ESTABLISHING A QUANTITAIVE PERFORMANCE EVALUATION FRAMEWORK THROUGH FUZZY MEMBERSHIP FUNCTIONS

Lai-Kit Chow [1] **and S. Thomas Ng** [2]

[1]*Research Student, Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong.*
[2]*Associate Professor, Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong.*

Past practice of proposal-oriented Prequalification and Selection (PS) of Engineering Consultants (EC) is found problematic. Nowadays, public sectors and quasi-governmental organizations in many advanced countries emphasize on gauging ECs' performance at different project stages to congregate a set of Consultant Performance Evaluation (CPE) data for future PS. It is believed that the performance of ECs in previous projects could directly reflect its real strength and ability for succeeding assignments. To further enhance the reliability, it is desirable to merge all the CPE data from different clients into an integrated repository. A universal set of CPE criteria and corresponding Quantitative Indicators (QI) were compiled in previous research to provide a common evaluation platform. However, evaluation variations still exist as the QI require assessors' subjective value judgement based on their own expectations. The paper summarized the CPE criteria and QI identified in previous research, and illustrated the discrepancies arisen from applying subjective judgement in the evaluation. To tackle this deficiency, the paper further described how Quantitative Requirements (QR) of each QI against different performance levels could be established, by utilizing modified horizontal approach to develop fuzzy membership functions for each QI. Finally, an objective evaluation framework for CPE is proposed.

Keywords: consultant performance evaluation, quantitative indicators, quantitative requirements, modified horizontal approach, fuzzy membership functions.

INTRODUCTION

Construction projects are complex and risky, and any clients who do not have a good understanding on the design and management of a project may suffer severe loss. To safeguard the interests of the clients, Engineering Consultants (ECs) are appointed to provide professional advice, carry out the design and administer the contract (Cooley, 1994). The engagement of capable ECs could, therefore, improve the chance of delivering a project on time and within budget (Hattan and Lalani, 1997).

Researchers suggested that the practice of adopting the two-envelope system (Ng *et al*, 2001) for consultant selection is found unreliable and problematic, as the system relies heavily on the attractiveness of technical and financial proposals to determine whether the ECs could complete their services effectively and satisfactorily. Yet, the true proficiency of an EC may not necessarily correlate to the merits of the technical proposal and/or the consultancy fee they charge.

[1] chowlk@hkusua.hku.hk

Chow, L-K and Ng, S T (2005) Establishing a quantitaive performance evaluation framework through fuzzy membership functions. *In:* Khosrowshahi, F (Ed.), *21st Annual ARCOM Conference*, 7-9 September 2005, SOAS, University of London. Association of Researchers in Construction Management, Vol. 2, 1015-24.

The quality of the service outcomes in previous projects is believed to be the most confiding indicator to reflect one's real strength and ability. Therefore, greater emphasis should be allocated on the past performance when the design team is formulated (Ingram and Peltier, 2001). With an aim to compile a set of reliable past performance record for the decision support in future EC selection, an appropriate Consultant Performance Evaluation (CPE) system is inevitable. Nowadays, many clients measure the performance of ECs regularly throughout the entire project duration.

Previous relevant studies on CPE include a performance measurement framework which focuses on the non-financial criteria (Geanuracos and Meiklejohn, 1993), the CPE for architectural practices (Cheung *et al*, 2002), and the task and contextual performance of ECs in design and build projects (Ling, 2002, 2003).

In Hong Kong, public (KCRC 2004; HKHA, 1997; Works Bureau, 2001, 2002) and quasi-governmental (MTRC, 2003) clients carry out CPE in a regular manner throughout the entire project phases (feasilibilty, design, bidding, construction and maintenance). The evaluation results are paramount important which contribute up to 30% of the overall technical assessment score during ECs selection. Therefore, it is desirable to broaden the coverage of the performance data by assembling the CPE results from different clients, in order to have a more objective and comprehensive view on the performance of ECs. However, without a mutually acceptable CPE standard, it is arduous to merge and compare the performance data from different clients as variations may occur. Therefore, it is inevitable to establish a rational framework for CPE.

Since the consultancy services could cover a number of stages within the project life cycle, it is necessary to determine which stage(s) would have the greatest impact to the project. According to Pilcher (1994), 80% of the construction costs are taken when the sketch design is formulated, and any design errors and omissions, if undetected or unresolved, could undoubtedly be the origins of serious claims and reworks once the construction begins. Burati *et al* (1992) advocated that the cost of design errors exceeds that attributable to those generated by construction (9.5% as opposed to 2.5% of the total project cost). This indicates that the design stage has crucial influence to the project, and the performance of ECs at the design stage should therefore be carefully scrutinized.

To unveil the criteria used for measuring the design performance, Chow and Ng (2003) and Ng and Chow (2004b) conducted a comprehensive review on the evaluation criteria currently adopted in UK, US, Australia and Hong Kong, and eight predominant CPE criteria that reflect the performance of ECs at the design stage were identified. The establishment of a universal set of evaluation criteria is only the first step towards a common platform for CPE information exchange. Yet, the CPE criteria are fuzzy in nature which may exhibit a significant level of vagueness in the semantic meanings / interpretations (Zimmermann, 1991) for each CPE criterion. There are some research studies which emphasize on defining concrete semantic meanings for evaluating the effectiveness of product from the discrepancies between designers' and clients' perceptions (*cf*: Shang *et al*, 2000; Sagawa and Inooka, 2002). A well-defined set of semantic description(s) or systematic guideline(s) for the evaluation criteria is crucial as it would help prevent assessors from applying their own interpretation during the evaluation process. Furthermore, the guidelines should be quantitative in nature to achieve an evidential-based evaluation instead of a qualitative one. With these aims, Chow and Ng (2004) further developed eighteen

Quantitative Indicators (QI) which specify the quantitative interpretation(s) for each CPE criterion and a rational evaluation guideline is thus established. The CPE criteria and the corresponding QI identified in previous research are presented in Table 1.

Table 1: CPE criteria and QI

Criteria (Design Stage)	*Corresponding Quantitative Indicator(s)*
Compliance and Understandings to Clients' Brief	o Percentage of consultancy services that adhere to the design brief
Compliance to Legislative Requirements	o Number of design submission(s) required to get the design approval from all related government departments
Identification of Client's Requirements and Project Objectives	o Percentage of consultancy services that fulfill the technical requirements of the project o Percentage of consultancy services that fulfill the financial requirements of the project o Percentage of consultancy services that fulfill the time requirements of the project
Quality of Design	o Percentage of ECs' design solutions that fulfill the technical standard of the project o Percentage of ECs' design solutions that fulfill the financial standard of the project o Percentage of ECs' design solutions that fulfill the environmental standard of the project o Percentage of ECs' design solutions that fulfill the safety standard of the project o Percentage of ECs' design solutions that fulfill the quality standard of the project o Percentage of project risks that effectively covered (avoided / minimized) in ECs' design recommendations o Percentage of ECs' design solutions that comprehensively examine all constructability aspects
Availability of Innovative and Alternative Solutions	o Percentage of ECs' design elements that carefully examine all possible solutions o Percentage reduction of construction duration from the client's expected value o Percentage reduction of construction cost from the client's expected value
Approach to Overall Cost Effectiveness	o Percentage reduction of all project resources from the client's expected value
Quality of Drawings / Documents	o Percentage of drawings / documents that are comprehensive, clear and well-defined
Adequacy of Cost Estimate	o Percentage deviation of the overall project cost from client's expected value

The universal set of CPE criteria could make the exchange of CPE data possible and the QI further reduce the vagueness in determining the semantic meaning(s) or interpretation(s) for each CPE criterion. However, the QI are fuzzy and requires assessors' subjective value judgement (Nahapiet and Nahapiet, 1985; Hamilton, 1987) based on their personal expectation on each QI. The excessive reliance on assessors' own expectation could significantly affect the evaluation result (Chow and Ng, 2004).

Fuzzy set theory (Zadeh, 1965) can be applied to model the QI in CPE. The first step towards the development of a fuzzy CPE is to interpret the CPE criterion in a logical way which has been achieved by Chow and Ng (2004) through the identification of QI. The subsequent step is to define the fuzzy membership functions for each QI. With the fuzzy membership functions, Quantitative Requirements (QR) for each QI

against different performance levels could be identified. The evaluation could then be carried out based on the QR to prevent the assessors from applying their subjective value judgement. This paper aims to describe the research methodology and data analysis technique to develop the QR by utilizing modified horizontal approach (Ng *et al*, 2002) to develop fuzzy membership functions of each QI. Finally, an objective evaluation framework is proposed.

EVALUATION DISCREPANCIES ARISEN FROM SUBJECTIVE VALUE JUDGEMENT

The fuzziness of QI is a reason for assessors to apply their value judgement based on their own expectation and this could lead to serious evaluation discrepancies. In order to illustrate the effects caused by subjective judgments, a hypothetical case was set up for two independent experts to evaluate. The criterion "identification of client's requirements and objectives" was chosen for this exercise, and Tables 2a & 2b show the details of the hypothetical case and the results of evaluation respectively.

Table 2a: Assessors' personal expectation on each QI against three performance levels

CPE Criteria (Design Stage)	***Assessor A's Expectation***			***Assessor B's Expectation***		
Identification of client's requirements and objectives	E	A	P	E	A	P
o Percentage of consultancy services that fulfill the technical requirements of the project	85%	75%	70%	75%	70%	65%
o Percentage of consultancy services that fulfill the financial requirements of the project	85%	80%	70%	80%	70%	65%
o Percentage of consultancy services that fulfill the time requirements of the project	90%	80%	75%	80%	75%	70%

Note: E-excellent; A-average; P-poor

Table 2b: Evaluation conducted by two assessors based on personal expectation

CPE Criteria (Design Stage)	Consultant's Actual Service Outcome	Grade Given by Assessor A	Grade Given by Assessor B
Identification of client's requirements and objectives			
o Percentage of consultancy services that fulfill the technical requirements of the project	70%	P	A
o Percentage of consultancy services that fulfill the financial requirements of the project	80%	A	E
o Percentage of consultancy services that fulfill the time requirements of the project	80%	A	E
Translated score		1+3+3 = 7	3+5+5 = 13

Note: E-excellent; A-average; P-poor

The CPE system of a quasi-governmental organization in Hong Kong was employed in this exercise. The system makes use of a three-grade approach for CPE which categorizes ECs' performance into three different levels (i.e. excellent, average and poor). The grades given by the assessors were then be translated into the performance score of the ECs where 5 represents "excellent performance", 3 denotes "average performance", and 1 for "poor performance".

In this exercise, there were no systematic and quantitative guidelines for the assessors to follow when they performed their assessment, which means subjective value judgment was needed. It is reasonable that the grade to each QI would be based on assessors' own expectation which is illustrated in Table 2b. Since assessor A expects a poorly performed EC could only fulfill 70% of the technical requirements of the project, a grade "poor" was assigned by assessor A to the EC against the QI "percentage of consultancy services that fulfill the technical requirements of the project". Instead, a grade "average" was given by assessor B to the same QI.

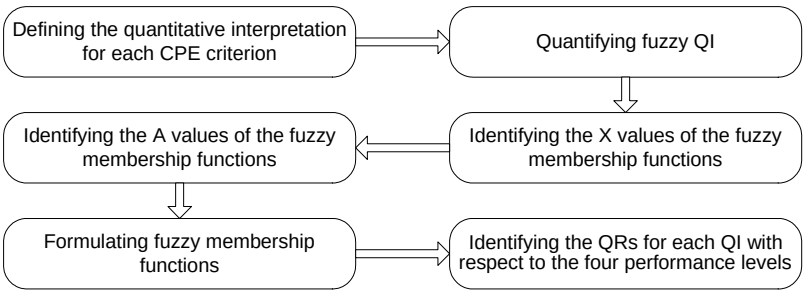
It is found that the translated performance scores for the ECs from assessor A (translated score = 7) and assessor B (translated score = 13) are quite different. It is obvious that a higher score can normally be sought from assessor B. As a result, applying subjective value judgment to the QI is undoubtedly problematic, as it could seriously affect the reliability of the CPE information.

RESEARCH METHODOLOGY AND DATA ANALYSIS TECHNIQUE FOR THE DEVELOPMENT OF QR

In order to eliminate the evaluation variations resulted by subjective judgement, QR of each QI against different performance levels should be established. Fuzzy set theory is deemed appropriate to model the fuzzy QI, and QR could be established by developing fuzzy membership functions for each QI. Four common methods are normally used for establishing the fuzzy membership functions including (i) the horizontal approach (Godal and Goodman, 1980; Bharathi-Devi and Sarma, 1985), (ii) the vertical approach (Civanlar and Trussel, 1986), (iii) the pairwise comparison method (Saaty, 1980) and (iv) the membership function estimation approach with the aid of probabilistic characteristics (Dubois and Prade, 1983).

A modified horizontal approach (Ng *et al*, 2002) based on an amalgamation of the horizontal approach and the graphical approach (Bandemer and Gottwald, 1995) is adopted in this paper for developing the fuzzy membership functions. This approach is simple and accurate as the horizontal approach allows the final outcome to be derived from simple probability functions. Therefore, clients could establish their own membership functions to carry out the CPE conveniently. In addition, the horizontal approach could allow the computation of an optimal value of k (number of bands) which is crucial to the accuracy of estimation (Bharathi-Devi and Sarma, 1985). The graphical approach can further overcome the problem of discontinuity in the transition from full membership to absolute exclusion in pure horizontal methods (Otnes and Enochson, 1972). Modified horizontal approach was first proposed and adopted for analysis by Ng *et al* (2002). The procedures of defining the QR of each QI against different performance levels are shown in Figure 1:

Figure 1: Procedures in defining the QR



Defining the Quantitative Interpretation for each CPE Criterion

The CPE criteria should first be transformed into meaningful QI. The eighteen QI (Figure 1) pertinent to the eight CPE criteria for the design stage of a project identified by Chow and Ng (2004) are utilized as the quantitative interpretation in developing the fuzzy membership functions.

Quantifying Fuzzy QI

Experts' opinions expectation on each QI against different performance levels should be sought to build up the pool of data. The design of a research questionnaire could effectively meet this purpose. In the questionnaire, experts should be requested to apply numerical figures (f_0) for each QI with respect to different performance levels, in order to gather their expectations for further data analysis. The number of performance levels included in the research should be carefully considered, but basically three (excellent, average and poor), four (excellent, good, average and poor) or five (excellent, good, average, below average and poor) grading systems are commonly adopted by clients.

Identifying the *X* values of the Fuzzy Membership Functions

A fuzzy membership function could basically be formulated by two values: X and A. X represents the value in the universe of discourse that defines the fuzzy set while A stands for the degree of membership of that fuzzy set. X_i values are defined as the means of bands B_i (i=1,2,…, k), where B_i (i=1,2,…,k) are the bands of values f_0 given by the respondents of the questionnaire survey to the QI against different performance levels. The X_i values are defined according to the lowest and greatest values of f_0 for each QI and the number of bands k. Bharathi-Devi and Sarma (1985) defined an equation to calculate the number of bands for the estimation:

$$k = 1.87(N-1)^{\frac{2}{5}} \qquad \text{----------} \quad (1)$$

where N is the total number of valid replies to the corresponding QI.

Identifying the *A* values of the Fuzzy Membership Functions

The degree of membership A_i could be computed according to the following equation:

$$A_i = \frac{n(B_i)}{n_{\max}} \qquad i=1,2,\ldots,k \qquad \text{----------} \quad (2)$$

In this equation, $n(B_i)$ corresponds to the number of valid replies that have the values of f_0 belonging to a certain band B_i, and n_{max} represents the maximum value of all the $n(B_i)$ with i=1,2,…,k.

In order to examine whether the estimation of membership is valid, the standard deviation, *st.d.*(*A*) should be calculated. The result could be considered acceptable in case the *st.d.*(A_i) have a lower value than A_i calculated in Eqn. 2. The *st.d.*(*A*) were found out by the following equation:

$$st.d.(A_i) = A_i \times \left(\frac{1 - A_i^{\frac{1}{2}}}{N} \right) \qquad i=1,2,\ldots,k \qquad \text{----------} \quad (3)$$

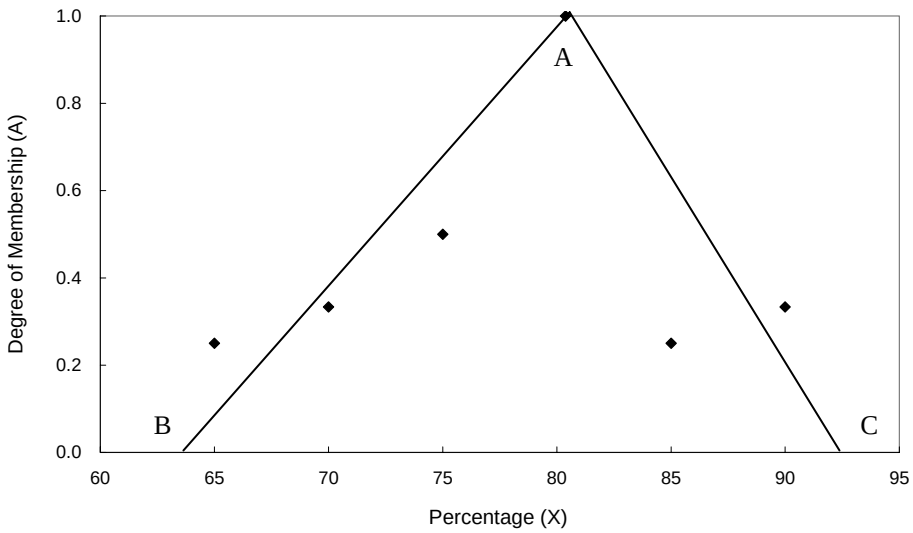
Formulating Fuzzy Membership Functions

The fuzzy membership functions developed could be first presented in a tabular manner (see Table 3 as an example). Based on the *X* and *A* values, scatter diagrams for the membership functions could be plotted (see Figure 2 as an example), with the horizontal axis representing the *X* values and the vertical axis representing the *A* values. After the degree of membership for all bands k had been determined, best-fit lines should then be plotted to join all the discrete points (line *AB* and *AC* in Figure 2). This could be achieved by using the best-fit line plotting function in EXCEL™. The best-fit lines should pass through the point with full membership (point *A*). Therefore, the transformation of coordinates should be conducted before constructing the best-fit lines in EXCEL™. In this process, the point with full membership (point *A*) should first be shifted to the origin and all other points should be transformed accordingly. Best-fit lines could then be constructed in EXCEL™ by activating the plotting criterion of passing through the zero intercept.

Table 3: *X* and *A* values of the membership functions

Percentage (X)	65	70	75	80.38	85	90
Degree of membership (A)	0.250	0.333	0.500	1.000	0.250	0.333

Figure 2: Scatter diagram of the membership function



Identifying the QR of each QI pertinent to Different Performance Levels

After plotting the best-fit lines of each QI against the different performance levels on the same graph (Figure 3), the intersections of the best-fit lines between two

consecutive performance levels indicates a same degree of membership occurs for both performance levels (Points *A*, *B* and *C* in Figure 3). For example, point *A* in Figure 3 is the intersection of the best-fit lines of average and good performance levels with the average performance level best-fit line having a decreasing degree of membership and the good performance best-fit line having an increasing degree of membership. Therefore, it is logical to choose these intersecting points to identify the QR of each QI against different performance levels. For example, the QR for the fuzzy membership functions in Figure 3 could be defined as follows:

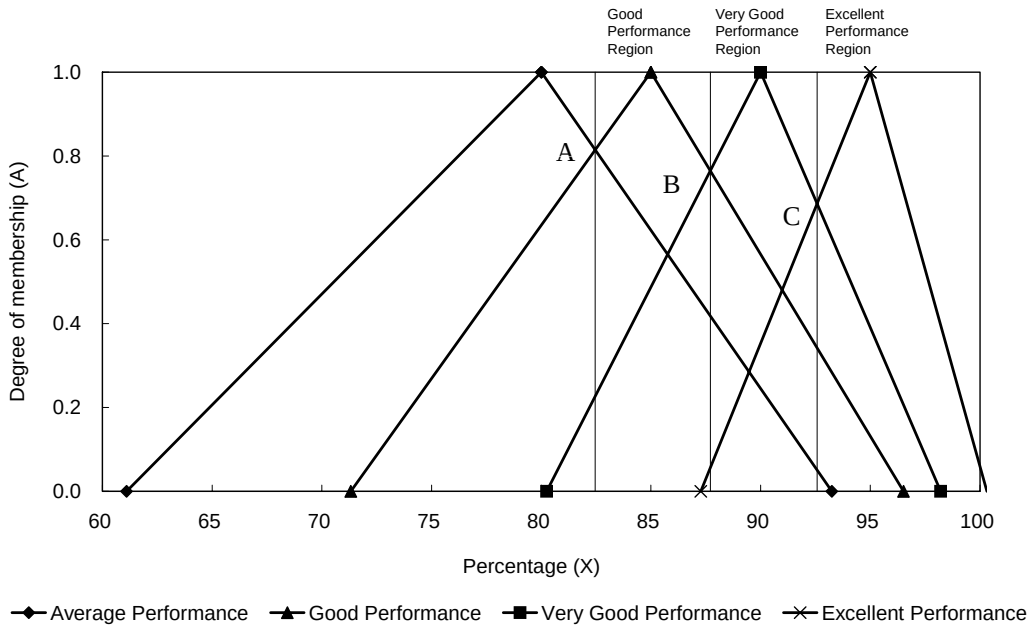
QR for *excellent* performance level: C_x X 100%

QR for *very good* performance level: B_x $X < C_x$

QR for *good* performance levels: A_x $X < B_x$

where A_x, B_x and C_x represent the X values of points *A*, *B* and *C* respectively.

Figure 3: QR defined through fuzzy membership functions



CONCLUSION

In order to have a comprehensive view on the past performance of ECs, it is desirable to have a common platform to facilitate the exchange of CPE results among different clients. In addition, a well-defined set of interpretations for each evaluation criterion should also exist to prevent any evaluation discrepancies. In this paper, eight CPE criteria pertinent to the performance of ECs at the design stage and eighteen corresponding QI have been identified from previous research.

However, the fuzziness of QI requires assessors' own value judgement during evaluation. In order to tackle this deficiency, a modified horizontal approach is proposed in this paper to develop the fuzzy membership functions of each QI, thus establishing the QR of each QI against different performance levels.

The QR allow the assessors to perform their appraisal based on objective evidence instead of their own value judgement. Consequently, the QR eliminate the evaluation discrepancies arisen from employing assessors' own expectation in CPE. The CPE criteria, QI and QR together form an objective evaluation CPE framework to gauge the design performance of ECs. This framework enhance the reliability of past performance data, and thus improve the chance of prequalifying and selecting suitable and capable ECs for future consultancy assignments.

REFERENCES

Bandemer, H. and Gottwald, S. (1995) Fuzzy sets, fuzzy logic, fuzzy methods with applications, Wiley, Chichester.

Burati, J., Farrington, J. and Ledbetter, W. (1992) Causes of quality deviation in design and construction, Journal of Construction Engineering and Management, 118(1), 34-49.

Bharathi-Devi, B. and Sarma, V.V.S. (1985) Estimation of fuzzy memberships from histograms, Information Sciences, 35, 43-59.

Cheung, F.K.T., Leung, J.F.K. and Skitmore R.M. (2002) Multi-criteria evaluation model for the selection of architectural consultants, Construction Management and Economics, 20(7), 569-580.

Chow, L.K. and Ng, S.T. (2003) Performance-based evaluation for engineering consultants – a study on the assessment criteria, Proceedings, ARCOM 19th Annual Conference 2003, September 3-5, University of Brighton, Brighton, 433-442.

Chow, L.K. and Ng, S.T. (2004) A framework for evaluating the performance of engineering consultants in Hong Kong, Proceedings, 3rd Scientific Conference on Project Management "Clustering in Construction Project Management", September 24-25, Thessaloniki, Greece, 231-238.

Civanlar, M.R. and Trussel, H.J. (1986) Constructing membership functions using statistical data, Fuzzy Sets and Systems, 18(1), 1-13.

Cooley, M.S. (1994) Selecting the right consultants, HR Magazine, 39(8), 100-103.

Dubois, D and Prade, H. (1983) Unfair coins and necessity measures: towards a probabilities interpretation of histograms, Fuzzy Sets and Systems, 10(1), 15-20.

Geanuracos, J. & Meiklejohn, I. (1993) Performance measurement, the new agenda, London: Business Intelligence.

Godal, R.C. & Goodman, T.J. (1980) Fuzzy sets and borel, Transactions on Systems, Man, and Cybernetics, 10(10), 637-640.

Hamilton, I.W. (1987) Developing expert systems for management applications, Building Cost Modelling and Computers, P.S. Brandon (ed.), E. & F. N. Spon, London, 441-451.

Hattan, D.E. and Lalani, N. (1997) Selecting the right consultant team, Institute of Transportation Engineering Journal, 67(9), 40-46.

HKHA (1997) Consultant reporting system, Housing Department, Hong Kong Housing Authority, Hong Kong

HM Treasury (1997) Appointment of consultants and contractors, Procurement Guidance No. 3, Procurement Group, UK Government Procurement Group, London.

Ingram, D.C. and Peltier, B. (2001), MTO adopts performance-based consultant selection system, Engineering Dimensions, March/April, 11-15.

KCRC (2004) Guidance notes for consultants' interim / final performance appraisal report, Kowloon-Canton Railway Corporation, Hong Kong.

Ling, Y.Y. (2002) Model for predicting performance of architects and engineers, Construction Engineering and Management, 128(5), 446-455.

Ling, Y.Y. (2003) A conceptual model for selection of architects by project managers in Singapore, International Journal of Project Management, 21(2), 135-144.

MTRC (2003) Consultants/Contractors Performance Report: General Project Procedure, Mass Transit Railway Corporation, Hong Kong.

Nahapiet, H. and Nahapiet, J. (1985) The management of construction projects: case studies from the UK and USA, Chartered Institute of Building, Ascot.

Ng, S.T. and Chow, L.K. (2004a) Evaluating engineering consultants' general capabilities during the pre-selection process – a Hong Kong Study, Engineering, Construction and Architectural Management, 11(3),150-158

Ng, S.T. and Chow, L.K. (2004b) Framework for evaluating the performance of engineering consultants, Journal of Professional Issues in Engineering Education and Practice, ASCE, 130(4), 280-288

Ng, S.T., Kumaraswamy, M.M. and Chow, L.K. (2001) Selecting consultants through combined technical and fee assessment – a Hong Kong study, Proceedings, ARCOM 17th Annual Conference 2001, Sept. 5-7, University of Salford, Salford, Manchester, 639-648.

Ng, S.T., Luu, D.T., Chen, S.E. and Lam, K.C. (2002) Fuzzy membership functions of procurement selection criteria, Construction Management and Economics, 20, 685-696.

Otnes, R. and Enochson, L. (1972) Digital Times Series Analysis, Wiley, New York.

Pilcher, R. (1994) Project cost control in construction, Blackwell Scientific Publications, Oxford, UK.

Saaty, T.L. (1980) The analytical hierarchy processes, McGraw, NewYork.

Sagawa, K. and Inooka, H. (2002) Ride quality evaluation of an actively-controlled stretcher for an ambulance, Journal of Engineering in Medicine, 216 (4), 247-256.

Shang, H.H., Ming, C.C. and Chien, C.C. (2000) A semantic differential study of designers' and users' product from perception, International Journal of Indutrial Ergonomics, 25, 375-391.

Ullman, D.G. (2001) Robust decision-making for engineering design, Journal of Engineering Design, 12(1), 3-13.

Works Bureau (2001) Reporting and management of consultants' performance, Technical Circular No. 29/2001, Works Bureau, HKSAR Government, Hong Kong.

Works Bureau (2002) Consultants' performance report and performance index system, Technical Circular No. 27/2002, Works Bureau, HKSAR Government, Hong Kong.

Zadeh, L. A. (1965) Fuzzy sets, Information and Control, 8, 338-353.

Zimmermann, D.G. (1991) Fuzzy set theory and its applications, 2nd edition, Kluwer, Dordrecht.