defining exactly this - how much we are willing to pay and for what solution.

The single question that needs to be answered when using RVfM is “how much are we willing to pay for each point on the technical scoring scale?” The answer to this question is then used to define a “Value for Money gradient” that defines what we mean by equal value for money. From this it is possible to compare any bids and determine their *real value* for money. In other words, how much better or worse value for money is offered by any bid in relation to another.

It provides a robust way of ranking each and every bid that does not change depending on what other bids are submitted. It provides a robust method for buyers to define what they mean by value for money, without the requirement for a great deal of analysis to test and get right. It also provides buyers with robust techniques for undertaking sensitivity analysis and testing of their technical criteria and weights.

Lastly, RVfM provides bidders with a transparent mechanism for the scoring of bids and, most importantly, it enables each bidder to optimise their solution to achieve the highest value for money possible.

Conclusions

Traditional techniques for scoring tenders have significant flaws that not only make it difficult (if not impossible)

to reliably achieve value for money when selecting suppliers; but also jeopardise the procurement process by introducing opportunity for legal challenge.

Procurement teams can mitigate these risks by undertaking suitable analysis and testing when preparing their scoring methodology to ensure it has the effect they expect.

Alternatively, procurement teams can make use of alternative methods like RVfM from Commerce Decisions, to provide a more robust and transparent supplier selection mechanism.

Commerce Decisions’ White Paper ‘Measuring Value for Money in a tender evaluation’ covers our VfM methodology in greater detail. *A copy can be downloaded from the white papers section of our website: cd.QinetiQ.com.*

Weaknesses of traditional supplier evaluation methods

About Peter Marshall

Peter Marshall is Professional Services Director, responsible for managing Commerce Decisions' consultancy and training services and assisting clients in the deployment of best-practice process and tool support on their procurement projects. Peter's expertise includes procurement strategy, RFX development, requirements management, planning and executing evaluations, negotiation and bidder debriefing.

About Commerce Decisions

Commerce Decisions provides software and services to support the procurement and post contract review processes for both buyers and suppliers. For buyers we improve the efficiency and effectiveness of the evaluation process in order to make the best buying decision based on all the relevant criteria. For bidders we improve the quality and timeliness of proposals to best meet the needs of the potential buyer and thereby give them the best chance of securing the contract. Commerce Decisions' AWARD® software has been used internationally, by the defence, aerospace and utilities industries, healthcare and local and central government in a wide range of procurements of differing sizes and complexity, worth collectively over \$300 billion. Commerce Decisions is an acknowledged global leader in procurement best practice, with extensive experience in working with clients to successfully build appropriate and robust assessment schemes. In the Defence domain, Commerce Decisions has the unique position of providing support to the majority of UK MOD acquisition projects; internationally, Commerce Decisions has supported strategic programmes in Australia, Canada, UAE, Finland, USA, Saudi Arabia, Brunei and for NATO.

Weaknesses of traditional supplier evaluation methods

Commerce Decisions Ltd

101 Park Drive
Milton Park
Oxfordshire
OX14 4RY
United Kingdom

For more information contact us:

Tel: +44 (0) 123 543 1100

Email: sales@cd.QinetiQ.com

www.cd.QinetiQ.com