SUMMIT
CANADA'S MAGAZINE ON PUBLIC SECTOR PURCHASING

Know your vendors

www.jupiterimages.com

Evaluate your consultant's performance

by Chris F Jones MBA FCMC

In his book *"Managing the Professional Services Firm"* David Maister stresses the importance of evaluating the levels of service delivery and the results in connection with a consultant's client work. Yet few consulting firms, and even less of their clients, seem to make any serious attempt to evaluate their consultants' performance objectively after the engagement is finished – and go on to use the resulting information.

Many thousands of management consulting engagements are performed every year, but very few are objectively evaluated. Is no one interested in their performance? And if past performance can be measured, might it be useful to help qualify consultants for future engagements?

Evaluations are sometimes performed on whether or not the engagement itself is successful, but from the perspective of the future buyers and the consulting firms themselves, this is not the same thing at all.

Why is a consultant's performance so rarely measured after the engagement? The answer probably lies in whose interest the measurement lies. The consultant certainly has an interest in his/her own performance but a client who does not use consultants regularly has a lesser interest since he/she may never need the services again. Another reason is that the necessary performance criteria and indicators tend to be qualitative and thus quite difficult to measure.

If the performance was less than satisfactory, the client might just have to live with it.

Who cares?

Other than the consultants themselves, who *should* care about measuring consultants' performance? This answer is anyone who uses consultants on a regular basis and expects to use them regularly in the future – and yes, governments in particular.

In the corporate world, a client experiencing unsatisfactory consultant performance will most likely seek out another vendor next time. And the consultant will not use the client as a reference. In the public sector, the same thing happens, but the performance experience is less likely to be conveyed to other government buyers in any usable form. Word gets around but it is anecdotal at best and biased at worst. Poor consultant performance is rarely penalised.

Governments as a group are the largest single users of management consulting services in Canada. They buy the services, manage the engagements and the consultants, receive the deliverables, and then (for the most part) move on. But measurement of the consultants' performance is anecdotal unless the consultant presses the issue and follows Maister's advice in performing a formal performance evaluation afterwards. Some do, most do not. And when the performance is documented – good or bad – no one knows but the two parties involved.

Government clients should have a strong interest in evaluating their consultants' performance so they and their colleagues can engage only the higher performing consultants next time. Then the taxpayers' interests can be satisfied too. Yet almost every time they buy, governments start from scratch to determine and assess the vendors' qualifications. There is no generally accepted way to rate a consultant's prior performance, except perhaps through the references included in the proposal. More anecdotes.

Governments usually have a recognizable group of consultants and service providers – most consultants who provide services to governments do so regularly. Consultants complain about the time-consuming processes they have to follow to demonstrate their qualities in their proposals every time they write one. Is there a way to facilitate the RFP/proposal process by measuring and recording consultants' past performance with some objectivity so that government buyers can access and understand these qualities? Why not collect the consultants' performance information on a database and make it available to all government buyers?

What does "good performance" mean?

What does "good performance" mean in a consultant, and how can it be measured?

Maister lists fifteen qualitative elements of consultant performance in addition to the obvious ones of "on time and on budget." These are (in no particular order) as follows:

1. Did the consultants understand your business and your organization?

2. Did they listen to what you and your colleagues had to say and what you wanted, and not substitute their judgement for yours on what needed to be done?
3. Did they give good explanations on what they were doing, and why?
4. Did they tell you in advance what they planned to do?
5. Did they help you reach your own conclusions on the basis of the work they did?
6. Did they keep you sufficiently informed on their progress?
7. Did they document their activities well?
8. Did they avoid confusing jargon?
9. Were they accessible to you when you needed them?
10. Did they notify you promptly of any changes in scope, and then seek your approval?
11. Did they keep their promises on deadlines?
12. Did they involve you and your staff appropriately?
13. Did they make you feel as if you were important to them?
14. Did they show an interest in you and your organization beyond their specifics of their tasks?
15. Were they helpful to you beyond the specifics of their engagement?

There are additional qualities that are just as important in assessing consultants' performance, for example:

16. Were they properly qualified for the job in the first place?
17. Did they follow the "right" approach to the engagement?
18. Did they produce what you expected them to produce?
19. Did the consultants provide the benefits expected and defined in their terms of reference?

Not every question will be appropriate to every consulting engagement, but most will. If a client can truthfully say "yes" to all of the above questions, the consultant can be reasonably well satisfied with his/her performance on the engagement.

The matter of degree – evaluating shades of grey

In real life, we know that not all the above questions can be answered with a straight "yes" or "no" – there are invariably shades of grey – "yes, but…."

As well, one objective in measuring the consultant's performance is for both the client and the consultant to learn from their experience on the engagement. So we not only have to make accurate, transparent, defensible and *explicit* measurements about the strengths and weaknesses of the consultant's performance, we also need to combine the measurements to provide an overall assessment which can rank the performance of one consultant against others. We would also like to know their specific strengths and weaknesses. Only then can the evaluation be of comparative value to the buyers.

The pitfalls of using numeric scores to measure qualitative criteria are well documented.[1] As well, while most people are quite familiar with the idea of explicit criteria, few use *explicit measuring devices* to evaluate performance against the criteria they have defined. Instead they might rank them 1 to 5.

A Toronto company, ProGrid Evaluation Solutions (www.progrid.info), solved the

[1] For a discussion on why numeric scores fail to measure qualitative criteria adequately see C W Bowman, "*Intangibles*", Grafics Publishing, Sarnia, Ontario, 2005 available at www.clembowman.info. Dr. Bowman's book also contains a description of the ProGrid® theory and methodology, together with nineteen application case studies.

technicalities of the measurement problem in an innovative way. They developed and market an evaluation and assessment process that not only uses explicit measurement criteria but also explicit qualitative measurements against the criteria, written in plain language.

The ProGrid® approach overcomes the pitfalls by using a rubric[2] methodology. Their assessment process is used extensively in the science and technology grant-making business[3] where a grant application often reflects no more than "a glint in the eye" of an applicant and which contains no quantitative factors which can be measured with numeric ratings. The ProGrid process relies instead on a matrix containing the assessment criteria and a set of "Language Ladders™" which themselves contain the measurements.

Language Ladders – the measuring devices – contain four generic measurements for each qualitative criterion:

> *A level* - does not meet requirements/expectations;
> *B level* - meets some requirements/expectations but not all of them;
> *C level* - meets all requirements/expectations; and
> *D level* - exceeds all requirements/expectations.

There are no other options. While the Language Ladders themselves contain wording specific to the context, they *always* follow this same generic theme.

For example, if you were interested in assessing the consultant's performance in respect of a criterion called "understanding of our organization," a Language Ladder/rubric might be compiled as shown in *Table 1* on page 5.

Note that the generic measure is translated into an explicit measure designed specifically to suit the client's explicit context and circumstances. The assessment – the selection of level A, B, C or D – provides a measure of the consultant's performance with regard to his/her understanding of the organization expressed in plain language, not in numbers. The selection of a C level ranking does not mean 7/10 or 65 percent, or any other number. It means explicitly that "the consultant showed a good understanding of our culture and organization and fit in well with our staff" – something the client felt was a sensible component of performance.

So we not only have to make accurate, transparent, defensible and explicit measurements about the strengths and weaknesses of the consultant's performance, we also need to combine the measurements to provide an overall assessment which can rank the performance of one consultant against others.

[2] A rubric is a tool for scoring subjective assessments. It is a set of criteria and *standard measures* linked to explicit objectives that can be used to assess performance on topics that do not allow for quantitative measurement. Rubrics allow for a standardised evaluation according to specified criteria and measures, making assessment simpler, more consistent, more transparent and replicable.

[3] To the extent of over $1 billion a year in assessed projects.

Table 1 – Language Ladder Example – “Consultant’s Understanding of Our Organization”		
Generic Measure	**Explicit Measure**	**Select**
Did not meet expectations	The consultant showed little understanding of our culture and organization and did not fit in well with our staff	A
Met some, but not all expectations	The consultant showed a limited understanding of our culture and organization and there were occasions when he/she did not fit in well with our staff	B
Met all expectations	The consultant showed a good understanding of our culture and organization and fit in well with our staff	C
Exceeded expectations	The consultant showed a good understanding of our culture and organization and fit in well with our staff. **In addition**, the consultant made positive and realistic suggestions along the way to assist us in solving potential organizational problems	D

Putting it together – the performance evaluation

Three “overarching” criteria[4] or are used as the foundation for the qualitative evaluation. These are the:

1. technical quality of the consultants;
2. understanding of the consultants; and
3. the quality of the service delivered by the consultants

Each of these overarching criteria is divided into a series of individual performance criteria under headings that capture the essence of the qualitative criteria listed earlier. For example:

Technical quality

- Credentials of the consultant
- Project management
- Results of the engagement

Consultant’s understanding

- Understanding of the engagement
- Understanding of the client organization
- Understanding of the client’s context and business environment

Service quality

- Accessibility and responsiveness of the consultant
- Communications
- Attitude of the consultant
- Benefits arising from the engagement

In practice, each organization might determine specific performance criteria for its own consultants. However, in a competitive environment, there are obvious advantages in standardizing the performance criteria as much as possible to enable comparisons to be drawn between the consultants.[5]

[4] See Bowman, *op cit*, for an explanation of why there are almost always three primary or “overarching” criteria sets in any qualitative evaluation/ranking process.

[5] While such comparisons might be of greater interest to governments than to private sector clients, could client or consultant industry associations play a role in collecting this kind of comparative performance data?

The overarching criteria and the detailed performance criteria can now be placed in the first two columns of an evaluation worksheet (*Table 2*, page 7, shows an example of a worksheet for the technical quality component, the first of the three). Column 3 of the worksheet explains exactly what each performance criterion means, so there can be no ambiguity or misunderstanding of what is to be measured.

The performance measurements themselves (the Language Ladders) appear in column 4. There are four options for each criterion and each option is associated with a performance ranking A to D (column 5).

The evaluators (client and consultant) independently review the options in the Language Ladders and indicate their preference by circling A. B, C or D[6] for each criterion. Comments may be added as appropriate.

Both rankings are entered in a ProGrid® evaluation database and the results are calculated and reported. The consultant and the client then discuss (and resolve) the assessment variances before the data is finalized and captured in the consultants' performance database. If any variances are not resolved, explanatory comments may be added to the evaluation report and permanently archived.

Consultants' performance measurement report

The example in *Table 3* (page 8) shows how the results are reported by the ProGrid® evaluation software.

The report contains two components:

1. **Chart 1** – an "Overall Performance" graph showing the consultant's overall performance measured by the two most important of the "overarching" criteria, technical quality and service quality, with the numeric scores by the side,[7] and
2. **Chart 2** - a "Performance Profile" bar-chart showing the consultant's performance in respect of each of ten performance criteria determined by the client.

Once the assessment is completed, the reports are accumulated in a database so that any individual with "permitted" access might review the rankings for any consulting firm that has done business with the organization in the past. Both sets of rankings, and the comments on any unresolved variances, are permanently

[6] The ranking and report production processes are fully automated in the ProGrid® software and performed at the computer screen.

[7] Placement of the red diamond in the top right hand corner of the graph and a numeric score of 100 percent would indicate that the consultant achieved the D level by "exceeding requirements" in every one of the ten performance criteria.

retained in the database. This feature alone could cut down the need for consultants' "boiler plate" in competitive proposals very significantly. It would be well supported and quite objective.

Table 2 – Example of Consultants' Performance Evaluation Grid and Data Input Sheet				
1. Overarching Criterion	**2. Criterion**	**3. Meaning of the Criterion**	**4. Explicit Measures**	**5. Ranking (Select 1)**
Technical Quality	Credentials of the consultant	Education, professional certifications, credibility in the client's community, testimonials.	Were limited and not of sufficient quality to undertake this assignment	A
			Met most, but not all, the requirements for this engagement	B
			Fully met all the requirements for this engagement	C
			Exceeded the requirements for this engagement	D
	Project management	Approach to and management of the tasks required to complete the engagement	Did not manage the tasks well and needed assistance throughout the engagement	A
			Satisfactorily managed most tasks but fell short on some where assistance was needed	B
			Successfully managed all tasks required to deliver the engagement results	C
			AND presented innovative solutions when faced with difficult challenges	D
	Results of the engagement	Delivery of the engagement on time, on budget and according to the specifications in the contract	Did not deliver according to the engagement specifications	A
			Delivered the engagement according to the specifications, but was over budget or delivered the results late	B
			Successfully fully completed the engagement on-time, on-budget and according to specifications	C
			Delivered the engagement results exceeding specifications, and/or below budget, and/or ahead of schedule	D

Table 3 – Example of Consultant's Performance Report

Evaluating a Consultant's Performance: Post-engagement

Client Name: A Client
Consultant's Name: A Consultant

Summary of the Consultant's Engagement:

An Engagement

Chart 1 - Overall Performance

This chart shows the overall performance of the consultant in terms of both the Technical Quality and Service Qaulity provided.

The closer to the upper right-hand corner of the grid, the better the overall performance of the consultant.

Performance:

Technical Quality = 59%

Service Quality = 48%

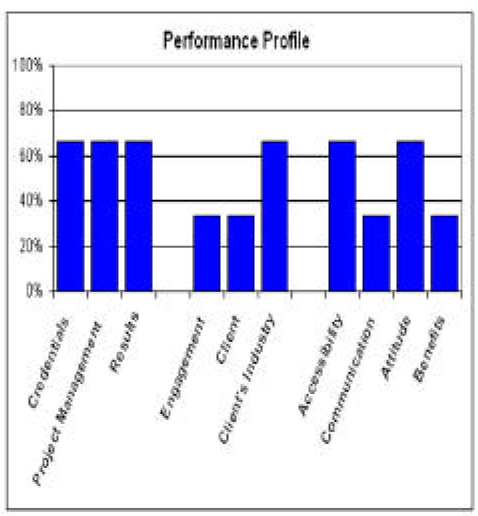
Overall Score = 54%

(The overall score represents the distance towards the point x=10 and y=10, with the point 10,10 representing 100% and the point 0,0 representing 0%)

Chart 2 - Performance Profile

This chart shows the ratings for each of the 10 evaluation criteria.

You can now print a copy of this evaluation and then return to the start to evaluate another consultant for comparison.



Conclusions - turning performance knowledge into action

Indicators of performance for management consulting firms tend to be qualitative. However, they can be measured and the results used to inform buyers for subsequent consulting engagements. Inexpensive evaluation and database software packages for these purposes are widely available.

Collection of the comparative evaluation data in a permanent database will not only facilitate the selection of consulting firms to perform future engagements, but also the preparation and evaluation of competitive proposals written in response to government RFPs.

Indicators of performance... can be measured

With minor amendments to the criteria and the measures, the process described will suit other types of professional service engagement. Obviously, the existence of multiple providers delivering a similar service to the same client, as is sometimes the case, adds another dimension to the potential evaluative value.

Chris F. Jones is president of FORUM Consulting Group www.forumconsulting.info a Fellow of the British Columbia Institute of Management Consultants and an associate faculty member, Faculty of Management, Royal Roads University, Victoria BC. He may be reached at info@forumconsulting.info. The author wishes to acknowledge the kind assistance of Fraser Barnes MBA CMC and John Kramers PhD of ProGrid Evaluation Solutions, Toronto and Edmonton, who developed the evaluation matrix and Language Ladders™ for the model described in this article