Establishing quantitative indicators for measuring the partnering performance of construction projects in Hong Kong

John F. Y. Yeung , Albert P. C. Chan & Daniel W. M. Chan

Establishing quantitative indicators for measuring the partnering performance of construction projects in Hong Kong

JOHN F. Y. YEUNG*, ALBERT P. C. CHAN and DANIEL W. M. CHAN

Department of Building and Real Estate, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Research into Key Performance Indicators (KPIs) for partnering projects in construction becomes vital because an increasing trend of client organizations has been observed to introduce a partnering approach to their building and construction works internationally during the last decade. A Partnering Performance Index (PPI) has been developed for construction projects in Hong Kong. The PPI can assist in developing a benchmark for measuring the performance of their partnering projects. However, it is worth noting that assessors may have their own semantic interpretations on each KPI. The aim of this paper is to establish suitable quantitative indicators (QIs) and reasonable quantitative ranges (QRs) for each KPI in order to avoid any possible discrepancies in interpreting the meaning of each KPI and provide objective evaluation results based on quantitative evidence. By conducting five structured face-to-face interviews and two rounds of a Delphi questionnaire survey in Hong Kong, a set of QIs were developed to measure the seven most important KPIs, including: (1) time performance; (2) cost performance; (3) top management commitment performance; (4) quality performance; (5) trust and respect performance; (6) effective communications performance; and (7) innovation and improvement performance. The identified QIs and QRs could assess and compare different partnering projects on a common basis objectively, thus helping to set a benchmark for measuring the performance level of partnering projects in Hong Kong. Construction senior executives and project managers can thus apply the QIs and QRs to measure, evaluate and improve the existing performance of their partnering projects in order to strive for construction excellence.

Keywords: Key Performance Indicators, partnering, quantitative indicators, Delphi method, Hong Kong.

Introduction

Partnering has been acknowledged as an innovative and non-confrontational relationship-based approach to the procurement of construction projects over the past decade (Construction Industry Institute, 1991 and 1996; Cowan *et al.*, 1992; Abudayyeh, 1994; Harback *et al.*, 1994; Lazar, 1997; Thompson and Sanders, 1998; Bayliss, 2000; Black *et al.*, 2000; Li *et al.*, 2001; Chan *et al.*, 2003). Many research studies reported that partnering is one of the conceivable solutions for improving overall project performance (Moore *et al.*, 1992; Mohr and Spekman, 1994; Construction Industry Board, 1997; Bresnen and Marshall, 2000; Lazar, 2000; Chan *et al.*, 2002 and 2006). In fact,

during the past decade, client organizations have indicated a wider application of introducing a partnering approach to their building and construction works both locally and internationally (Chan *et al.*, 2002). With the perceived benefits of adopting a partnering approach (Construction Industry Institute, 1991; Black *et al.*, 2000; Li *et al.*, 2001; Chan *et al.*, 2003), research into Key Performance Indicators (KPIs) for partnering projects in construction becomes important because it can help to develop a benchmark for measuring the performance of partnering projects.

A Partnering Performance Index (PPI) for construction projects in Hong Kong has been developed by Yeung *et al.* (in press) (Figure 1). The PPI is composed of seven most important KPIs, including: (1) time performance, with the weighting of 0.167; (2) cost performance, with the weighting of 0.160; (3) top

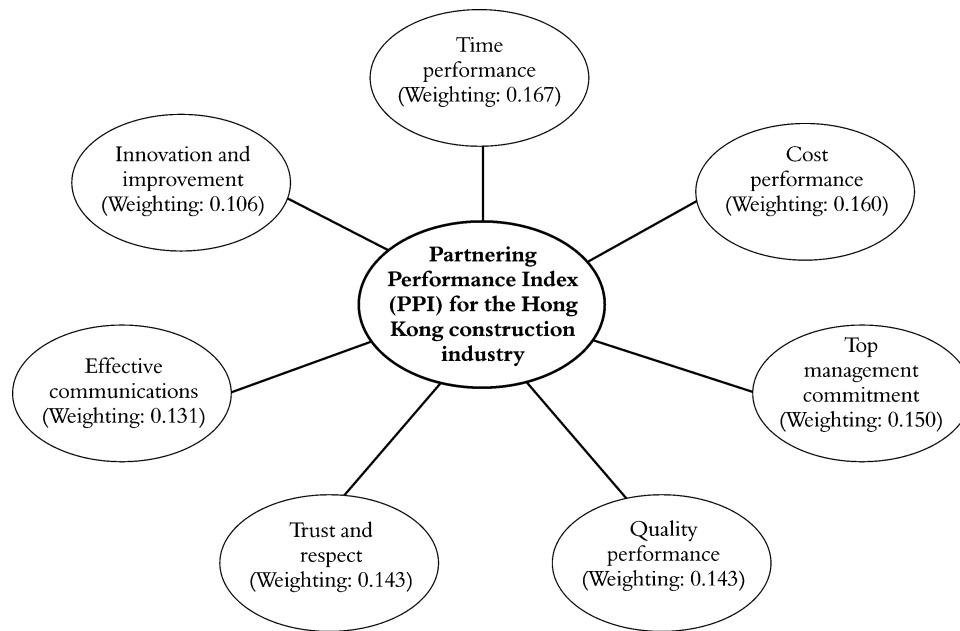*Author for correspondence. E-mail: bsjyeung@inet.polyu.edu.hk

Figure 1 The Partnering Performance Index (PPI) for the Hong Kong construction industry (Yeung *et al.*, in press)

management commitment, with the weighting of 0.150; (4) quality performance, with the weighting of 0.143; (5) trust and respect, with the weighting of 0.143; (6) effective communications, with the weighting of 0.131; and (7) innovation and improvement, with the weighting of 0.106. The coefficients of the KPIs are their individual weightings, which are calculated by their mean ratings divided by the total mean ratings. The Index is derived based on the assumption that this is a linear and additive model. It is logical and valid to derive this linear and additive model because the correlation matrix as shown in Table 1 reveals that nearly all the seven weighted KPIs are not highly correlated with each other at 5% significance level (more than half of them are even insignificantly correlated with each other). In addition, the units of

measurement for the seven weighted KPIs are different so there is unlikely to be any multiplier effect between them. Though it seems more sophisticated to use a non-linear model to fit the data obtained, overfitting is a common problem with non-linear models especially when the sample size is not sufficiently large (Neter *et al.*, 2005; Weisberg, 2005). That is why a linear, but not non-linear model is recommended if the relationship among variables is not proved to be non-linear. In fact, a linear model is assumed to be a linearized model of an unknown non-linear model if it really exists (Morrison, 1991; Griffiths, 1993). Practically speaking, it is simpler and easier to use this linear model to measure the partnering performance of construction projects in the Hong Kong construction industry.

Table 1 Correlation matrix among the seven selected weighted KPIs (for Round 4) (Yeung *et al.*, in press)

Correlation matrix	Time performance	Cost performance	Quality performance	Trust and respect	Top management commitment	Effective communications	Innovation and improvement
Time performance	1	0.464**	−0.193	0.414*	−0.181	−0.213	−0.166
Cost performance		1	−0.231	0.528**	−0.416*	−0.416*	−0.278
Quality performance			1	−0.271	0.804**	0.426*	0.205
Trust and respect				1	−0.256	−0.121	0.185
Top management commitment					1	0.571**	0.273
Effective communications						1	0.495**
Innovation and improvement							1

Notes: ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

However, it is worth noting that different assessors may have their own semantic interpretations on each KPI identified. For example, an assessor may use 'Percentage of conformance to the contract specifications' to measure quality performance while another assessor may adopt 'Number of non-conformance reports generated per month' to measure it. Even if a mutually agreed set of linguistic interpretations exists, its qualitative nature could lead to subjective judgment instead of evidence-based consideration. Thus, it is desirable to identify suitable quantitative indicators (QIs) for each KPI so as to avoid any possible discrepancies in interpreting the meaning of each KPI and provide objective evaluation results based on quantitative evidence.

The aim of this paper is to establish quantitative indicators (QIs) and quantitative ranges (QRs) appropriate for measuring each of the seven most important KPIs sought in Hong Kong. To achieve this, four objectives have been set, including (1) identifying a list of KPIs for measuring the partnering performance of construction projects in Hong Kong; (2) compiling a list of potential QIs to measure each of the seven selected weighted KPIs; (3) selecting the most vital QIs to measure each of the selected KPIs; and (4) defining reasonable quantitative ranges for different performance levels of each of the QIs. To do so, five structured face-to-face interviews were conducted with leading industrial practitioners who have been involved in partnering projects in Hong Kong to identify a list of potential QIs and subsequently, two rounds of Delphi questionnaire survey were undertaken with 31 relevant construction experts in Hong Kong in order to assess the appropriateness of the selected QIs by rating them against their level of importance, measurability and obtainability based on five-point Likert scales. After identifying appropriate QIs, reasonable quantitative ranges for measuring different performance levels of each of the QIs were defined through another empirical questionnaire survey. Finally, different partnering projects could be assessed and compared objectively on the same basis, thus assisting in setting a benchmark for measuring the performance level of partnering projects. The findings of this paper will be discussed, followed by highlighting the significance and limitations of the study.

Research method

The research methods employed in this paper included: (1) structured face-to-face interviews; (2) Delphi questionnaire survey; and (3) empirical questionnaire survey. A total of five structured face-to-face interviews

were conducted with leading industrial practitioners in Hong Kong who all have gained extensive hands-on experience in procuring partnering projects. The interviewees were invited via a set of structured open-ended questions to propose two most important QIs to evaluate the previously developed seven selected weighted KPIs for the Hong Kong construction industry. A total of 39 QIs for construction partnering projects were proposed by the five interviewees. The meanings of some QIs are similar in nature so they are combined and rephrased into one statement. And the QIs with the highest frequencies identified by the interviewees were selected for further study. Finally, 21 QIs (three QIs per each KPI) were formulated and consolidated for further analysis. Afterwards, two rounds of Delphi questionnaire survey were undertaken with the same 38 construction experts (one was retired) in Hong Kong, who were previously identified to conduct the four rounds of the first Delphi survey (Yeung *et al.*, in press), in order to assess the appropriateness of the selected QIs by rating them against their levels of importance, measurability and obtainability based on five-point Likert scales. Ultimately, the QIs with the highest mean rating for each of the seven selected KPIs were selected to measure the performance of partnering projects. Figure 2 shows the process of this research stage.

A major reason for employing the Delphi method in this research is that this method is a highly formalized method of communication that is designed to extract the maximum amount of unbiased information from a panel of experts (Chan *et al.*, 2001). In fact, it has been increasingly adopted in many complex areas in which a consensus is needed to be reached (Chan *et al.*, 2001), for example: (1) the development of residential areas (Anatharajan and Anataraman, 1982); (2) theory and design application (Corotis *et al.*, 1981); (3) bridge condition rating and effects of improvements (Saito and Sinha, 1991); (4) procurement selection (Chan *et al.*, 2001); and (5) sustainable development (Manoliadis *et al.*, 2006). Thus, it is suitable to adopt the Delphi method to obtain appropriate QIs to evaluate the performance of construction partnering projects because it is a rather subjective and new area of research.

Manoliadis *et al.* (2006) stated that the key issues in preparing a Delphi survey study were: (1) the definition of experts and their selection; (2) the number of rounds; and (3) the questionnaire structure (i.e. number of questions) in each study round. The Delphi method used in this research was composed of two rounds. In Round 1 of the Delphi questionnaire, the respondents were asked to provide ratings against the levels of importance, measurability and obtainability on each of the proposed QIs, based on five-point
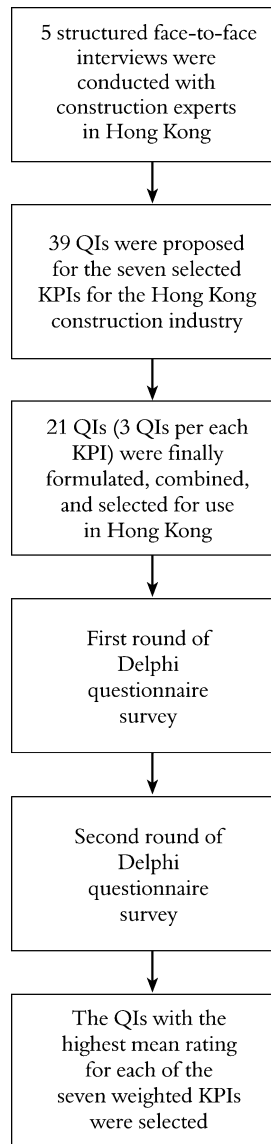
Figure 2 The process of this research stage

Likert scales, to measure the performance of partnering projects. While analysing the data, the focus ought to be on the opinion of the whole group rather than of individuals. In Round 2 of the Delphi questionnaire, respondents were provided with the consolidated results from Round 1. They were asked to reconsider the ratings of each QI to see if they would like to adjust their original ratings in the light of the consolidated results.

After identifying a set of QIs for measuring the partnering performance of construction projects, another empirical questionnaire survey was conducted in order to capture the perception of professionals as to the expectation of different performance levels for each of the QIs. The seven QIs selected were used for the design of a questionnaire. The questionnaire is divided

into two parts. The first part shows the results of previous Round 4 of the first Delphi Survey (Yeung *et al.*, in press) and the second part focuses on asking for the expectations of experts on each of the seven selected QIs with respect to five different performance levels namely 'poor', 'average', 'good', 'very good' and 'excellent'. A survey questionnaire together with a covering letter stating the objectives of the study was delivered to the same 31 respondent construction experts in Hong Kong. Of the 31 questionnaires distributed, 22 valid replies were received representing a response rate of 70.97%.

The success of Delphi method depends principally on the careful selection of the panel members (Chan *et al.*, 2001). The same 38 construction experts who contributed to the four rounds of the first Delphi questionnaire survey were again approached for this second stage of the research study. They represent a wide spectrum of construction professionals in Hong Kong, with eighteen from the private sector, eight from the public sector, six from the infrastructure sector and six from the academic sector. The infrastructure and academic sectors are considered as sectors in the same way as private building and public building sectors except that the infrastructure sector, which is different from the private and public building sectors, only includes mass transportation service providers, e.g. Mass Transit Railway Corporation Limited (MTRCL) and Kowloon–Canton Railway Corporation (KCRC). The composition of this group of experts provides a balanced view for the Delphi study.

Analysis of interview dialogues

Table 2 indicates the 39 QIs proposed by the five leading industrial practitioners during the face-to-face interviews. In fact, the transcriptions of the interview dialogues were sent back to all the interviewees for their verification before conducting the Delphi study. The meanings of some QIs are similar in nature so they are combined and rephrased into one statement. And the QIs with the highest frequencies identified by the interviewees were selected for further analysis. Finally, 21 QIs (three QIs per each KPI) were formulated and consolidated for further study. It should be emphasized that the reason why the QI survey did not comprise more instances with several items per KPI construct with these being selected for multivariate data analysis is that this performance evaluation model (PPI) has already included seven weighted KPIs after conducting four rounds of the first Delphi questionnaire survey. In order to make the PPI model more practical and easier to use, it is not good to have more than one QI per KPI

Table 2 The QIs proposed by the five leading industrial practitioners in Hong Kong

KPIs	The proposed QIs
Time performance	QI 1: Variation of project completion time against programme expressed as a percentage of project completion time
	QI 2: Variation of project completion time against completion time of best-in-class projects expressed as a percentage of completion time of best-in-class projects
	QI 3: Variation of project completion time against completion time of standard projects in similar type as a percentage of completion time of standard projects in similar type
	QI 4: Variation of initially mutually agreed completion time expressed as a percentage of finally mutually agreed completion time
	QI 5: Time predictability for design: measuring change between actual design time and predicted design time, expressed as a percentage of the estimated design time
	QI 6: Time predictability for construction: measuring change between actual construction time and predicted construction time, expressed as a percentage of the estimated construction time
	QI 7: Time improvement: measuring how much time improvement of a project is delivered to previous projects
	QI 8: Percentage of meeting milestone dates of a project by a main contractor
	QI 9: Composite time performance score by using Likert scale
Cost performance	QI 1: Variation of project completion cost against budget expressed as a percentage of project completion cost
	QI 2: Variation of project completion cost against completion cost of best-in-class projects expressed as a percentage of completion cost of best-in-class projects
	QI 3: Variation of project completion cost against completion cost of standard projects in similar type expressed as a percentage of completion cost of standard projects in similar type
	QI 4: Cost predictability for design: measuring change between actual design cost and predicted budget, expressed as a percentage of the estimated design budget
	QI 5: Cost predictability for construction: measuring change between actual construction cost and predicted construction cost, expressed as a percentage of the estimated construction cost
	QI 6: Cost improvement: measuring how much cost improvement of a project is delivered to the previous projects
	QI 7: Composite cost performance score by using Likert scale
Top management commitment	QI 1: Partnering development cost[1] of project expressed as a percentage of project completion cost
	QI 2: Ratio of time spent by project director in partnering steering/progress monitoring meetings to time by project director in project steering/progress monitoring meetings
	QI 3: Percentage of partnering steering/progress monitoring meetings attended by company director[2]
	QI 4: Percentage of partnering steering/progress monitoring meetings attended by director's/deputy director's representative (very often by project managers/deputy project managers)
	QI 5: Measuring level of top management commitment by using Likert scale
Quality performance	QI 1: Cost of rectifying major defects or non-conformances before project completion expressed as a percentage of project completion cost
	QI 2: Cost of rectifying major defects or non-conformances during defect liability period expressed as a percentage of project completion cost
	QI 3: Cost of rectifying major defects of a project expressed as a percentage of project completion cost
	QI 4: Ratio of number of non-conformance reports per month to the average number of non-conformance reports per month
	QI 5: Number of non-conformance reports (focusing on the trend over a period of time)
	QI 6: Number of complaints received by customers
	QI 7: Composite satisfaction scores of end users by using Likert scale
Trust and respect	QI 1: Average speed of resolving variations (for example, there are 5 major variations and the total duration to resolve them is 70 days, the average speed of resolving them is 14 days per variation)
	QI 2: Average speed of settling EOT claims (for example, there are 10 EOT claims and the total duration to settle them is 120 days, the average speed of settling them is 12 days per claim
	QI 3: Composite satisfaction scores of key stakeholders by using Likert scale
	QI 4: Frequency of meeting one's expectation about another party's behaviour and/or having confidence in another party

Table 2   (Continued.)

KPIs	The proposed QIs
Effective communications	QI 1: The difference between number of formal letter (per year) sent between parties and standard number of formal letter (per year) sent between parties
	QI 2: Number of formal letters and e-mails sent between parties both internally and externally per month
	QI 3: Composite satisfaction scores of key stakeholders by using Likert scale
Innovation and improvement	QI 1: Innovation and improvement cost saving expressed as a percentage of project completion cost
	QI 2: Innovation and improvement time saving expressed as a percentage of project completion time
	QI 3: Number of new initiatives for improvement introduced (construction techniques)
	QI 4: Composite satisfaction scores of key stakeholders by using Likert scale

Notes: [1] Partnering development cost is defined as a dedicated resource allocation cost from the total project completion cost, which includes (1) cost of employing facilitators; and (2) cost of organizing partnering. [2] Project director is defined as the most senior executive in a company responsible for managing the project.

construct in the model because it will make the model become too complex and difficult to use. In fact, to do so would require much more necessary data to be collected to compile the PPI. Therefore, a more practical way was adopted in which the most representative one out of the three most vital QIs per KPI according to their levels of importance, measurability and obtainability was chosen to measure the partnering performance of construction projects in Hong Kong.

Perceived QIs for time performance

Table 3 indicates the QIs for measuring the time performance of partnering projects, which were proposed by the five interviewees. Since the meanings of QIs 1, 4, 5 and 6 are similar in nature (the meanings of QIs 1 and 4 have already included the meanings of QIs 5 and 6), the four QIs are rephrased and combined into one. By the same logic, QIs 2, 3 and 7 are consolidated into one statement. Although 'best-in-class projects' as mentioned in QI 2 is different from 'standard projects' as mentioned in QI 3, the wording 'previous similar projects' as mentioned in the combined statement included both concepts. In addition, most interviewees perceived that it may not be easy to collect relevant data and during the interview, they were inclined to suggest using a composite time performance score by using a Likert scale to measure the time performance. Finally, the three most vital QIs identified for further study were: (1) 'Variation of actual completion time expressed as a percentage of finally agreed completion time' (also suggested by Collin, 2002); (2) 'Time improvement: measuring how much time improvement of a project is delivered to previous similar projects'; and (3) 'Composite time performance score by using Likert scale'.

Perceived QIs for cost performance

A similar approach was applied to identify QIs for other KPIs. Table 4 indicates the QIs for measuring the cost

performance of partnering projects, which were proposed by the five interviewees. As with the method for measuring the time performance, QIs 1, 4 and 5 are combined and rephrased into one statement because their meanings are similar (the meaning of QI 1 has already included the meanings of QIs 5 and 6). By the same logic, the QIs 2, 3 and 6 are consolidated into one statement. Although 'best-in-class projects' as mentioned in QI 2 is different from 'standard projects' as mentioned in QI 3, the wording 'previous similar projects' as mentioned in the combined statement included both concepts. In addition, all interviewees perceived that it is not easy to collect relevant data and they were inclined to recommend using a composite cost performance score by using a Likert scale to measure the cost performance. Finally, the three most vital QIs identified for measuring the cost performance of partnering projects were: (1) 'Variation of actual project cost expressed as a percentage of finally agreed project cost' (also suggested by Cheung *et al.*, 2003); (2) 'Cost improvement: measuring how much cost improvement of a project is delivered to previous similar projects'; and (3) 'Composite cost performance score by using Likert scale'.

Perceived QIs for top management commitment performance

A similar approach was applied to identify QIs for other KPIs. Table 5 indicates that the three most important QIs for measuring the top management commitment performance of partnering projects were: (1) 'Partnering development cost of project expressed as a percentage of total project cost'; (2) 'Percentage of top management attendance in partnering meetings'; and (3) 'Measuring level of top management commitment by using Likert scale' (say high level, moderate level or low level).

Table 3 Proposed and newly selected quantitative indicators (QIs) for measuring the time performance of partnering projects

Proposed QIs for measuring the time performance of partnering projects	Private sector		Public sector	Infrastructure sector		Total
	Interviewee A	Interviewee B	Interviewee C	Interviewee D	Interviewee E	
QI 1: Variation of project completion time against programme expressed as a percentage of project completion time	X	X	X	X	X	5
QI 2: Variation of project completion time against completion time of best-in-class projects expressed as a percentage of completion time of best-in-class projects				X		1
QI 3: Variation of project completion time against completion time of standard projects in similar type as a percentage of completion time of standard projects in similar type					X	1
QI 4: Variation of initially mutually agreed completion time expressed as a percentage of finally mutually agreed completion time			X			1
QI 5: Time predictability for design: measuring change between actual design time and predicted design time, expressed as a percentage of the estimated design time		X				1
QI 6: Time predictability for construction: measuring change between actual construction time and predicted construction time, expressed as a percentage of the estimated construction time		X				1
QI 7: Time improvement: measuring how much time improvement of a project is delivered to previous projects		X				1
QI 8: Percentage of meeting milestone dates of a project by a main contractor	X					1
QI 9: Composite time performance score by using Likert scale	Implied, but not directly identified	Implied, but not directly identified		Implied, but not directly identified		1

Remarks: QIs 1, 4, 5 and 6 are combined because their meanings are similar (selected with some wording rewritten)

QIs 2, 3 and 7 are combined because their meanings are similar (selected)

QI 8 is disposed because only 1 interviewee considered it to be an important QI for measuring time performance

QI 9 is selected because it is vital, easy to estimate and obtain (most interviewees implied that it is a useful QI to measure the time performance)

Newly selected QIs for measuring the time performance of partnering projects

QI 1: Variation of actual completion time expressed as a percentage of finally agreed completion time

QI 2: Time improvement: measuring how much time improvement of a project is delivered to previous similar projects

QI 3: Composite time performance score by using Likert scale

Perceived QIs for quality performance

Table 6 shows the three most vital QIs for measuring the quality performance of partnering projects. They were: (1) 'Cost of rectifying major defects or non-conformances of a project expressed as a percentage of total project cost'; (2) 'Average number of non-conformance reports generated per month' (also suggested by Bayliss *et al.*, 2004); and (3) 'Perceived end users' satisfaction scores by using Likert scale' (also suggested by Chan *et al.*, 2006).

Perceived QIs for trust and respect performance

Table 7 indicates that the three most important QIs for measuring the trust and respect performance of partnering projects were: (1) 'Average duration for

Table 4 Proposed and newly selected quantitative indicators (QIs) for measuring the cost performance of partnering projects

Proposed QIs for measuring the cost performance of partnering projects	Private sector		Public sector	Infrastructure sector		Total
	Interviewee A	Interviewee B	Interviewee C	Interviewee D	Interviewee E	
QI 1: Variation of project completion cost against budget expressed as a percentage of project completion cost		X	X	X	X	4
QI 2: Variation of project completion cost against completion cost of best-in-class projects expressed as a percentage of completion cost of best-in-class projects				X		1
QI 3: Variation of project completion cost against completion cost of standard projects in similar type expressed as a percentage of completion cost of standard projects in similar type					X	1
QI 4: Cost predictability for design: measuring change between actual design cost and predicted budget, expressed as a percentage of the estimated design budget		X				1
QI 5: Cost predictability for construction: measuring change between actual construction cost and predicted construction cost, expressed as a percentage of the estimated construction cost		X				1
QI 6: Cost improvement: measuring how much cost improvement of a project is delivered to the previous projects		X				1
QI 7: Composite cost performance score by using Likert scale	X	Implied, but not directly identified	Implied, but not directly identified	Implied, but not directly identified	Implied, but not directly identified	1

Remarks: QIs 1, 4 and 5 are combined because their meanings are similar (selected with some wording rewritten)
        QIs 2, 3 and 6 are combined because their meanings are similar (selected)
        QI 7 is selected because it is vital, easy to estimate and obtain (most interviewees implied that it is a useful QI to measure the cost performance)
Newly selected QIs for measuring the cost performance of partnering projects
QI 1: Variation of actual project cost expressed as a percentage of finally agreed project cost
QI 2: Cost improvement: measuring how much cost improvement of a project is delivered to previous similar projects
QI 3: Composite cost performance score by using Likert scale

settling variation orders and EOT claims'; (2) 'Frequency of meeting another party's expectation'; and (3) 'Perceived key stakeholders' satisfaction scores by using Likert scale' (also suggested by Cheung *et al.*, 2003).

Perceived QIs for effective communications performance

Table 8 shows the three most vital QIs for measuring the effective communications performance of partnering projects. They were: (1) 'Reduction of written communication: measuring how much written communication is reduced as compared to previous similar projects'; (2)

'Variation of the number of formal letters and e-mails sent between parties per month against the number with previous similar projects'; and (3) 'Perceived key stakeholders' satisfaction scores by using Likert scale'.

Perceived QIs for innovation and improvement performance

Table 9 shows that the three most important QIs for measuring the innovation and improvement performance of partnering projects were: (1) 'Cost saving resulting from innovation expressed as a percentage of total project cost'; (2) 'Number of innovative initiatives introduced (e.g. construction techniques, procurement

Table 5 Proposed and newly selected quantitative indicators (QIs) for measuring the top management commitment performance of partnering projects

Proposed QIs for measuring the top management commitment performance of partnering projects in	Private sector		Public sector	Infrastructure sector		Total
	Interviewee A	Interviewee B	Interviewee C	Interviewee D	Interviewee E	
QI 1: Partnering development cost of project expressed as a percentage of project completion cost				X		1
QI 2: Ratio of time spent by project director in partnering steering/progress monitoring meetings to time by project director in project steering/ progress monitoring meetings				X		1
QI 3: Percentage of partnering steering/progress monitoring meetings attended by company director	X					1
QI 4: Percentage of partnering steering/progress monitoring meetings attended by director's/deputy director's representative (very often by project managers/deputy project managers)			X			1
QI 5: Measuring level of top management commitment by using Likert scale		X			X	2

Remarks: QI 1 is selected because it is vital, and not difficult to measure and obtain

       QIs 2, 3 and 4 are combined because their meanings are similar (selected with some wording rewritten)

       QI 5 is selected because it is vital, easy to estimate and obtain

Newly selected QIs for measuring the top management commitment performance of partnering projects

QI 1: Partnering development cost of project expressed as a percentage of total project cost

QI 2: Percentage of top management attendance in partnering meetings

QI 3: Measuring level of top management commitment by using Likert scale (say high level, moderate level or low level)

approaches, management strategies)' (also suggested by Zhao, 2002); and (3) 'Perceived key stakeholders' satisfaction scores by using Likert scale'.

Two rounds of Delphi questionnaires

Round 1 of the Delphi questionnaire: ratings obtained from experts

Format

In the first round of the Delphi questionnaire, experts were requested to assess the appropriateness of the identified QIs by rating them against their levels of importance, measurability and obtainability based on five-point Likert scales. In addition, they were encouraged to suggest additional QIs for each KPI wherever deemed appropriate. The five-point Likert scales, ranging from 1=very unimportant/very difficult to measure and obtain, to 5=very important/very easy to measure and obtain, is used because the dimensions for measuring QIs should be bipolar, referring to the presence of opposite attributes, not unipolar, referring to different degrees of the same attribute (Schwarz, 1996). Only about half of the experts completed the questionnaire within one month. An individual e-mail

was sent to remind those experts who had not yet returned their completed questionnaires in time, and there was a follow-up phone call. Finally, 27 experts returned their completed questionnaires in late June of 2006.

Results and analysis

Factor analysis (FA) was conducted on each set of QIs to determine if they load on a single KPI. Table 10 shows the results of the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity in which the value of KMO is 0.433, which is smaller than 0.5. A small value for the KMO measure indicates that the appropriateness of factor analysis of the variables is weak because correlations between pairs of variables cannot be explained by the other variables (Norusis, 2005). Table 11 shows the results of the correlation matrix among QIs and it was found that most of the correlations between QIs were low and statistically insignificant at 5% level. These results support the propostion that each QI should not be combined with the other QIs.

In order to obtain a measure of consistency in ranking the QIs by the survey respondents, a statistical test was applied involving the calculation of Kendall's

Table 6 Proposed and newly selected quantitative indicators (QIs) for measuring the quality performance of partnering projects

Proposed QIs for measuring the quality performance of partnering projects	Private sector		Public sector	Infrastructure sector		Total
	Interviewee A	Interviewee B	Interviewee C	Interviewee D	Interviewee E	
QI 1: Cost of rectifying major defects or non-conformances before project completion expressed as a percentage of project completion cost				X		1
QI 2: Cost of rectifying major defects or non-conformances during defect liability period expressed as a percentage of project completion cost				X		1
QI 3: Cost of rectifying major defects of a project expressed as a percentage of project completion cost					X	1
QI 4: Ratio of number of non-conformance reports per month to the average number of non-conformance reports per month			X			1
QI 5: Number of non-conformance reports (focusing on the trend over a period of time)		X				1
QI 6: Number of complaints received by customers	X					1
QI 7: Composite satisfaction scores of end users by using Likert scale		X	Implied, but not directly identified	Implied, but not directly identified	Implied, but not directly identified	4

Remarks: QI 3 is selected because it is vital and easier to measure and obtain when compared with QIs 1 and 2
QI 4 is not selected because QI 5 has already reflected it in a better way
QI 5 is selected with some wording rewritten
QI 7 is selected because it is vital, easy to estimate and obtain

Newly selected QIs for measuring quality performance of partnering projects
QI 1: Cost of rectifying major defects or non-conformances of a project expressed as a percentage of total project cost
QI 2: Average number of non-conformance reports generated per month
QI 3: Perceived end users' satisfaction scores by using Likert scale

Coefficient of Concordance (W) for the QIs (Chan *et al.*, 2001). If the Concordance Coefficient is equal to 1, it means that all the experts rank the QIs identically. In contrast, if the Concordance Coefficient is equal to 0, it means that all the experts rank the QIs totally differently. Table 12 also shows that Kendall's Coefficient of Concordance (W) for the rankings of all the QIs was 0.290, which was statistically significant at 1% significance level. The null hypothesis that the respondent's ratings within the group are unrelated to each other has to be rejected. Therefore, it can be concluded that a significant amount of agreement among the respondents within the group of panel experts in ranking the QIs was found.

A statistical analysis was further performed on the 27 questionnaires received in which the mean ratings against the levels of importance, measurability and obtainability for each of the proposed QIs were computed. Table 12 shows the results of Round 1 of the Delphi questionnaire. It is indicated that the QIs with the highest mean ratings for the seven selected weighted KPIs were: (1) 'Variation of actual completion time expressed as a percentage of finally agreed completion time', with the mean rating of 4.47 (for measuring time performance); (2) 'Variation of actual project cost expressed as a percentage of finally agreed project cost', with the mean rating of 4.42 (for measuring cost performance); (3) 'Percentage of top management attendance in partnering meetings', with the mean rating of 4.44 (for measuring top management commitment); (4) 'Average number of non-conformance reports generated per month' (for measuring civil works and building works), with the mean rating of 4.02 (for measuring quality performance); (5) 'Perceived key stakeholders' satisfaction

Table 7 Proposed and newly selected quantitative indicators (QIs) for measuring the trust and respect performance of partnering projects

Proposed QIs for measuring the trust and respect performance of partnering projects	Private sector		Public sector	Infrastructure sector		Total
	Interviewee A	Interviewee B	Interviewee C	Interviewee D	Interviewee E	
QI 1: Average speed of resolving variations (for example, there are 5 major variations and the total duration to resolve them is 70 days, the average speed of resolving them is 14 days per variation)	X					1
QI 2: Average speed of settling EOT claims (for example, there are 10 EOT claims and the total duration to settle them is 120 days, the average speed of settling them is 12 days per claim	X					1
QI 3: Composite satisfaction scores of key stakeholders by using Likert scale	X	X	X	X	X	5
QI 4: Frequency of meeting one's expectation about another party's behaviour and/or having confidence in another party		X				1

Remarks: QIs 1 and 2 are combined and selected with some wording rewritten

　　　　QI 3 is selected with some wording rewritten

　　　　QI 4 is selected with some wording rewritten

Newly selected QIs for measuring the trust and respect performance of partnering projects

QI 1: Average duration for settling variation orders and EOT claims

QI 2: Frequency of meeting another party's expectation

QI 3: Perceived key stakeholders' satisfaction scores by using Likert scale

Table 8 Proposed and newly selected quantitative indicators (QIs) for measuring the effective communications performance of partnering projects

Proposed QIs for measuring the effective communications performance of partnering projects	Private sector		Public sector	Infrastructure sector		Total
	Interviewee A	Interviewee B	Interviewee C	Interviewee D	Interviewee E	
QI 1: The difference between number of formal letter (per year) sent between parties and standard number of formal letter (per year) sent between parties					X	1
QI 2: Number of formal letters and e-mails sent between parties both internally and externally per month			X			1
QI 3: Composite satisfaction scores of key stakeholders by using Likert scale	X	X		X		3

Remarks: QI 1 is selected with some wording rewritten

　　　　QI 2 is selected with some wording rewritten

　　　　QI 3 is selected with some wording rewritten

Newly selected QIs for measuring the effective communications performance of partnering projects

QI 1: Reduction of written communication: measuring how much written communication is reduced as compared to previous similar projects

QI 2: Variation of the number of formal letters and e-mails sent between parties per month against the number with previous similar projects

QI 3: Perceived key stakeholders' satisfaction scores by using Likert scale

Table 9 Proposed and newly selected quantitative indicators (QIs) for measuring the innovation and improvement performance of partnering projects

Proposed QIs for measuring the innovation and improvement performance of partnering projects	Private sector		Public sector	Infrastructure sector		Total
	Interviewee A	Interviewee B	Interviewee C	Interviewee D	Interviewee E	
QI 1: Innovation and improvement cost saving expressed as a percentage of project completion cost			X	X		2
QI 2: Innovation and improvement time saving expressed as a percentage of project completion time			X	X		2
QI 3: Number of new initiatives for improvement introduced (construc tion techniques)	X				X	2
QI 4: Composite satisfaction scores of key stakeholders by using Likert scale		X				1

Remarks: QI 1 is selected with some wording rewritten because it is better to reflect than the QI 2
        QI 3 is selected with some wording rewritten
        QI 4 is selected with some wording rewritten

Newly selected QIs for measuring the innovation and improvement performance of partnering projects
QI 1: Cost saving resulting from innovation expressed as a percentage of total project cost
QI 2: Number of innovative initiatives introduced (e.g. construction techniques, procurement approaches, management strategies)
QI 3: Perceived key stakeholders' satisfaction scores by using Likert scale

scores by using Likert scale', with mean rating of 3.74 (for measuring trust and respect performance); (6) 'Perceived key stakeholders' satisfaction scores by using Likert scale', with the mean rating of 3.49 (for measuring effective communications performance); and (7) 'Number of innovative initiatives introduced', with the mean rating of 3.81 (for measuring innovation and improvement performance). It may be of interest to note that a new QI for measuring the effective communications performance of partnering projects, which had not been proposed by the five interviewees, was suggested by one of the panel experts. The new QI is 'Integrated offices; frequency of/attendance at meetings'. Since only one panel member suggests it as a measure of effective communications, it was not selected for further study.

Round 2 of the Delphi questionnaire: reassessing the ratings

Format

For Round 2 of the Delphi survey, the experts were provided with the consolidated results obtained in Round 1. The average ratings of the 27 experts against the levels of importance, measurability and obtainability for each QI and the respondent's own ratings in Round 1 were shown. The respondents were asked to reassess their ratings in the light of the mean scored by the 27 experts. Round 2 of the Delphi questionnaire was distributed to the same group of panel experts by both postal mail and e-mail in late June of 2006. As in the previous round, an individual e-mail was forwarded to remind all the experts who had not yet returned their completed questionnaires in time, followed up by a

Table 10 Results of KMO measure and Bartlett's test of sphericity (for Round 1)

Kaiser–Meyer–Olkin measure of sampling adequacy		0.433
Bartlett's test of sphericity	Approx. chi-square	389.616
	df	210
	Sig	0.000

Table 11 Results of the correlation matrix among QIs (for Round 1)

										Correlations											
		TP1	TP2	TP3	CP1	CP2	CP3	TMC1	TMC2	TMC3	QP1	QP2	QP3	TR1	TR2	TR3	EC1	EC2	EC3	II1	II2	II3
TP1	Pearson Correlation	1	−.330	.135	.562**	.248	.131	.384*	.258	.282	.075	.472*	.035	.331	.115	.307	.183	.287	.284	.144	.043	.167
	Sig. (2-tailed)		.093	.503	.002	.212	.514	.048	.193	.155	.708	.013	.862	.092	.566	.119	.360	.147	.152	.475	.832	.405
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
TP2	Pearson Correlation	−.330	1	−.043	−.050	.438*	−.019	.102	−.319	−.076	.252	−.136	.147	−.032	.310	−.045	.224	.059	−.053	−.171	−.184	−.078
	Sig. (2-tailed)	.093		.831	.803	.022	.927	.611	.104	.707	.205	.498	.465	.875	.116	.823	.261	.771	.791	.394	.359	.697
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
TP3	Pearson Correlation	.135	−.043	1	.491**	.050	.537**	.039	−.071	.445*	−.171	.134	.419*	.207	.168	.436*	.213	.100	.600**	−.322	−.400*	.466*
	Sig. (2-tailed)	.503	.831		.009	.805	.004	.847	.727	.020	.394	.506	.030	.299	.401	.023	.285	.620	.001	.102	.039	.014
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
CP1	Pearson Correlation	.562**	−.050	.491**	1	.336	.228	.125	.053	.214	.018	.168	.021	.189	.213	.136	.216	.153	.237	.128	−.215	.021
	Sig. (2-tailed)	.002	.803	.009		.087	.253	.533	.792	.283	.929	.402	.915	.345	.287	.499	.279	.446	.234	.524	.281	.916
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
CP2	Pearson Correlation	.248	.438*	.050	.336	1	.183	.111	−.174	.149	.067	−.017	.021	.049	.242	−.046	.010	.003	−.077	.133	.049	−.203
	Sig. (2-tailed)	.212	.022	.805	.087		.362	.583	.385	.459	.738	.933	.918	.810	.224	.821	.961	.988	.703	.508	.807	.310
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
CP3	Pearson Correlation	.131	−.019	.537**	.228	.183	1	−.086	−.018	.686**	−.513**	.162	.467*	.251	.047	.363	−.034	−.023	.398*	−.282	−.277	.512**
	Sig. (2-tailed)	.514	.927	.004	.253	.362		.668	.927	.000	.006	.419	.014	.207	.818	.063	.868	.910	.040	.155	.161	.006
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
TMC1	Pearson Correlation	.384*	.102	.039	.125	.111	−.086	1	.308	.236	.316	.443*	.240	.386*	.287	.536**	.265	.405*	.399*	.168	−.049	.337
	Sig. (2-tailed)	.048	.611	.847	.533	.583	.668		.118	.237	.108	.021	.229	.046	.147	.004	.182	.036	.039	.403	.808	.086
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
TMC2	Pearson Correlation	.258	−.319	−.071	.053	−.174	−.018	.308	1	.213	.021	.253	−.174	.043	−.046	.368	.012	.147	.103	−.020	.241	.082
	Sig. (2-tailed)	.193	.104	.727	.792	.385	.927	.118		.286	.916	.204	.384	.831	.819	.059	.952	.466	.611	.922	.225	.684
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
TMC3	Pearson Correlation	.282	−.076	.445*	.214	.149	.686**	.236	.213	1	−.371	.380	.594**	.289	.113	.662**	.117	.275	.673**	−.163	−.199	.704**
	Sig. (2-tailed)	.155	.707	.020	.283	.459	.000	.237	.286		.056	.050	.001	.144	.575	.000	.560	.165	.000	.415	.318	.000
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27

Table 11 (Continued.)

												Correlations										
		TP1	TP2	TP3	CP1	CP2	CP3	TMC1	TMC2	TMC3	QP1	QP2	QP3	TR1	TR2	TR3	EC1	EC2	EC3	II1	II2	II3
QP1	Pearson Correlation	.075	.252	−.171	.018	.067	−.513**	.316	.021	−.371	1	−.068	−.177	.067	−.086	−.128	−.056	−.073	−.059	.014	−.217	−.124
	Sig. (2-tailed)	.708	.205	.394	.929	.738	.006	.108	.916	.056		.736	.377	.738	.671	.526	.780	.718	.771	.943	.277	.538
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
QP2	Pearson Correlation	.472*	−.136	.134	.168	−.017	.162	.443*	.253	.380	−.068	1	.362	.480*	.129	.573**	.182	.266	.571**	.104	.078	.543**
	Sig. (2-tailed)	.013	.498	.506	.402	.933	.419	.021	.204	.050	.736		.063	.011	.520	.002	.364	.180	.002	.605	.699	.003
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
QP3	Pearson Correlation	.035	.147	.419*	.021	.021	.467*	.240	−.174	.594**	−.177	.362	1	.436*	.283	.547**	.243	.305	.759**	−.125	−.371	.801**
	Sig. (2-tailed)	.862	.465	.030	.915	.918	.014	.229	.384	.001	.377	.063		.023	.153	.003	.221	.122	.000	.533	.056	.000
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
TR1	Pearson Correlation	.331	−.032	.207	.189	.049	.251	.386*	.043	.289	.067	.480*	.436*	1	.330	.467*	.359	.348	.613**	.047	−.206	.453*
	Sig. (2-tailed)	.092	.875	.299	.345	.810	.207	.046	.831	.144	.738	.011	.023		.093	.014	.066	.075	.001	.816	.302	.018
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
TR2	Pearson Correlation	.115	.310	.168	.213	.242	.047	.287	−.046	.113	−.086	.129	.283	.330	1	.397*	.519**	.536**	.342	.221	.064	.155
	Sig. (2-tailed)	.566	.116	.401	.287	.224	.818	.147	.819	.575	.671	.520	.153	.093		.040	.006	.004	.081	.268	.749	.441
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
TR3	Pearson Correlation	.307	−.045	.436*	.136	−.046	.363	.536**	.368	.662**	−.128	.573**	.547**	.467*	.397*	1	.458*	.594**	.836**	.083	.054	.768**
	Sig. (2-tailed)	.119	.823	.023	.499	.821	.063	.004	.059	.000	.526	.002	.003	.014	.040		.016	.001	.000	.681	.789	.000
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
EC1	Pearson Correlation	.183	.224	.213	.216	.010	−.034	.265	.012	.117	−.056	.182	.243	.359	.519**	.458*	1	.893**	.397*	.253	.129	.216
	Sig. (2-tailed)	.360	.261	.285	.279	.961	.868	.182	.952	.560	.780	.364	.221	.066	.006	.016		.000	.040	.203	.522	.280
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
EC2	Pearson Correlation	.287	.059	.100	.153	.003	−.023	.405*	.147	.275	−.073	.266	.305	.348	.536**	.594**	.893**	1	.439*	.310	.147	.317
	Sig. (2-tailed)	.147	.771	.620	.446	.988	.910	.036	.466	.165	.718	.180	.122	.075	.004	.001	.000		.022	.116	.465	.107
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
EC3	Pearson Correlation	.284	−.053	.600**	.237	−.077	.398*	.399*	.103	.673**	−.059	.571**	.759**	.613**	.342	.836**	.397*	.439*	1	.036	−.153	.888**
	Sig. (2-tailed)	.152	.791	.001	.234	.703	.040	.039	.611	.000	.771	.002	.000	.001	.081	.000	.040	.022		.860	.445	.000
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27

Table 11 (Continued.)

Correlations

		TP1	TP2	TP3	CP1	CP2	CP3	TMC1	TMC2	TMC3	QP1	QP2	QP3	TR1	TR2	TR3	EC1	EC2	EC3	II1	II2	II3
II1	Pearson Correlation	.144	−.171	−.322	.128	.133	−.282	.168	−.020	−.163	.014	.104	−.125	.047	.221	.083	.253	.310	.036	1	.622**	−.022
	Sig. (2-tailed)	.475	.394	.102	.524	.508	.155	.403	.922	.415	.943	.605	.533	.816	.268	.681	.203	.116	.860		.001	.915
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
II2	Pearson Correlation	.043	−.184	−.400*	−.215	.049	−.277	−.049	.241	−.199	−.217	.078	−.371	−.206	.064	.054	.129	.147	−.153	.622**	1	−.152
	Sig. (2-tailed)	.832	.359	.039	.281	.807	.161	.808	.225	.318	.277	.699	.056	.302	.749	.789	.522	.465	.445	.001		.450
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
II3	Pearson Correlation	.167	−.078	.466*	.021	−.203	.512**	.337	.082	.704**	−.124	.543**	.801**	.453*	.155	.768**	.216	.317	.888**	−.022	−.152	1
	Sig. (2-tailed)	.405	.697	.014	.916	.310	.006	.086	.684	.000	.538	.003	.000	.018	.441	.000	.280	.107	.000	.915	.450	
	N	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27

** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

phone call. Finally, 25 experts completed their questionnaires in late August of 2006.

Results and analysis

The same factor analysis (Table 13), correlation matrix between QIs (Table 14) and measure of consistency through the calculation of the Kendall's Coefficient of Concordance (W) (Table 13) were conducted for Round 2 of the Delphi questionnaire and the results were similar to Round 1 and therefore the same conclusions were drawn (i.e. each QI should not be combined with the other QIs and there was significant amount of agreement among the respondents within the group of panel experts in ranking the QIs.)

Most experts had reconsidered their ratings provided in the previous round and had made adjustments to their ratings. However, Table 15 shows that all the QIs with the highest mean ratings are still the same when compared with the consolidated results in Round 1, except that 'Number of innovative initiatives introduced (e.g. construction techniques, procurement approaches, management strategies)' was replaced by 'Cost saving resulting from innovation expressed as a percentage of total project cost' to measure the innovation and improvement performance of partnering projects.

Questionnaire results: mean expectations of the quantitative assessment against the five different performance levels

Although a set of QIs established can provide a mutually agreed set of linguistic interpretations and lead to more objective performance evaluation for partnering projects, it cannot fully eliminate the subjectivity of evaluation as different assessors may perceive the same performance level with different numerical figures. For example, a 2% reduction in project cost may represent 'good performance' to someone who is not too demanding; but a 5% reduction in project cost may be perceived as 'average performance' to someone who has a higher expectation. To remedy this deficiency, an empirical questionnaire survey was conducted in order to capture the perception of professionals as to the expectation of different performance levels for each of the QIs. The results of the questionnaire are summarized in Table 16 in which the mean expectation (ME) and coefficient of variation (CV) of each QI against the five performance levels are listed. A closer inspection to the coefficient of variation (CV) reveals that there are slight to moderate

Table 12   Result of Round 1 of the Delphi questionnaire in Hong Kong

Quantitative indicators for measuring the performance of partnering projects in Hong Kong	Average ratings of experts in Round 1			
Quantitative indicators for measuring time performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Variation of actual completion time expressed as a percentage of finally agreed completion time	4.52	4.56	4.33	4.47
Time improvement: measuring how much time improvement of a project is delivered to previous similar projects	3.74	3.00	3.04	3.26
Composite time performance score by using Likert scale	3.15	3.41	3.44	3.33
Quantitative indicators for measuring cost performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Variation of actual project cost expressed as a percentage of finally agreed project cost	4.48	4.48	4.30	4.42
Cost improvement: measuring how much cost improvement of a project is delivered to previous similar projects	3.96	3.19	3.33	3.49
Composite cost performance score by using Likert scale	3.22	3.67	3.96	3.62
Quantitative indicators for measuring top management commitment performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Partnering development cost of project expressed as a percentage of total project cost	3.19	4.00	4.04	3.74
Percentage of top management attendance in partnering meetings	4.19	4.59	4.56	4.44
Measuring level of top management commitment by using Likert scale (say high level, moderate level or low level)	3.70	3.85	4.00	3.85
Quantitative indicators for measuring quality performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Cost of rectifying major defects or non-conformances of a project expressed as a percentage of total project cost	4.19	3.56	3.11	3.62
Average number of non-conformance reports generated per month	3.93	4.19	3.96	4.02
Perceived end users' satisfaction scores by using Likert scale	3.81	3.37	3.59	3.59
Quantitative indicators for measuring trust and respect performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Average duration for settling variation orders	3.70	3.59	3.56	3.62
Frequency of meeting another party's expectation	3.52	2.44	2.67	2.88
Perceived key stakeholders' satisfaction scores by using Likert scale	3.78	3.63	3.81	3.74
Quantitative indicators for measuring effective communications performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Reduction of written communication: measuring how much written communication is reduced as compared to previous similar projects	3.37	2.89	2.78	3.01
Variation of the number of formal letters and e-mails sent between parties per month against the number with previous similar projects	3.04	2.78	2.78	2.86
Perceived key stakeholders' satisfaction scores by using Likert scale	3.63	3.30	3.56	3.49
Quantitative indicators for measuring innovation and improvement performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Cost saving resulting from innovation expressed as a percentage of total project cost	4.19	3.48	3.37	3.68
Number of innovative initiatives introduced (e.g. construction techniques, procurement approaches, management strategies)	3.89	3.81	3.74	3.81
Perceived key stakeholders' satisfaction scores by using Likert scale	3.48	3.37	3.63	3.49
Kendall's Coefficient of Concordance (W)			0.290	
Level of significance			0.000	

Notes: Remarks 1 : Rating 1=very unimportant/very difficult and 5=very important/very easy.

Table 13 Results of KMO measure and Bartlett's test sphericity (for Round 2)

Kaiser–Meyer–Olkin measure of sampling adequacy		0.350
Bartlett's test of sphericity	Approx. chi-square	393.567
	df	210
	Sig	0.000

deviations from the mean value in most of the performance levels describing the QIs. Nevertheless, the deviations are high for 'Variation of actual completion time expressed as a percentage of finally agreed completion time' (CV for the average performance=−2.90 and CV for the good performance= 1.18); 'Variation of actual project cost expressed as a percentage of finally agreed project cost' (CV for the average performance=−2.36 and CV for the good performance=1.01); 'Average number of non-conformance reports generated per month for civil works' (CV for the excellent performance=2.17); and 'Average number of non-conformance reports generated per month for building works' (CV for the excellent performance=1.52). The results show that differences in expectation exist between the construction experts in the perceived performance level of each QI. Thus, despite the fact that the mean value can serve as a quick rule of thumb for evaluators to differentiate an 'average' and 'good' performance of a partnering project, it is more appropriate to identify a quantitative range (QR) of reasonable expectation for each performance level as shown in Figure 3. A similar approach was adopted by Chow and Ng (2007). Therefore, a partnering project with 'good' time performance would be one with for example ahead of schedule by 0.68% to 8.82%. In this example, the lower boundary for the 'good' time performance was simply taken as the average of the mean expectation for the 'average' time performance (mean expectation for the 'average' performance=ahead of schedule by 1.25%) and 'good' time performance (mean expectation for the 'good' performance=ahead of schedule by 3.86%) and the average of the mean expectation for the 'good' time performance (mean expectation for the 'good' performance=ahead of schedule by 3.86%) and 'very good' time performance (mean expectation for the 'very good' performance=ahead of schedule by 9.91%). Table 17 shows all the QRs for each of the seven QIs.

Further research studies should be carried out to establish a more scientific way to define a range for different performance levels. It is clear that accurate estimation of the performance ranges would provide a greater flexibility for assessors to objectively, reliably and practically evaluate the partnering performance of construction projects.

Discussion and validation of research findings

The research findings indicate that three QIs with the highest mean ratings were cross-referenced with the reported literature. They were: (1) 'Variation of actual completion time expressed as a percentage of finally agreed completion time' (for measuring time performance) (also suggested by Collin, 2002); (2) 'Variation of actual project cost expressed as a percentage of finally agreed project cost' (for measuring cost performance) (also suggested by Cheung et al., 2003); and (3) 'Average number of non-conformance reports generated per month' (for measuring civil works and building works) (for measuring quality performance) (also suggested by Bayliss et al., 2004). It is worth noting that the current Delphi survey results are in line with earlier research results with similar research topics (Crane et al., 1999; Cheung et al., 2003 and Lo et al., 2006).

Crane et al. (1999) conducted detailed interviews with 21 successful partnering relationships and then classified partnering measures into three types: result, process and relationship measures. Result and process measures include (1) cost; (2) schedule; (3) safety; (4) quality; and (5) litigation. Relationship measures encompass (1) internal communication; (2) external communication; (3) meeting effectiveness; (4) worker morale; (5) internal trust; (6) external trust; (7) internal leadership; (8) external leadership; (9) accomplishment of objectives; (10) utilization of resources; (11) problem solving; (12) creativity and synergy; (13) timely evaluation and appropriate response; (14) definition and adherence to roles and responsibilities; (15) continuous improvement; and (16) teamwork. However, no further performance index was developed and appropriate QIs were not identified for assessing the partnering performance of construction projects, thus making benchmarking difficult.

Cheung et al. (2003) adopted eight partnering measures suggested by the New South Wales Public Works Department of Australia. The eight partnering measures were (1) time; (2) cost; (3) quality; (4) safety; (5) communication; (6) claim and issue resolution; (7) environment; and (8) contract relations. A Partnering Temperature Index was developed and an IT system was used to measure the performance of partnering

Table 14 Results of the correlation matrix among QIs (for Round 2)

										Correlations											
		TP1	TP2	TP3	CP1	CP2	CP3	TMC1	TMC2	TMC3	QP1	QP2	QP3	TR1	TR2	TR3	EC1	EC2	EC3	II1	II2	II3
TP1	Pearson Correlation	1	−.325	.156	.703**	.178	.163	.205	.402*	.092	.259	.414*	.026	.399*	.193	.231	.291	.286	.281	.175	−.057	.021
	Sig. (2-tailed)		.112	.457	.000	.395	.436	.325	.046	.661	.211	.039	.902	.048	.356	.266	.159	.166	.173	.404	.785	.920
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
TP2	Pearson Correlation	−.325	1	−.043	−.204	.306	−.206	.242	−.168	−.113	.175	−.100	.157	−.125	.116	.082	−.032	−.101	−.002	−.166	−.220	.000
	Sig. (2-tailed)	.112		.839	.328	.137	.324	.244	.421	.589	.403	.634	.454	.550	.582	.696	.878	.632	.992	.427	.291	.999
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
TP3	Pearson Correlation	.156	−.043	1	.348	−.078	.708**	−.158	−.050	.456*	−.145	.186	.564**	.284	.232	.281	.094	−.075	.565**	−.382	−.522**	.405*
	Sig. (2-tailed)	.457	.839		.088	.712	.000	.450	.814	.022	.488	.374	.003	.169	.265	.174	.654	.720	.003	.060	.007	.044
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
CP1	Pearson Correlation	.703**	−.204	.348	1	.382	.412*	−.021	.099	.149	.090	.249	.188	.319	.196	.185	.140	.096	.268	.206	−.329	.038
	Sig. (2-tailed)	.000	.328	.088		.060	.041	.921	.636	.478	.670	.229	.369	.121	.348	.375	.506	.649	.195	.324	.108	.858
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
CP2	Pearson Correlation	.178	.306	−.078	.382	1	.070	−.214	−.109	−.102	.181	−.049	−.087	.023	.118	−.084	−.205	−.256	−.133	.164	−.075	−.367
	Sig. (2-tailed)	.395	.137	.712	.060		.738	.305	.605	.628	.387	.815	.680	.913	.573	.689	.326	.218	.525	.435	.720	.071
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
CP3	Pearson Correlation	.163	−.206	.708**	.412*	.070	1	−.252	−.103	.610**	−.290	.250	.501*	.402*	−.003	.325	−.066	−.242	.518**	−.210	−.375	.517**
	Sig. (2-tailed)	.436	.324	.000	.041	.738		.224	.624	.001	.160	.229	.011	.046	.987	.113	.753	.244	.008	.315	.065	.008
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
TM-C1	Pearson Correlation	.205	.242	−.158	−.021	−.214	−.252	1	.354	−.086	.150	.373	.084	.323	.173	.382	.265	.331	.271	.140	−.020	.176
	Sig. (2-tailed)	.325	.244	.450	.921	.305	.224		.082	.683	.473	.067	.691	.115	.409	.059	.201	.107	.189	.504	.923	.401
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
TM-C2	Pearson Correlation	.402*	−.168	−.050	.099	−.109	−.103	.354	1	.121	.296	.385	−.217	.130	−.088	.152	.120	.030	.057	−.111	.068	.023
	Sig. (2-tailed)	.046	.421	.814	.636	.605	.624	.082		.563	.151	.058	.298	.534	.677	.467	.568	.888	.789	.598	.745	.912
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
TM-C3	Pearson Correlation	.092	−.113	.456*	.149	−.102	.610**	−.086	.121	1	−.320	.369	.621**	.260	−.013	.545**	.121	.083	.706**	−.259	−.328	.755**
	Sig. (2-tailed)	.661	.589	.022	.478	.628	.001	.683	.563		.119	.070	.001	.210	.950	.005	.563	.692	.000	.212	.109	.000
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25

Table 14   (Continued.)

									Correlations												
		TP1	TP2	TP3	CP1	CP2	CP3	TMC1	TMC2	TMC3	QP1	QP2	QP3	TR1	TR2	TR3	EC1	EC2	EC3	II1	II2	II3
QP1	Pearson Correlation	.259	.175	−.145	.090	.181	−.290	.150	.296	−.320	1	−.059	−.329	−.030	−.032	−.008	−.153	−.090	−.155	−.091	−.146	−.274
	Sig. (2-tailed)	.211	.403	.488	.670	.387	.160	.473	.151	.119		.778	.108	.888	.881	.971	.467	.669	.459	.666	.487	.185
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
QP2	Pearson Correlation	.414*	−.100	.186	.249	−.049	.250	.373	.385	.369	−.059	1	.442*	.686**	.273	.611**	.286	.373	.617**	−.025	−.083	.523**
	Sig. (2-tailed)	.039	.634	.374	.229	.815	.229	.067	.058	.070	.778		.027	.000	.187	.001	.165	.066	.001	.907	.694	.007
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
QP3	Pearson Correlation	.026	.157	.564**	.188	−.087	.501*	.084	−.217	.621**	−.329	.442*	1	.470*	.529**	.595**	.268	.259	.815**	−.156	−.551**	.746**
	Sig. (2-tailed)	.902	.454	.003	.369	.680	.011	.691	.298	.001	.108	.027		.018	.007	.002	.195	.212	.000	.456	.004	.000
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
TR1	Pearson Correlation	.399*	−.125	.284	.319	.023	.402*	.323	.130	.260	−.030	.686**	.470*	1	.311	.539**	.202	.152	.592**	.033	−.206	.414*
	Sig. (2-tailed)	.048	.550	.169	.121	.913	.046	.115	.534	.210	.888	.000	.018		.131	.005	.332	.469	.002	.875	.323	.040
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
TR2	Pearson Correlation	.193	.116	.232	.196	.118	−.003	.173	−.088	−.013	−.032	.273	.529**	.311	1	.476*	.498*	.507**	.433*	.406*	−.142	.169
	Sig. (2-tailed)	.356	.582	.265	.348	.573	.987	.409	.677	.950	.881	.187	.007	.131		.016	.011	.010	.031	.044	.498	.419
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
TR3	Pearson Correlation	.231	.082	.281	.185	−.084	.325	.382	.152	.545**	−.008	.611**	.595**	.539**	.476*	1	.526**	.528**	.825**	.148	−.143	.681**
	Sig. (2-tailed)	.266	.696	.174	.375	.689	.113	.059	.467	.005	.971	.001	.002	.005	.016		.007	.007	.000	.479	.496	.000
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
EC1	Pearson Correlation	.291	−.032	.094	.140	−.205	−.066	.265	.120	.121	−.153	.286	.268	.202	.498*	.526**	1	.861**	.326	.359	.123	.184
	Sig. (2-tailed)	.159	.878	.654	.506	.326	.753	.201	.568	.563	.467	.165	.195	.332	.011	.007		.000	.111	.078	.559	.379
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
EC2	Pearson Correlation	.286	−.101	−.075	.096	−.256	−.242	.331	.030	.083	−.090	.373	.259	.152	.507**	.528**	.861**	1	.306	.358	.062	.188
	Sig. (2-tailed)	.166	.632	.720	.649	.218	.244	.107	.888	.692	.669	.066	.212	.469	.010	.007	.000		.137	.078	.767	.367
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
EC3	Pearson Correlation	.281	−.002	.565**	.268	−.133	.518**	.271	.057	.706**	−.155	.617**	.815**	.592**	.433*	.825**	.326	.306	1	.012	−.306	.864**
	Sig. (2-tailed)	.173	.992	.003	.195	.525	.008	.189	.789	.000	.459	.001	.000	.002	.031	.000	.111	.137		.953	.136	.000
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25

Table 14 (Continued.)

Correlations

		TP1	TP2	TP3	CP1	CP2	CP3	TMC1	TMC2	TMC3	QP1	QP2	QP3	TR1	TR2	TR3	EC1	EC2	EC3	I1	I2	I3
I1	Pearson Correlation	.175	−.166	−.382	.206	.164	−.210	.140	−.111	−.259	−.091	−.025	−.156	.033	.406*	.148	.359	.358	.012	1	.565**	−.081
	Sig. (2-tailed)	.404	.427	.060	.324	.435	.315	.504	.598	.212	.666	.907	.456	.875	.044	.479	.078	.078	.953		.003	.701
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
I2	Pearson Correlation	−.057	−.220	−.522**	−.329	−.075	−.375	−.020	.068	−.328	−.146	−.083	−.551**	−.206	−.142	−.143	.123	.062	−.306	.565**	1	−.254
	Sig. (2-tailed)	.785	.291	.007	.108	.720	.065	.923	.745	.109	.487	.694	.004	.323	.498	.496	.559	.767	.136	.003		.220
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
I3	Pearson Correlation	.021	.000	.405*	.038	−.367	.517**	.176	.023	.755**	−.274	.523**	.746**	.414*	.169	.681**	.184	.188	.864**	−.081	−.254	1
	Sig. (2-tailed)	.920	.999	.044	.858	.071	.008	.401	.912	.000	.185	.007	.000	.040	.419	.000	.379	.367	.000	.701	.220	
	N	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25

** Correlation is significant at the 0.01 level (2-tailed).   * Correlation is significant at the 0.05 level (2-tailed).

projects. However, the weightings by default were treated as equal for each measure.

Lo *et al.* (2006) used a Balanced Scorecard (BSC) approach to measure the partnering project performance in a holistic manner through an extensive literature review and data analysis (principal components factor analysis) through a questionnaire survey. Four performance measurement perspectives: (1) benefits; (2) attitudes of project stakeholders; (3) attitudes enhancement process; and (4) strategic learning and growth, and 36 strategic objectives (including improved quality, reduced project cost and time, open communication, mutual trust, and top level commitment) were developed. Although it was comprehensive for this approach to assess partnering performance of construction projects in Hong Kong, different industrial practitioners might interpret the same strategic objectives differently. In addition, corresponding weightings were not derived for different strategic objectives, thus making benchmarking difficult as well.

It should also be pointed out that these researchers encountered the problem of subjectivity in selecting the most vital KPIs for partnering projects without good resolution methods. In contrast, the Delphi survey method used in this research study could assist in extracting the maximum amount of unbiased information from a panel of experts (Chan *et al.*, 2001).

It should also be noted that the Delphi method by its inherent nature serves as a self-validating mechanism because individual experts are given the chance to reassess their scores with reference to the consolidated mean scores as assessed by other experts (Yeung *et al.*, in press). And it is logical and reasonable to define a range for different performance levels by taking the average of two consecutive performance levels. By doing so, assessors can have greater flexibility to evaluate the partnering performance of construction projects without sacrificing objectiveness and reliability.

Significance and limitations of the study

This paper established a series of QIs and reasonable QRs for different performance levels for each of the QIs to measure the previously developed most important seven weighted KPIs for partnering projects in the Hong Kong construction industry. These QIs and QRs could prevent various assessors from applying their subjective interpretation to each KPI during evaluation. In order to compile a rational list of QIs for each of the seven weighted KPIs, five structured face-to-face interviews were launched to suggest appropriate QIs,

Table 15 Result of Round 2 of the Delphi questionnaire in Hong Kong

Quantitative indicators for measuring the performance of partnering projects in Hong Kong	Average ratings of experts in Round 2			
Quantitative indicators for measuring time performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Variation of actual completion time expressed as a percentage of finally agreed completion time	4.56	4.64	4.38	4.53
Time improvement: measuring how much time improvement of a project is delivered to previous similar projects	3.80	2.92	3.04	3.25
Subjective assessment by using Likert scale (say ahead of schedule, on time or behind schedule)	3.12	3.30	3.46	3.29
Quantitative indicators for measuring cost performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Variation of actual project cost expressed as a percentage of finally agreed project cost	4.56	4.48	4.32	4.45
Cost improvement: measuring how much cost improvement of a project is delivered to previous similar projects	4.00	3.08	3.32	3.47
Subjective assessment by using Likert scale (say within budget, on budget or overrun budget)	3.12	3.62	3.68	3.47
Quantitative indicators for measuring top management commitment performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Partnering development cost of project expressed as a percentage of total project cost	3.00	4.08	4.12	3.73
Percentage of top management attendance in partnering meetings	4.32	4.60	4.60	4.51
Measuring level of top management commitment by using Likert scale (say high level, moderate level or low level)	3.62	3.80	3.88	3.77
Quantitative indicators for measuring quality performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Cost of rectifying major defects or non-conformances of a project expressed as a percentage of total project cost	4.26	3.76	3.22	3.75
Average number of non-conformance reports generated per month	3.94	4.28	4.08	4.10
Perceived end users' satisfaction scores by using Likert scale	3.74	3.46	3.74	3.65
Quantitative indicators for measuring trust and respect performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Average duration for settling variation orders	3.64	3.56	3.48	3.56
Frequency of meeting another party's expectation	3.50	2.46	2.64	2.87
Perceived key stakeholders' satisfaction scores by using Likert scale	3.82	3.66	3.84	3.77
Quantitative indicators for measuring effective communications performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Reduction of written communication: measuring how much written communication is reduced as compared to previous similar projects	3.28	2.92	2.64	2.95
Variation of the number of formal letters and e-mails sent between parties per month against the number with previous similar projects	2.84	2.76	2.70	2.77
Perceived key stakeholders' satisfaction scores by using Likert scale	3.62	3.34	3.58	3.51
Quantitative indicators for measuring innovation and improvement performance	*Importance*	*Measurability*	*Obtainability*	*Mean ratings*
Cost saving resulting from innovation expressed as a percentage of total project cost	4.26	3.72	3.56	3.85
Number of innovative initiatives introduced (e.g. construction techniques, procurement approaches, management strategies)	3.80	3.80	3.80	3.80
Perceived key stakeholders' satisfaction scores by using Likert scale	3.44	3.52	3.72	3.56
Kendall's coefficient of concordance (W)			0.401	
Level of significance			0.000	

Notes: Remark 1: Rating 1=very unimportant/very difficult and 5=very important/very easy.

Table 16 Mean value of the quantitative assessment figures

KPIs	The most important quantitative indicator (QI) for each of the seven selected KPIs	Performance level									
		Poor		Average		Good		Very good		Excellent	
		Mean	CV	Mean	CV	Mean	CV	Mean	CV	Mean	CV
Time performance	Variation of actual completion time expressed as a percentage of finally agreed completion time	−11.32%	−0.54	−1.25%	−2.90	3.86%	1.18	9.91%	0.58	15.55%	0.45
Cost performance	Variation of actual project cost expressed as a percentage of finally agreed project cost	−12.50%	−0.55	−1.55%	−2.36	3.89%	1.01	8.77%	0.60	14.07%	0.43
Top management commitment	Percentage of top management attendance in partnering meetings	48.00%	0.34	63.64%	0.25	74.23%	0.19	82.73%	0.15	84.09%	0.14
Quality performance (civil works)	Average number of non-conformance reports generated per month (for civil works)	11.78	0.46	6.67	0.49	4.17	0.60	2.17	0.78	0.41	2.17
Quality performance (building works)	Average number of non-conformance reports generated per month (for building works)	25.74	0.52	14.95	0.50	10.58	0.59	5.58	0.84	1.95	1.52
Trust and respect	Perceived key stakeholders' satisfaction scores by using Likert scale	3.73	0.40	5.52	0.18	7.00	0.13	7.84	0.10	8.06	0.09
Effective communications	Perceived key stakeholders' satisfaction scores by using Likert scale	4.19	0.30	5.93	0.15	6.93	0.12	7.71	0.09	8.71	0.07
Innovation and improvement	Cost saving resulting from innovation expressed as a percentage of total project cost	0.38%	2.02	2.55%	0.60	4.28%	0.49	6.49%	0.36	8.21%	0.38

Note: CV=Coefficient of Variation.

and two subsequent rounds of Delphi questionnaire survey were carried out to validate the suitability of the identified QIs. The QI with the highest mean ratings for each of the seven weighted KPIs was finally selected. By incorporating these indicators into the evaluation process, assessors could perform their evaluation based on quantitative evidence.

However, having a set of QIs cannot fully eliminate the subjectivity of evaluation. To remedy this deficiency, another empirical questionnaire survey was conducted in order to capture the perception of professionals as to the expectation of different performance levels for each of the QIs. The results show that differences in expectation exist between the construction experts in the perceived performance level of each QI. Thus, in spite of the fact that the mean value can serve as a quick rule of thumb for evaluators to differentiate an 'average' and 'good' performance of a partnering project, it is more appropriate to identify a range of reasonable expectations for each performance level. In this research study, a simple but practical way to define the performance range is suggested. Further research studies should be carried out to establish a more scientific way to define a range for different performance levels. Since fuzzy set theory is used to model vagueness intrinsic in the human cognitive
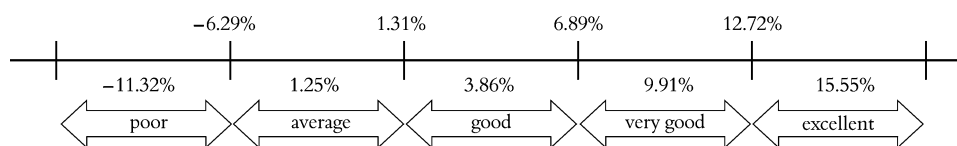

Figure 3 Range for each performance level in relation to QI of time performance

Table 17 Quantitative ranges for each of the selected QIs

The selected KPIs (with their individual weighting)	The selected quantitative indicator (QI) for each of the seven selected weighted KPIs	Quantitative range (QR) for each QI				
		Poor	Average	Good	Very good	Excellent
Time performance, with the weighting of 0.167	Variation of actual completion time expressed as a percentage of finally agreed completion time	<6.29%	−6.29% to 1.31%	1.31% to 6.89%	6.89% to 12.73%	>12.73%
Cost performance, with the weighting of 0.167	Variation of actual project cost expressed as a percentage of finally agreed project cost	<−7.0%	−7.0% to 1.17%	1.17% to 6.33%	6.33% to 11.42%	>11.42%
Top management commitment, with the weighting of 0.167	Percentage of top management attendance in partnering meetings	<55.82%	55.82% to 68.94%	68.94% to 78.48%	78.48% to 83.41%	>83.41%
Quality performance (civil works), with the weighting of 0.167	Average number of non-conformance reports generated per month (for civil works)	>9.23	5.42 to 9.23	3.17 to 5.42	1.29 to 3.17	<1.29
Quality performance (building works), with the weighting of 0.167	Average number of non-conformance reports generated per month (for building works)	>20.35	12.77 to 20.35	8.08 to 12.77	3.77 to 8.08	<3.77
Trust and respect, with the weighting of 0.167	Perceived key stakeholders' satisfaction scores by using Likert scale	<4.63	4.63 to 6.26	6.26 to 7.42	7.42 to 7.95	>7.95
Effective communications, with the weighting of 0.167	Perceived key stakeholders' satisfaction scores by using Likert scale	<5.06	5.06 to 6.43	6.43 to 7.32	7.32 to 8.21	>8.21
Innovation and improvement, with the weighting of 0.167	Cost saving resulting from innovation expressed as a percentage of total project cost	<1.47%	1.47% to 3.42%	3.42% to 5.39%	5.39% to 7.35%	>7.35%

process and it has been used to tackle ill-defined and complex problems due to incomplete and imprecise information that characterize the real-world systems (Baloi and Price, 2003), it is more appropriate to use the approach of fuzzy set theory to establish well-defined ranges of 'quantitative requirements' for each QI against the five performance levels. In fact, the development of appropriate ranges of quantitative requirements for each QI identified by using the fuzzy set theory in this research study has been completed and the overall research outcomes will be disseminated via subsequent publications.

Conclusion

This research study has established a set of quantitative indicators (QIs) and has identified a range of reasonable quantitative ranges (QRs) for the five performance levels by conducting five structured face-to-face interviews; two rounds of the second Delphi questionnaire survey; and one empirical questionnaire survey. The QIs identified with the highest mean ratings for each KPI were found to be: (1) 'Variation of actual completion time expressed as a percentage of finally agreed completion time', with the mean rating of 4.53 (for measuring time performance); (2) 'Variation of actual project cost expressed as a percentage of finally agreed project cost', with the mean rating of 4.45 (for measuring cost performance); (3) 'Percentage of top management attendance in partnering meetings', with the mean rating of 4.51 (for measuring top management commitment); (4) 'Average number of non-conformance reports generated per month' (for measuring civil works and building works), with the mean rating of 4.10 (for measuring quality performance); (5) 'Perceived key stakeholders' satisfaction scores [on trust and respect] by using Likert scale', with the mean rating of 3.77 (for measuring trust and respect performance); (6) 'Perceived key stakeholders' satisfaction scores [on effective communications] by using Likert scale', with the mean rating of 3.51 (for measuring effective communications performance); and (7) 'Cost saving resulting from innovation expressed as a percentage of total project cost', with the mean rating of 3.85 (for measuring innovation and improvement performance). After identifying a set of QIs, the quantitative ranges (QRs) for each of them against the five performance levels have been defined by taking the average value of two consecutive performance levels.

By incorporating these quantitative indicators and quantitative ranges into the evaluation process, different assessors could perform their evaluation process based on quantitative evidence. Different partnering projects can then be evaluated and compared on an objective basis with reference to this set of QIs and QRs. As a result, construction senior executives and project managers can adopt the identified QIs and QRs to measure, evaluate and upgrade the existing performance level of their partnering projects. It also enriches the current body of knowledge and understanding of both academics and practitioners in the construction industry on partnering practices to achieve outstanding partnering performance. Since the model was developed locally in Hong Kong, further research should be conducted in other geographical locations to seek their similarities and differences for international comparisons by applying the same research method.

References

Abudayyeh, O. (1994) Partnering: a team building approach to quality construction management. *ASCE Journal of Management in Engineering*, **10**(6), 26–9.

Anatharajan, T. and Anataraman, V. (1982) Development of residential areas: Delphi technique for decision making. *International Journal for Housing Science and Its Application*, **6**(4), 329–41.

Baloi, D. and Price, A.D.F. (2003) Modelling global risk factors affecting construction cost performance. *International Journal of Project Management*, **21**, 261–9.

Bayliss, R.F. (2000) Project partnering: a case study on MTR Corporation Ltd's Tseung Kwan O Extension, in Hkie *et al.*, (eds) *Proceedings of the Millennium Conference on Construction Project Management—Recent Developments and the Way Forward 2000, October*, Hong Kong, pp. 1–6.

Bayliss, R., Cheung, S.O., Suen, H.C.H. and Wong, S.P. (2004) Effective partnering tools in construction: a case study on MTRC TKE contract 604 in Hong Kong. *International Journal of Project Management*, **22**(3), 253–63.

Black, C., Akintoye, A. and Fitzgerald, E. (2000) Partnering in construction: an analysis of success factors and benefits. *International Journal of Project Management*, **18**(6), 423–34.

Bresnen, M. and Marshall, N. (2000) Motivation, commitment and the use of incentives in partnerships and alliances. *Construction Management and Economics*, **18**(5), 587–98.

Chan, A.P.C., Yung, E.H.K., Lam, P.T.I., Tam, C.M. and Cheung, S.O. (2001) Application of Delphi method in selection of procurement systems for construction projects. *Construction Management and Economics*, **19**, 699–718.

Chan, A.P.C., Chan, D.W.M. and Ho, K.S.K. (2002) An analysis of project partnering in Hong Kong. Research Monograph, Department of Building and Real Estate, The Hong Kong Polytechnic University.

Chan, A.P.C., Chan, D.W.M. and Ho, K.S.K. (2003) An empirical study of the benefits of construction partnering in Hong Kong. *Construction Management and Economics*, **21**(5), 523–33.

Chan, A.P.C., Chan, D.W.M., Fan, L.C.N., Lam, P.T.I. and Yeung, J.F.Y. (2006) Partnering for construction

excellence—a reality or myth. *Building and Environment*, **41**(12), 1924–33.

Cheung, S.O., Suen, H.C.H. and Cheung, K.K.W. (2003) An automated partnering monitoring system—Partnering Temperature Index. *Automation in Construction*, **12**(3), 331–45.

Chow, L.K. and Ng, S.T. (2007) Expectation of performance levels pertinent to consultant performance evaluation. *International Journal of Project Management*, **25**, 90–103.

Collin, J. (2002) *Measuring the Success of Building Projects—Improved Project Delivery Initiatives*, Report for the Queensland Department of Public Works, Australia.

Construction Industry Board (CIB) (1997) *Partnering in the Team: A Report by the Working Group 12 of the Construction Industry Board, UK*, Thomas Telford, London.

Construction Industry Institute (CII) (1991) *In Search of Partnering Excellence*, Publication No. 17-1, Report CII, Austin, TX.

Construction Industry Institute (CII) (1996) *Partnering: Models for Success*, Partnering Task Force Report, Construction Industry Institute, Australia.

Corotis, R., Fox, R. and Harris, J. (1981) Delphi methods: theory and design load application. *ASCE Journal of the Structural Division*, **107**(6), 1095–1105.

Cowan, C., Gray, C. and Larson, E. (1992) Project partnering. *Project Management Journal*, **22**(4), 5–12.

Crane, T.G., Felder, J.P., Thompson, P.J., Thompson, M.G. and Sanders, S.R. (1999) Partnering measures. *ASCE Journal of Management in Engineering*, **6**(4), 431–446.

Griffiths, H.B. (1993) *Mathematics of Models: Continuous and Discrete Dynamics Systems*, Ellis Horwood Ltd, New York.

Harback, H.F., Basham, D.L. and Buhts, R.E. (1994) Partnering paradigm. *ASCE Journal of Management in Engineering*, **10**(1), 23–7.

Lazar, F.D. (1997) Partnering: new benefits from peering inside the black box. *ASCE Journal of Management in Engineering*, **13**(6), 75–83.

Lazar, F.D. (2000) Project partnering: improving the likelihood of win/win outcomes. *ASCE Journal of Management in Engineering*, **16**(2), 71–83.

Li, H., Cheng, E.W.L., Love, P.E.D. and Irani, Z. (2001) Co-operative benchmarking: a tool for partnering excellence in construction. *International Journal of Project Management*, **19**(3), 171–9.

Lo, T., Wong, P.S.P. and Cheung, S.O. (2006) Using Balanced Scorecard (BSC) approach to measure performance of partnering projects. *Survey and the Built Environment*, **17**(1), 45–57.

Manoliadis, O., Tsolas, O. and Nakou, A. (2006) Sustainable construction and drivers of change in Greece: a Delphi study. *Construction Management and Economics*, **24**(2), 113–20.

Mohr, J. and Spekman, R. (1994) Characteristics of partnering success: partnering attributes, communication behaviour, and conflict resolution techniques. *Strategic Management Journal*, **15**(2), 135–52.

Moore, C., Mosley, D. and Slagle, M. (1992) Partnering guidelines for win-win project management. *Project Management Journal*, **22**(1), 18–21.

Morrison, F. (1991) *The Art of Modeling Dynamics Systems: Forecasting for Chaos, Randomness and Determinism*, John Wiley & Sons, New York.

Neter, J., Kutner, M., Nachtsheim, C. and Wasserman, W. (2005) *Applied Linear Statistical Models*, 5th edn, Boston, McGraw-Hill.

Norusis, M.J. (2005) *SPSS 14.0 Statistical Procedures Companion*, Prentice Hall, Upper Saddle River, NJ.

Saito, M. and Sinha, K. (1991) Delphi study on bridge condition rating and effects of improvements. *Journal of Transportation Engineering*, **117**(3), 320–34.

Schwarz, N. (1996) *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*, Lawrence-Erlbaum, Mahwah, NJ.

Thompson, P.J. and Sanders, S.R. (1998) Partnering continuum. *ASCE Journal of Management in Engineering*, **14**(5), 73–8.

Weisberg, S. (2005) *Applied Linear Regression*, 3rd edn, Hoboken, NJ, John Wiley & Sons.

Yeung, J.F.Y., Chan, A.P.C., Chan, D.W.M. and Li, L.K. (in press) Development of a Partnering Performance Index (PPI) for construction projects in Hong Kong: a Delphi study. *Construction Management and Economics*.

Zhao, F. (2002) Measuring inter-organizational partnership: the challenge of cultural discrepancy, in A. Nealy and A. Walters (eds) *Proceedings of the 3rd International Conference on Theory & Practice in Performance Measurement*, 17–19 July, The World Trade Centre, Boston, USA, pp. 627–634.