

Are the Experts Really Experts? A Cognitive Ergonomics Investigation for Project Estimations

Budi Hartono^{1*}, Bhery Aditya Saputra¹

Abstract: Uniqueness is a major characteristic of any project systems. Hence it is virtually infeasible for project analysts to utilize data from past projects as references for subsequent project planning and scheduling. Most project analysts would then depend on intuition, gut feeling and experiences to develop quantitative models for project scheduling and analysis which, according to past studies, is prone towards systematic errors. This study attempts to investigate the performance of both 'experts' and 'non-experts' when utilizing their cognitive capability to estimate project durations in group/non-group settings. A cognitive ergonomics perspective -which views human capability to make judgment as rationally bounded - is utilized in this investigation. An empirical approach is used to inquiry data from 'projects' on which 'experts' and 'non-experts' are required to provide prior estimate on project durations. The estimates are then gauged against the actual duration. Results show that some systematic cognitive judgmental errors (biases) are observable for both experts and non-experts. The identified biases include: anchoring bias as well as accuracy bias.

Keywords: Project management, cognitive ergonomics, estimate bias, expert, group.

Introduction

Project management methodology becomes more and more popular to be utilized by various organizations to deliver strategic objectives such as: change management, product development, research and development, and construction. One key factor attributable to project success is the effectiveness of project planning and preparation (Dvir [6]; Dvir *et al* [7]). Due to the project uniqueness (PMI [22]), however, limited historical data is accessible for project analysts to carry out quantitative analysis of project preparation and planning. Accordingly, in most occasions human analysts would rely on past experience and subjective judgment (Akintoye and MacLeod [1]; Shen [23]; Simu [25]; Wood and Ellis [29]) in project planning – especially during the project estimation stage.

Extensive empirical studies in cognitive psychology suggest that while human intend to be objective and rational, the rationality is bounded (Bazerman [2]; Simon [24]). Hence human estimations, according to studies, tend to prone towards systematic judgmental errors or biases (Gigerenzer and Goldstein [13]; Hartono and Yap [14]; Hastie and Dawes [15]; Kahneman [16]). From a cognitive ergonomics perspective, this could be seen as a condition where a cognitive task of project estimation needs to be carried out by human with cognitive limitations leading

to less than rational estimations. If such a condition persists, a poor project planning would be inevitable which in turn could result in poor project performance.

Common wisdom would suggest that in order to improve performance of human judgment, experts are utilized (e.g. Cooke and Goossens [5]; O'Hagan *et al* [21]). Another method to improve the performance is by conducting the exercise in a group setting (Boy *et al* [3]; McGraw and Seale [20]). However, the efficacy of such interventions is less than clear. Vose [28] for instance argues that in some occasions, a person deemed expert for this particular cognitive assignment is not the real experts – i.e. inexpert experts. Other studies suggest that even real experts are not immune towards judgmental biases (Cleaves [4]).

Numerous past empirical studies had been carried out with aims at identifying possible judgmental biases with positive results. Systematic errors related to judgment and estimations (such as: anchoring effects, over-optimism bias) were pervasively identifiable in various settings (Mak and Raftery [19]; Tversky and Kahneman [27]). Some research opportunities, however, are identifiable. Firstly, most past empirical studies were conducted within general management settings. Accordingly, similar studies to identify possible existence judgmental biases are required within project management settings. Secondly, most studies were carried out in a developed country setting, which may be different from the developing country environment. Thirdly, as stated earlier, the effects of experts/non-experts and indi-

¹ Department of Mechanical and Industrial Engineering, Industrial Engineering Program, Universitas Gadjah Mada, Jl. Grafika 2, Yogyakarta 55231. Email: boed@gadjahmada.edu

* Corresponding author

vidual/group settings toward judgmental errors need further clarifications by means of empirical study.

Accordingly, this study aims at examining performance of human estimators for a cognitive task of estimating project durations within various settings. Specifically, this study investigates whether expert/non-expert and group/non-group settings do affect the performance of project estimations in terms of biases reductions. The biases under observation are: anchoring effects and accuracy bias. Anchoring effect is a systematical error which occurs when estimation by a project estimator is affected by irrelevant cues presented prior to the estimation process. The cues become anchors on which a decision maker would base the estimate. Anchoring effect is observable when the decision maker does not make sufficient adjustment from the base estimate. This study would provide empirical evidence to the possible cognitive limitations within a specific setting of project management of a developing country – in this case Indonesia. If such limitations leading to biases exist, possible remedies (i.e. procedures, protocol, tools) need to be developed. This study lays a ground work to a development of a joint cognitive system assigned to a specific task of project estimations. Such a system consists of (groups) of human estimators, the estimation tasks, and other supporting artifacts (protocols, methods, tools) which assist human estimators to achieve better performance in a cognitive task of project estimations.

Methods

An experiment was administered for the study. Subjects were forty (40) undergraduate students who were given both individual and group assignment using Lego. Prior to the experiment, a screening procedure is applied to ensure that subjects do not have prior knowledge and experience in building Lego blocks and developing Lego software. Each group consists of two students which were randomly selected. Four rounds of mini projects were given to each group in succession of random sequence. Hence while all groups would have the same four mini projects, they may experience difference sequence. This protocol is utilized to reduce possible systematic noises from environment by means of randomization. Each mini project consisted of a slightly different task of: individual estimation for the project duration, group estimation of the project duration, and the actual group project execution. No formal feedback of information from preceding mini projects is given to the subjects during the experiment.

To improve external validity (Levin *et al* [17]), a protocol following the induced value theory (Friedman and Cassar [12]) is applied by providing

Table 1. Summary of design of experiment

Design of experiments	
Subjects	Undergraduate students
Total number of subjects	40
Total number of groups	20 (a group of two)
Number of miniprojects	four (4) rounds for each group
Sequence of miniprojects	Random
Anchoring stimuli	Once for each group
Assignment on each mini project	
a. Individual estimation for the forthcoming project	
b. Group estimation for the forthcoming project	
c. Actual project execution by the respective group	
Protocol to improve external validity	Monetary reward utilizing induced-value theory

rewards of book vouchers to subjects who perform best during the mini projects. The fact that students are having lower opportunity cost could also improve the external validity. Table 1 shows a summary of the design of experiment.

Results and Discussions

Three cognitive phenomena are observable from the experiments. Those are the learning curve, the anchoring effect, and the accuracy bias. To investigate the performance of individual and group estimations, sub-group analysis were carried out. An additional analysis is added to observe whether a simple average of individual estimations would provide different performance from individual and group estimations.

Learning Curve

Figure 1 (a) shows the result for the average time required to actually complete the projects for round one to four. The data is compared with the average values of individual and group estimations respectively. Figure 1 (b) depicts the estimation accuracy for individual and groups. The calculation is carried out by computing the average of the difference between project estimations and actual durations, respectively.

Figure 1 (a) clearly indicates that learning curve exists during project execution. On average, subjects in the experiment require less time for completing projects than those for the succeeding projects. A significant amount of time reduction is observable especially for the round 1 to round 2, while a much smaller time reduction is observable for the subsequent orders. The finding confirms a prominent assertion of learning curve as extensively reviewed by Yelle [30]. Based on the result, subjects are classified as ‘non-experts’ during the first round of experiment and ‘experts’ for the second round onwards for the subsequent analysis.

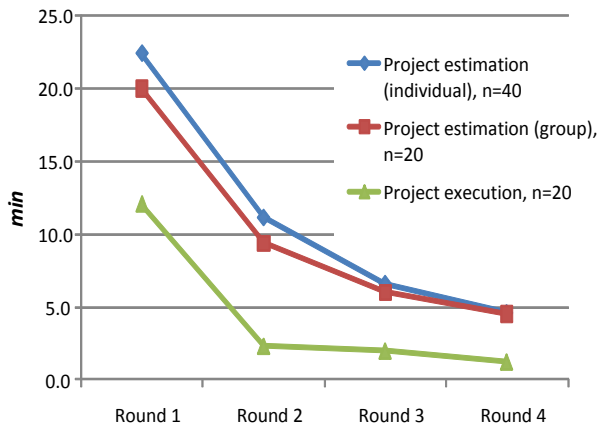


Figure 1(a). Learning curve

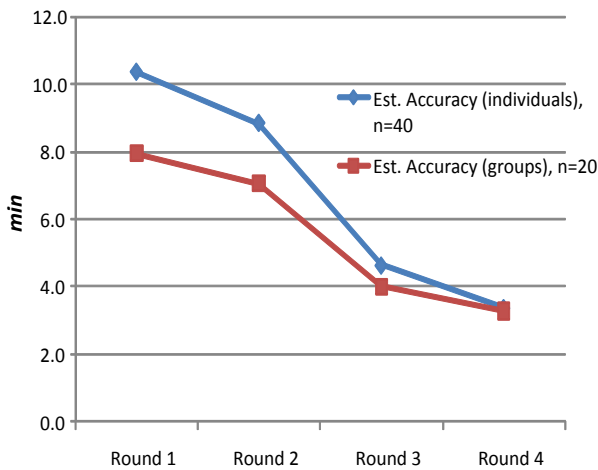


Figure 1 (b). Estimation inaccuracy

Figure 1 (a) also shows that both individual and group subjects provide estimates which comply with the learning curve concept. A follow-up questionnaire shows, however, that majority of subjects do not aware of the concept of learning curve despite its application when conducting project estimations.

From the perspective of estimate accuracy, a learning curve phenomenon is also observable. As can be seen in Figure 1 (b) accuracy improves over time as indicated by the smaller inaccuracy. It applies for both individual and group estimations. Figure 1 (b) also visually indicates that group estimations are generally more accurate than individual estimations. Individual estimates seem to be more conservative than group estimates. Such accuracy gaps decrease over time.

Anchoring Effects

Anchoring stimuli is given to the respondents in individual and group settings prior to a (randomly selected) particular round of mini projects. Each

individual or group of respondents is given with an irrelevant cue which would become the ‘anchoring stimuli’ prior to the estimation assignment. The cue is in the form of a request to provide a duration estimate of an activity which is unrelated yet similar to the mini project – in this case: time required to travel from their respective residence to the laboratory. This procedure is followed by the actual task of estimating the mini project duration.

Anchoring effect is said to be observable when there is a significant correlation between the estimates of ‘anchoring stimuli’ and estimates of the mini projects. Figure 2 depicts a pictorial representation of the experiment result for a particular scenario of individual, non-expert estimators. It is shown that, across the individual and non expert respondents, a correlation between the duration estimate of the ‘anchoring stimuli’ and this of the mini project is visually observable.

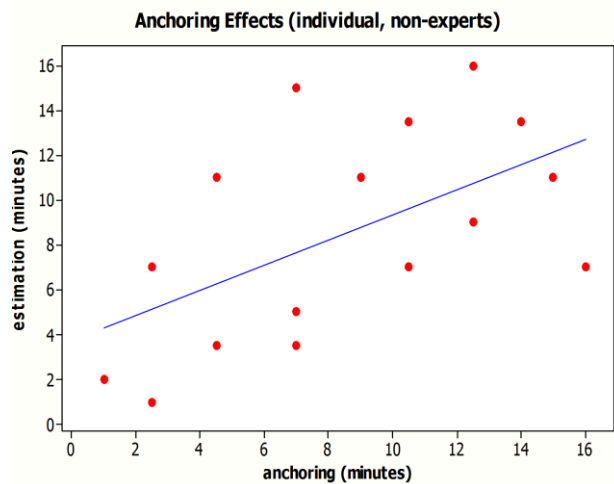


Figure 2. Anchoring stimuli vs. project estimates

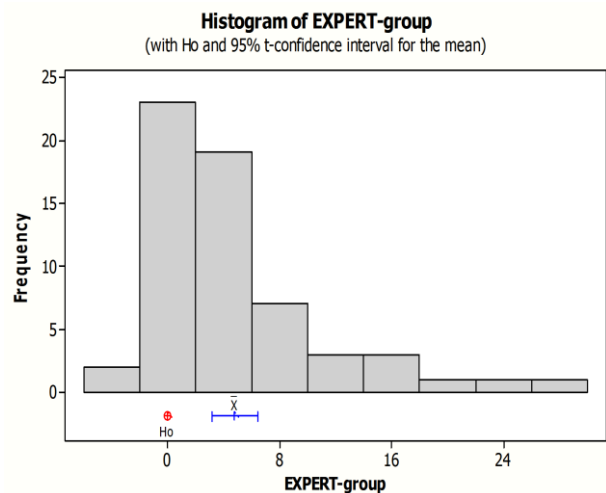


Figure 3. Frequency distribution of project estimate accuracy (Group, Expert)

Table 2. Results of anchoring effects (spearman’s rank correlation)

	Individual	Average in Ind	Group
Non experts			
Anchoring and adjustment	n = 16	n = 7	n = 7
Ho: rho=0	rho=0.562 (pvalue=0.024)	rho=0.721 (pvalue=0.068)	Rho=0.5333 (pvalue=0.218)
Experts			
Anchoring and adjustment	n = 17	n = 7	n = 10
Ho: rho=0	rho=0.416 (pvalue=0.109)	rho=0.559 (pvalue=0.192)	Rho=0.083 (pvalue=0.821)

Detailed results of the statistical analysis for individual/group against expert/non-expert scenarios are presented in Table 2. The first column of Table 2 shows the correlation analysis (non-parametric) between the estimates of ‘anchoring stimuli’ and those of the mini project duration for individual estimator. The third column presents the result for estimators working in a group setting. To observe whether simple average calculations of individual estimates for each group would provide a good performance against anchoring effect biases, results of the simple averaging computation is also presented in the second column of Table 2.

As can be seen in Table 2, significant anchoring effects ($p=0.024$) are observable for individual estimators of non-experts. A simple average computation of the individual estimators within respective groups could improve the performance of the estimates (i.e. anchoring effect is no longer significant; $p = 0.068 > 0.05$). Group estimators of non-experts indicates even better performance ($p=0.218$).

Table 2 also suggests that experts (either individual or group) do not indicate anchoring effects in their estimations (i.e. $p > 0.05$). From the p-values, it can be seen that group estimators provide the most superior performance of reeducing anchoring effects which is followed by simple averaging computation procedure and individual estimators.

To sum up, the empirical study in the anchoring effects shows that ‘level of expertise’ and ‘group/individual’ are two important determinants affecting the occurrence (or lack) of the anchoring biases. Experts (either individual or group) seem to have a better capability to eliminate anchoring effects in project duration estimations. Non-experts, on the other hand, would be able to eliminate such biases when the estimation process is carried out in a group environment.

Accuracy Bias

To observe the possible existence of accuracy biases, project estimates (individual or group) are compared to the respective, actual project durations. Errors in

accuracy are computed by subtracting project estimates with the actual project durations. Accuracy bias is observable statistically when the mean value is significantly different from zero. Figure 3 depicts the histogram of the errors in accuracy for the group expert estimators which indicate a systematic accuracy bias (mean > 0).

Table 3 shows detailed, experiment results for both experts and non-experts estimators in individual and group settings. As in the previous anchoring analysis, to observe whether simple average calculations of individual estimates for each group would provide a good performance against the biases, results of the simple averaging computation is also presented in the second column of Table 3.

Table 3 provides evidence that in terms of estimation accuracy, experts perform better than non-experts. It is reflected by smaller median values of errors for experts (ranging from 2.83 to 3.66) than those of non-experts (4.47 to 4.96). From another perspective, it is also observable that group estimators perform better than individual estimators in terms of accuracy. Results of the simple averaging protocol do not provide a consistently better accuracy than those of individual estimators. Hence it can be summarized that group, expert estimators perform best in terms of accuracy while individual, non-expert estimators perform worst. Other settings fall in between the two extreme performers.

From the perspective of bias, accuracy bias is pervasively observable for nearly all estimation settings. Both experts and non-experts are prone towards accuracy bias as indicated by results of non-parametric Wilcoxon tests of the accuracy errors. With Ho: two distributions are similar, most scenarios yield p-values < 0.05. Exceptions are for the non-expert estimators working in a group assignment. For this particular group, accuracy bias is not statistically observable for $p=0.089$. Nevertheless, the p-value is not remote from 0.05. A similar pattern is found for the estimates from the simple averaging protocol.

Table 3. Results of accuracy bias

	Individual	Average in Ind	Group
Non experts			
Accuracy	n = 40	n = 20	n = 20
biases			
Wilcoxon test	median=4.96; IQR= 32.71 (pvalue=0.011)	median=7.47; IQR= 32.44 (pvalue=0.055)	median=4.47; IQR= 25.12 (pvalue=0.089)
Experts			
Accuracy	n = 3x40	n = 3x20	n = 3x20
biases			
Wilcoxon test	median=3.66; IQR=7.56 (pvalue=0.000)	median=3.51 IQR=7.05 (pvalue=0.000)	median=2.83; IQR=6.04 (pvalue=0.000)

Accuracy = estimate - actual

A detailed observation shows that the median values of estimates for all settings are greater than zero (median > 0) suggesting that project estimates are systematically greater than the actual project durations. Hence estimators systematically take a conservative approach in making project estimations.

Accordingly it could be said that, in terms of accuracy, experts perform better than non-experts. In addition, group estimators provide better accuracy than individual estimators. On the contrary, from the accuracy bias perspective, neither experts or non-experts could effectively eliminate the systematic errors. Performance could not also be improved by working in a group context.

Managerial Insights

From the study it could be observed that learning curves are applicable for the experiments. After conducting a series of similar projects, over time, project practitioners are able to execute the projects in shorter durations until hitting the asymptotic line where further improvement is virtually unobservable. The same learning curve is also observable for the cognitive tasks of conducting project estimation. In general, over time, the errors of estimates decrease to an asymptotic line. Interestingly, both individual and group estimators experience the same phenomenon of learning curve while majority of the estimators do not realize it. The learning curve also indicates the transformation of the subjects from non-experts to experts. This is in line with the common wisdom that expertise could be learned from prior experiences whether or not the person is aware of the concepts. This learning process is especially applicable due to simplicity of the project assignment (Sterman [26]).

The empirical study provides evidence that the existence of anchoring effects is affected by the two currently observed dimensions, namely: level of expertise of the estimators and the group/non-group setting. Results suggest that experts perform better than non-experts in reducing the effects of anchoring. Group estimation is also found as an effective way to reduce such effects. In addition, a simple protocol of averaging the estimates of individual estimators works quite well in this case. The finding is consistent to the past studies such as those reported by Kahneman [16] and Hastie and Dawes [15].

In a practical context, anchoring stimuli could be easily observable in various settings prior to any project estimate exercise. Accordingly, practitioners need to be aware to the existence of such a bias. Moreover, project-based organizations need to develop a protocol to remedy or to eliminate the anchoring effects. From this study, the suggestion for

the remedy is to ask for the experts to carry out the estimates preferably working in a group setting. If such a remedy is not feasible, a simple protocol of averaging the estimates of two estimators would provide a better result in reducing the bias than those of a single estimator.

The study also suggests that accuracy of estimates is affected by estimators' level of expertise and the group/non-group setting. Experts in a group setting perform best in terms of estimate accuracy. Accordingly, to improve accuracy in project estimates, project-based organizations need to develop a procedure to encourage the utilization of expert estimation under a group environment.

Accuracy bias is found pervasive for various estimation settings. Both experts and non-experts are prone towards accuracy bias. This is consistent to various past reports such as those in Bazerman [2], Mak and Raftery [19], and Tversky and Kahneman [27]. Results from this empirical study also imply that estimators tend to take a more conservative approach. As seen in Table 3, statistical analyses indicate that accuracy (i.e. the difference between estimate and actual project duration) is significantly different from zero with medians greater than zero. In other words, majority of estimators provide project time estimations which are longer than the actual durations. Accordingly, estimators in this study behave conservatively. This practice found in the current study is on the contrary to those of various past studies (e.g. Flyvbjerg [8]; Flyvbjerg [9]; Flyvbjerg *et al* [11]; Lovallo and Kahneman [18]) suggesting that aggressive risk taking estimates are the more common observable phenomenon in accuracy bias. The different pattern of results may be attributable to various factors such as the risk culture in which the estimators operate. However, follow-up studies are required to provide scientific explanations for the phenomena.

To remedy the problem of accuracy bias, an enhanced, joint cognitive system needs to be developed by project-based organizations. The system consists of the cognitive task (i.e. project estimation), human estimators, and procedures of de-biasing. Similar joint cognitive systems had been developed in a slightly different context by scholars to overcome over-optimism biases (e.g. Flyvbjerg [10]). The proposed procedure for the Indonesian context would be taking a similar route of utilizing external references for de-biasing; but it should work on the opposite direction (i.e. to overcome over-conservatism bias).

Conclusion

An experimental study had been carried out in a project context to investigate the performance of

experts and non-experts in conducting a cognitive task of project duration estimation. The result shows that since human decision makers have limited cognitive capability, results of the project estimates are prone towards systematical errors (biases). In various scenarios of experiments, however, experts perform better than the non-experts. For the anchoring effect, biases could be reduced by having experts working in group estimation. The experience gained by experts in conducting past, similar cognitive tasks seems to contribute to the decreased impact of the bias in estimations. The same condition is observable for group estimations. Group discussions prior to the project estimate become an effective method of reducing anchoring effects. Experience and group discussions seem to increase the cognitive capability of the overall cognitive system leading to a better performance.

For the accuracy biases, however, experiences and group discussions do not contribute significantly in reducing the systematical errors. Hence to deal with accuracy bias, a specific intervention may be required. This includes the development of certain protocol, method, or tool to enhance the overall capability of the joint cognitive system. In cognitive ergonomics/psychology the procedure is termed debiasing. A series of follow-up studies is then required to provide a complete picture of the phenomenon and to develop joint cognitive system-level interventions to minimize errors in conducting project estimation tasks.

References

1. Akintoye, A. S., and MacLeod, M. J., Risk Analysis and Management in Construction, *International Journal of Project Management*, 15, 1997, pp. 31-38.
2. Bazerman, M. H., *Judgment in Managerial Decision Making*, New York: John Wiley & Sons, 1998.
3. Boy, G., Shadbolt, N., O'Hara, K., and Schreiber, G., The Group Elicitation Method: An introduction, in *Advances in Knowledge Acquisition*. 1076: Springer Berlin/Heidelberg, 1996, pp. 290-305.
4. Cleaves, D. A., Cognitive Biases and Corrective Techniques: Proposals for Improving Elicitation Procedures for Knowledge-based Systems, *International Journal of Man-Machine Studies*, 27, 1987, pp. 155-166.
5. Cooke, R. M., and Goossens, L. H. J., Expert Judgement Elicitation for Risk Assessments of Critical Infrastructures, *Journal of Risk Research*, 7, 2012/10/15 2004, pp. 643-656.
6. Dvir, D., Transferring Projects to Their Final Users: The effect of Planning and Preparations for Commissioning on Project Success, *International Journal of Project Management*, 23, 2005, pp. 257-265.
7. Dvir, D., Raz, T., and Shenhar, A. J., An Empirical Analysis of the Relationship between Project Planning and Project Success, *International Journal of Project Management*, 21, 2003, pp. 89-95.
8. Flyvbjerg, B., How Common and How Large are Cost Overruns in Transport Infrastructure Projects?, *Transport Reviews*, 23, 2003, pp. 71-88.
9. Flyvbjerg, B., From Nobel Prize to Project Management: Getting Risks Right, *Project Management Journal*, 37, 2006, pp. 5-15.
10. Flyvbjerg, B., Glenting, C., and Rønneest, A. K., Procedures for Dealing with Optimism Bias in Transport Planning: Guidance Document, The British Department for Transport, London 2004.
11. Flyvbjerg, B., Holm, M. K. S., and Buhl, S. L., How (In)accurate Are Demand Forecasts in Public Works Projects?: The Case of Transportation, *Journal of the American Planning Association*, 71, 2005, pp. 131-146.
12. Friedman, D., and Cassar, A., *First Principles: Induced Value Theory*, London: Routledge, 2004.
13. Gigerenzer, G., and Goldstein, D. G., Reasoning the Fast and Frugal Way: Models of Bounded Rationality, *Psychological Review*, 103, 1996, pp. 650-669.
14. Hartono, B., and Yap, C. M., Understanding Risky Bidding: a Prospect-Contingent Perspective, *Journal of Construction Management and Economics*, 29, 2011, pp. 579-593.
15. Hastie, R., and Dawes, R. M., *Rational Choice in an Uncertain World: The Psychology of Judgment and Decision Making*. Thousand Oaks, CA: Sage Publications, Inc, 2001.
16. Kahneman, D., *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux, 2011.
17. Levin, I. P., Louviere, J. J., Schepanski, A. A., and Norman, K. L., External Validity Tests of Laboratory Studies of Information Integration, *Organizational Behavior and Human Performance*, 31, 1983, pp. 173-193.
18. Lovallo, D., and Kahneman, D., Delusions of Success. How Optimism Undermines Executives' Decisions, *Harvard Business Review*, 81, 2003, pp. 56-63.
19. Mak, S., and Raftery, J., Risk Attitude and Systematic Bias in Estimating and Forecasting, *Construction Management and Economics*, 10, 1992, pp. 303-320.
20. McGraw, K. L., and Seale, M. R., Knowledge Elicitation with Multiple Experts: Considerations and Techniques, *Artificial Intelligence Review*, 2, 1988, pp. 31-44.
21. O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., and Rako, T., *Uncertain Judge-*

- ments: Eliciting Experts' Probabilities*: John Wiley & Sons Inc, 2006.
22. PMI, *A Guide to the Project Management Body of Knowledge*. Pennsylvania, USA: Project Management Institute, 2000.
 23. Shen, L. Y., Project Risk Management in Hong Kong, *International Journal of Project Management*, 15, 1997, pp. 101-105.
 24. Simon, H. A., *Models of Man*, New York: Wiley, 1957.
 25. Simu, K., Risk Management on Small Projects, in *Proceedings of 4th Nordic Conference on Construction Economics and Organisation* Luleå, Sweden: Luleå University of Technology, 2007.
 26. Sterman, J. D., *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Boston McGraw-Hill/Irwin, 2000.
 27. Tversky, A., and Kahneman, D., Judgment under Uncertainty: Heuristics and Biases, *Science*, 185, 1974, pp. 1124-1131.
 28. Vose, D., *Quantitative Risk Analysis: a Guide to Monte Carlo Simulation Modelling*, New York: Wiley, 1996.
 29. Wood, G. D., and Ellis, R. C. T., Risk Management Practices of Leading UK Cost Consultants, *Engineering, Construction and Architectural Management*, 10, 2003.
 30. Yelle, L. E., The Learning Curve: Historical Review and Comprehensive Survey, *Decision Sciences*, 10, 1979, pp. 302-328.