

Testing a PPP Performance Evaluation Framework

J. Liu^{*1}, P.E.D Love², J. Smith³ and C.P. Sing⁴

^{1,2} Curtin University, Perth, Australia

^{3,5} Bond University, Gold Coast, Australia

⁴ City University of Hong Kong, Hong Kong

* Corresponding Author

ABSTRACT Public-Private Partnerships (PPPs) have become an integral part of infrastructure procurement strategy in many governments across the world. However, the use of PPPs has been being plagued with controversy, especially in the UK and Australia, as some procured PPP projects have experienced significant time and cost overruns and poor or less than optimal operational performance. A perspective that has been raised is that the unsatisfactory delivery of PPPs over the past decade has resulted from the absence of an ineffective and incomplete performance evaluation. Typically evaluation has focused been *ex-post* measuring construction deliverables and thus has ignored the projects' inception phases. With this in mind, an innovative evaluation framework is presented and empirically tested using a case study and Confirmatory Factor Analysis to evaluate the outputs of the early formative stages of PPPs (i.e., initiation, planning and procurement). The developed evaluation framework can provide governments with a reliable tool for measuring and managing their PPP projects.

1 INTRODUCTION

A plethora of factors contribute to the successful delivery of a PPP project. Problematic issues related to time and cost overruns in PPPs over the past decade have resulted from an incomplete and ineffective performance evaluation throughout their life-cycle (Liu *et al.*, 2015a). In fact, most PPPs have not undergone a comprehensive evaluation in terms of what had been delivered (Regan *et al.*, 2011).

Performance evaluation is essential for business success, particularly at the corporate and project levels (Love and Holt, 2000). Liu *et al.* (2015b) have suggested that effective and efficient project evaluation is a critical success factor (CSFs) for PPPs. Essentially, evaluating project performance is a core activity of the contract management of PPPs in most developed countries (European Investment Bank – EIB, 2011). Despite the importance evaluation, limited research was undertaken to explore this critical

issue (Liu *et al.*, 2014). Against this contextual backdrop, this paper empirically develops and tests an innovative life cycle PPP performance measurement framework (PMF).

2 RESEARCH OF PPPs

PPPs have been perceived to be a time and cost efficient procurement approach for infrastructure projects. As a result, there has been an inordinate amount of PPP research. Liu *et al.* (2015) have identified that there are six key research areas in the normative literature of PPPs, including governments' roles and responsibilities, selection of concessionaire, risk identification and allocation, cost and time efficiencies of different contracts and project finance.

Although PPPs have attracted intensive attention of researchers, there have been limited studies that have empirically examined the performance of PPPs (Love *et al.*, 2015). Therefore, this study bridges this

knowledge gap and contributes to the normative literature with an introduction of a Performance Management Framework (PMF) that provides governments with a reliable tool for measuring the outputs of inception phases of their projects.

3 METHODOLOGY

A detailed case study relying on an array of documentary sources and semi-structured interviews was conducted to develop hypotheses for conceptualising a PMF. The case used in this paper was chosen by communicating with a project director, who was experienced in delivering infrastructure PPPs, during the initial stage of the study. Thus, the findings of the recommended case study can be assumed to be reliable and significant.

Following the case study, a conceptual PMF was developed, constituting five performance measurement perspectives and a series of relevant key performance indicators (KPIs). To empirically test the feasibility of the developed framework, a questionnaire was designed to solicit PPP practitioners' views and insights about the PMF. The hypotheses to be tested using the survey are indicated as follows:

$F^1 - H_0$: The measurement perspectives are not significant for measuring PPPs.

$F^1 - H_1$: All measurement perspectives are significant for measuring PPPs.

$F^2 - H_0$: The conceptually-derived KPIs are not significant for measuring PPPs.

$F^2 - H_1$: The majority of the conceptually-derived KPIs are significant for measuring PPPs.

A *Likert* scale ranging from 1 (strongly disagree) to 5 (strongly agree) and purposive sampling were applied to the survey. Purposive sampling is useful when the researchers require the expertise of individuals who are specialised in particular fields to deal with a topic that integrates with a high level of uncertainties (Foreman, 1991). Respondents who participated in this study had to be knowledgeable about multiple aspects of PPPs, many of which are considered to be dynamic and uncertain projects (Yuan *et al.*, 2009). Thus, purposive sampling was ideal for this type of study and the target population were senior management personnel who had been involved in the procurement of social PPPs (Hodge, 2004).

The developed questionnaire was distributed to the selected respondents via *SurveyMonkey*. The collected data was analysed by applying Confirmatory Factor Analysis (CFA), which is a multivariate process to statistically examine how well the variables being measured represent their constructs. CFA relies on pre-constructed theory and it is used for confirming the theoretical relationships, rather than exploring the linkages between the items (Schreiber *et al.*, 2006).

4 CASE STUDY OF A SOCIAL PPP PROJECT

4.1 Description of the Case Project

The selected PPP project was initiated with an aim of delivering a new facility to replace an existing regional prison, which was built in the 1980s with 100 beds. The business case of Project-P for a redevelopment of a new prison was approved in 2009 by the State Treasury, and comprised: (1) 200-bed male medium security prison with the capacity to accommodate 20 maximum security prisoners; (2) 60-bed male minimum security sector; (3) 40-bed male open minimum security sector; and (4) 50-bed female maximum, medium and minimum sector including a 6-bed unit for women with children.

The procurement arrangement that the selected PPP undertook was a Design, Build, Finance and Maintenance (DBFM) model. The State Government expected value for money (VfM) from the private sector, which involved asset durability, efficiency and productivity, innovation in design, lower life-cycle costs and quality outputs.

4.2 Delivery Process of Project-P

After an evaluation of the Expression of Interests (EOI) submissions, the PPP project proceeded to the Request for Proposal (RFP), whereby the short-listed respondents of were requested to submit fully-costed and complete proposals in 2012. The government then evaluated the received proposals against a range of criteria, which included organisational structure, stakeholder relationship management, design solution and management, project management, delivery of facility maintenance (FM) services.

Table 1 reports the delivery process and timeframe of Project-P.

Table 1: Project-P delivery process and timeframe

Phases of Delivery Process	Timeframe
<i>Initiation and Planning</i>	
Business case study	2009 - 2011
Definition and procurement option	
Invitation for EOI	
Evaluation of the EOIs	
<i>Procurement</i>	
Release of RFP	Jan-Dec, 2011
Tendering & final negotiation	
Contract and financial close	
<i>Design and Construction</i>	
Commencement of construction	2013 – 2015
Stage 1 & 2 works completion	
<i>Facility Maintenance</i>	
Operation and FM	Since mid. 2015
<i>Handover</i>	

Source: 'Project Summary' of Project-P, p.12

4.3 Practice in Performance Evaluation

The aim of this paper is to develop a PMF for evaluating the life cycle of PPP, which were handled by the government. Hence, interviews with the Procurement Director, who was responsible for overseeing such aforementioned phases of Project-P, were conducted. The questions used for the interview focused on how Project-P was evaluated throughout its early stages. These included:

- How did you measure the deliverables of the inception stages of the project?
- What do you consider to be limitations in the current approach you used to measure the project's early-stage outputs?

It was identified from the interviews that the approach adopted to evaluate Project-P during its formative stages, for example, focused on the reviews of the business case and tender decision. As the Director stated:

"In the pre-contracting phase, we had reviews for the business case development and the decision on the tendering to examine whether they are appropriate. Before these reviews, we also conducted a VfM assessment for the project."

This point of view was supported by the 'Project Summary', which stated that the evaluation approach of the inception phases of Project-P incorporated

three parts: (1) a VfM assessment; (2) a review of the business case; and (3) a review of the appropriateness of the tendering decision (Figure 1).

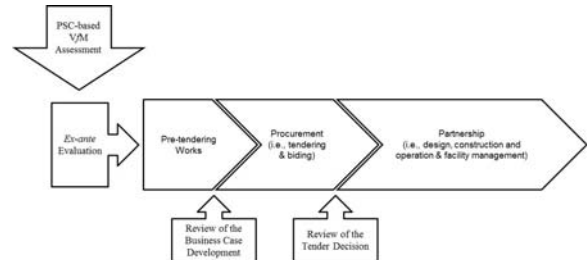


Figure 1: Performance evaluation in pre-contract of Project-P

4.4 Limitations of Current Evaluation

The Director identified that there were some limitations with the existing evaluation approach that was being used and stated:

"Performance evaluation is used for improvement. In the case of this project, improving is first about the efficiency of the Procurement. It sometimes took us a little bit long to pass an approval process. Competition is an aspect we expected in the Procurement phase. But this is missed in the project's performance evaluation and it might be necessary for us to have a mechanism to examine whether the competitiveness of bidding can achieve the level we expected."

The Director also considered that the management of evaluation information in Project-P needed to be improved and an effective and efficient internal process would be required to capture the lessons learned from the evaluation results. Additionally, as outlined in the 'Project Summary', the VfM assessment of Project-P still relied on the *Public Sector Comparator* (PSC). The PSC has been criticised by both academia and practitioners owing to its pure focus on financial benefits. The Director acknowledged that the VfM assessment of Project-P should be ameliorated, and stated:

"I think PSC gave us good results but we will have to continue refining it, of course, because it is true that PSC is a quantitative estimation and does not embrace important qualitative issues of PPPs."

To solve this problematic issue, the interview respondent tended to suggest that a process-based measurement with the measures capable of capturing the key stakeholders' expectations and strategic goals

would be a promising way for future amelioration. The Director, for example, suggested other aspects, such as ‘V/M for non-financial benefits’, ‘project planning’, ‘competitiveness of bidding’, ‘efficiency of approval process during the procurement phase’ and ‘capabilities of potential *Special Purpose Vehicles* (SPV)’, should be considered during the early formative stages of the project.

5 CONCEPTUAL PMF FOR PPPs

The case study provided an understanding of the current practice used to evaluate the performance of a PPP used to deliver a social infrastructure project. It was identified that the inception phases of a PPP evaluation (e.g., Project-P) possess a range of ‘gaps’, involving an incomplete V/M assessment, incomplete measurement for the essential deliverables and ineffective and inefficient internal process used to supporting learning.

As suggested from the interviews, a process-based performance evaluation with learning mechanisms would be useful for government to improving that the performance their infrastructure assets delivered by PPPs. This view is supported by the research conducted by Liu *et al.* (2015a), who have suggested that process-based measurement is robust to effectively evaluate the performance of PPPs over their life cycles as it can capture all essential works required to deliver them. A process-based PMF is conceptualised with an aim of measuring the inception stages of PPPs (Figure 1).

Key Performance Indicators (KPIs) form the heart of performance measurement systems (PMSs) (Neely *et al.*, 2005). Hence, the *Performance Prism* developed by Neely *et al.* (2001) is applied to derive performance measures for the proposed PMF. Neely *et al.* (2002) has stated that the Performance Prism is a stakeholder-focused framework that can shed light on organisational measurement within a multiple stakeholder environment. In addition, Liu *et al.* (2015c) have demonstrated that Performance Prism is suitable for deriving measures for the PMSs devised for PPPs as it can accommodate their inherent complexities and uncertainties that result from a highly sophisticated stakeholder network.

The Performance Prism encompasses *five* measurement perspectives: (1) Stakeholder Satisfaction;

(2) Strategies (i.e., project strategic goal); (3) Processes (i.e., project delivery process); (4) Capabilities (i.e., organisations’ capabilities in delivering the project); and (5) Stakeholder Contribution (Neely *et al.*, 2001). Noteworthy, a sequence of phase-based KPIs that address the stakeholders’ needs and expectations can be developed for the proposed PMF (Figure 2).

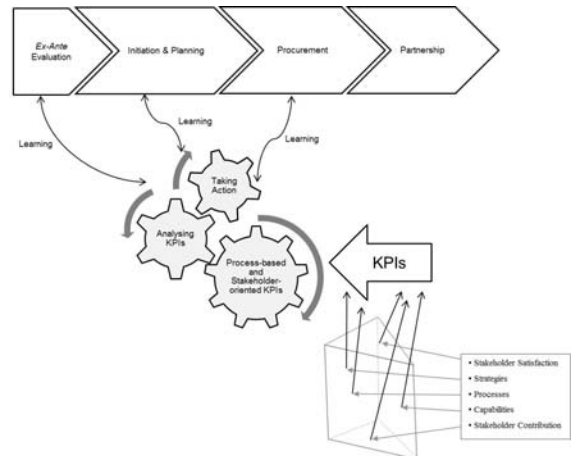


Figure 2: Process-based PMF of PPPs

6 FEASIBILITY TESTING

A questionnaire was designed to solicit senior practitioners’ perspectives about the hypotheses derived from the case study. A pilot survey was conducted to test the reliability of the developed research instrument. It was distributed to 28 experienced PPP practitioners. A total of 25 responses were received, equating to a responsive rate of 89%.

Throughout the data collection period, a total of 368 questionnaires were distributed to practitioners who were knowledgeable of about issues associated with both the public and private sectors involvement with the delivery of PPPs. A total of 141 responses were received; 6 responses had to be discarded because of incompleteness. Data completeness is a prerequisite of CFA. Hence, a total of 135 valid datasets were entered into SPSS for related analysis.

Cronbach’s α and the corrected item-total statistics were applied prior to conducting the CFA. All values generated from these two tests exceed 0.70 and 0.30, which indicate a high degree of internal consistency and reliability of the observed items (Nunnally, 1978; Nunnally and Bernstein, 1994).

After conducting the reliability tests, the CFA was undertaken. Essentially, CFA is a theory-driven factor analysis technique and the theoretical linkages drive its formulation between the observed and unobserved variables (Schreiber *et al.*, 2006). Within the configuration of the developed PMF (Figure 2), the five measurement perspectives and their relevant KPIs are viewed as the observed variables, while the outputs/deliverables of each PPP phase are considered to be the unobserved variables.

Initially, the CFA-hypothesised model was constructed to estimate a covariance matrix of the survey population in comparison with an observed covariance matrix. Accordingly, the formulated model was applied to examine whether the observed items (five performance measurement perspectives and their KPIs) were significant or could significantly contribute to the performance of a PPP. Items with comparatively low ‘factor loadings’ were eliminated to allow modifications to the initially hypothesised model to develop an optimal model.

The hypothesised model captures the features of the developed PMF, whereby the process-based KPIs are proposed according to the five measurement perspectives that were assumed to be causally significant to the performance of a PPP project.

The optimal models developed indicated that the coefficients of the five measurement perspectives are 0.78, 0.82, 0.77, 0.75 and 0.76, respectively, all of which are significant under 5% significance level. These suggest that such proposed perspectives are significant for measuring the performance of PPPs during their formative stages.

Apart from the findings relating to the five measurement perspectives, a sequence of critical implications can be drawn from the generated empirical evidence regarding the KPIs. Under the pre-contract phases (i.e., Phase 1: Initiation and Planning; Phase 2: Procurement), the coefficients of most of the derived KPIs are larger than 0.50 and are significant (under 5% significance level), which means that such observed KPIs are significant for the performance measurement of PPPs (Yuan *et al.*, 2012).

Most respondents acknowledged that the PPP market in Australia to be sophisticated and mature. Essentially, PPPs have become an integral part of both the federal and state governments’ procurement strategies in Australia (Regan *et al.*, 2011). Australian

PPP projects are procured by following strict well-developed guideline and processes (Infrastructure Australia, 2008). As a result, state governments and an array of private-sector entities within Australia have become experienced in procuring and delivering infrastructure projects via PPPs. The public and private sectors are familiar with solving process issues such as financing options and designing and determining an appropriate concession period, and/or how to organise and govern well the tendering and efficiently achieve financial close. This may explain why the KPIs relating to PPP project’s finance option concession period and financial close, and the government’s governance ability for the procurement phase were considered to be insignificant by the respondents of this survey.

After eliminating the four insignificant indicators an optimal model was constructed. The estimates of the optimal model are larger than 0.50 and significantly correlated to the performance of PPPs at a 5% significance level.

To examine the model, three Goodness-of-Fit Indexes (GFIs) were used: (1) *Chi-squared* statistic; (2) comparative fit index (CFI); and (3) root mean square error of approximation (RMSEA). Such GFIs are widely used for indicating how well the constructed structural model fits a set of observations (Sanders *et al.*, 2006). Table 2 identifies the benchmark values of the aforementioned GFIs. In other words, the constructed structural model is deemed as a ‘fitted’ model if its GFIs are within the intervals of the benchmark values.

Table 2: Benchmark values of the GFIs

Goodness-of-fit Indexes	Benchmark Values
<i>Chi-squared</i>	$1 \leq \chi^2/Df \leq 5$
CFI	≥ 0.90
RMSEA	$0.05 \leq \text{good model fit} \leq 0.1$

The empirical evidence relating to the three GFIs of the constructed optimal model are summarised as follows: 2.32 (*Chi-squared* statistic), 0.92 (CFI) and 0.076 (RMSEA). These indicate a good model fit. On the basis of the results derived above, it can be found that all of the five proposed measurement perspectives are significant for the performance measurement of PPPs, while 28 indicators out of the 32 proposed KPIs ‘passed’ the quantitative tests that relied on CFA. Therefore, the null hypotheses proposed above are rejected. The empirical findings confirmed

the main proposition, that the stakeholder-oriented measurement perspectives and their relevant KPIs are significant for future PPP performance evaluation.

7 CONCLUSION

This paper proposed and empirically tested a life cycle PMF for evaluating the outputs of PPPs with particular emphasis being placed on the formative phases, such as initiation, planning and procurement. A detailed case study was used to develop the conceptual framework. Then CFA was performed to testing its validity using a questionnaire survey. Four insignificant KPIs were eliminated, and it was identified that the main components of the developed PMF (five measurement perspectives and 28 process-based KPIs) were feasible and applicable for evaluating social infrastructure PPPs.

The research has contributes to the literature by filling the knowledge gap of PPP performance evaluation. As the PMF was developed from a 'real-world' project and validated by experienced practitioners, this paper provides governments that will embark on PPPs with a robust tool and conceptual foundation to design effective and efficient PMSs for effectively and efficiently evaluate their future projects.

ACKNOWLEDGEMENT

The authors would like to thank those participants who participated in this study. The authors would also like to acknowledge the financial support provided by the Australian Research Council (LP120100347).

REFERENCES

- EIB 2011. The guide to guidance: how to prepare, procure and deliver PPP projects, Luxembourg
- Foreman, E.K. 1991. *Survey sampling principles*, Marcel Dekker, NY, USA
- Hodge, G.A. 2004. The risky business of public-private partnerships. *Australian Journal of Public Administration*, **63**, 37-49.
- Liu, J., Love, P.E.D., Smith, J., Regan, M. & Sutrisna, M. 2014. Public-private partnerships: A review of theory and practice of performance measurement. *International Journal of Productivity and Performance Management*, **63**, 499-512.
- Liu, J., Love, P.E.D., Davis, P.R., Smith, J. & Regan, M. 2015a. Conceptual framework for the performance measurement of public-private partnerships, *Journal of Infrastructure Systems*, **21**, 04014023.
- Liu, J., Love, P.E.D., Smith, J., Regan, M. & Davis, P.R. 2015b. Life cycle critical success factors for public-private partnership infrastructure projects, *Journal of Management in Engineering*, **31**, 04014073
- Liu, J., Love, P.E.D., Smith, J., Regan, M. & Palaneeswaran, E. 2015c. Review of performance measurement: implications for public-private partnerships, *Built Environment Project and Asset Management*, **5**, 35-51.
- Neely, A., Adams, C. & Crowe, P. 2001. The performance prism in practice. *Measuring Business Excellence*, **5**, 6-12.
- Neely, A., Gregory, M., & Platts, K. 2005. Performance measurement system design: a literature review and research agenda. *International Journal of Operations & Production Management*, **25**, 1228-63.
- Nunnally, J.C. 1978. *Psychometric Theory*. 2nd ed., McGraw-Hill, NY, USA
- Nunnally, J.C. and Bernstein, I.H. 1994. *Psychometric Theory*, 3rd ed., McGraw-Hill, NY, USA
- Regan, M., Smith, J. & Love, P.E.D. 2011. Infrastructure procurement: Learning from public-private experience 'down under', *Environment Planning C: Government and Policy*, **29**, 363-378.
- Schreiber, J.B., Nora, A., Stage, F.K., Barlow, E.A., and King, J. 2006. Reporting structural equation modeling and confirmatory factor analysis results: A review, *The Journal of Educational Research*, **99**, 323-338.
- Yuan, J., Zeng, A.Y., Skibniewski, M.J., & Li, Q. 2009. Selection of performance objectives and key performance indicators in public-private partnership projects to achieve value for money. *Construction Management and Economics*, **27**, 253-270.
- Yuan, J., Wang, C., Skibniewski, M.J., and Li, Q. 2012. Developing key performance indicators for Public-Private Partnership projects: Questionnaire survey and analysis, *Journal of Management in Engineering*, **28**, 252-264.