

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Oliveros Romero, Jose Alejandro

Title:

Ex-post impact evaluation method for public-private partnership projects from the perspective of multiple stakeholders

Date:

2019

Persistent Link:

<https://hdl.handle.net/11343/239210>

Terms and Conditions:

Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.

Ex-post impact evaluation method for public-private partnership projects from the perspective of multiple stakeholders

Jose Alejandro Oliveros Romero

0000-0002-9948-0547

Doctor of Philosophy

December 2019

Faculty of Architecture, Building and Planning

Melbourne School of Design

Submitted in total fulfilment for the degree of

Doctor of Philosophy

ABSTRACT

Public-Private Partnership (PPP) model of procurement is a way to deliver public infrastructure using private funding and managing risk for public purposes. Currently, PPPs have been criticised in the media, social media and in the academic literature, notably regarding their limited contribution to real public welfare. The arguments are related to (1) the amount of profit that is conferred to the private entity, (2) the long-term implications for the public sector in term of service payments, (3) the quality of the service which can be driven by cost reduction goals; among others.

From a taxpayer's point of view, the impact of PPP projects, like any kind of public investment, should be assessed and analysed, especially if a private company is making profit with it. One relevant instrument to protect the taxpayers' interest is an ex-post impact evaluation. The problem is that a PPP ex-post impact evaluation is difficult to perform because usually: (1) projects are large and complex, (2) there are different perceptions of the delivery and the impacts, (3) the time horizon for the evaluation in some cases is more than 20 years, and (4) there are political, economic, and ideological incentives that can create biases towards their evaluation.

The problem to be addressed by this research is the lack of a method to evaluate, on an ex-post basis, the impact of an infrastructure PPP project from the perspective of multiple stakeholders. The arguments that support the problem are the existing criticism of PPPs regarding their performance for achieving public welfare, the nonexistence of a PPP ex-post evaluation method, and the difficulty associated with PPP evaluation due to their multiple perspectives from the different stakeholders. The research question is:

How can ex-post impact evaluation of Public-Private Partnership projects be conducted employing the perspective of its multiple stakeholders?

For developing the PPP ex-post impact evaluation method, a Design Science approach is adopted. Design Science differs from the traditional behavioural science approach because it aims to change nature by solving a problem through the design of an 'artefact'. The evaluation method is considered an artefact for the purpose of this research, grounded on the Project Success concept from the project management discipline.

The outcome of the research is an ex-post impact evaluation method presented as a procedure with five steps: (1) evaluation setup, (2) stakeholder workshop preparation, (3) stakeholder workshop, (4) data analysis, and (5) evaluation conclusion.

This research has a theoretical contribution by refining the concept of project as a project-system, and by proposing three high-level principles to inform a successful project. The research also has a practical contribution by delivering a documented evaluation method to assess existing PPP projects by systematically addressing their multiple perspectives towards them.

DECLARATION

This is to certify that:

- The thesis comprises only my original work towards the PhD except where indicated
- Due acknowledgement has been made in the text to all other material used,
- The thesis is fewer than 100.000 words in length, excluding of tables, biographies, and appendices

José Alejandro Oliveros Romero

December 2019

ACKNOWLEDGEMENT

I wish to acknowledge:

- Brookfield Multiplex for the scholarship that supported the research project
- My Supervisor, Dr Ajibade Aibinu, for his guidance, support, critiques and endless dedication to my work
- My committee members Roy Barton, Andre Stephan and James Mitchell for their commitment to the project, advice, support and help
- My wife, Constanza Gonzalez, for her patience, help and advise in almost every aspect of this endeavour
- My son, Emilio, for making this journey even more interesting and challenging
- And finally, to my colleagues and friends in the University, with whom the key ideas of the thesis were discussed, refined, and improved.

LIST OF PUBLICATIONS

The following articles were published during the course of the PhD study:

Refereed journal articles

Oliveros R., J.; Aibinu, A. (2019). Ex-post Value for Money assessment for PPP projects: A qualitative approach. *Built Environment Project and Asset Management Journal*. Manuscript submitted for publication.

Oliveros R., J.; Aibinu, A. (2019). Ex-post impact evaluation of PPP projects: An exploratory research. *Built Environment Project and Asset Management Journal*, 9(2), 315-330.

Barton, R.; Aibinu, A.; Oliveros R., J. (2019). Value for Money Concept in investment evaluation: Deconstructing its meaning for better decision-making. *Project Management Journal*, 50(2), 210-225.

Peer-reviewed academic conferences

Oliveros R., J. & Aibinu, A. (2017). The application of design science to AEC research. Paper presented at *Organization, Technology and Management in Construction (OTMC) Conference*, 27-30 September. Poreč, Croatia.

Oliveros R. J. & Aibinu, A. (2017); Towards an ex-post evaluation of PPP projects from the perspective of multiple stakeholders. Paper presented at *Advancement of Construction Management and Real Estate (CRIOCM) Conference*, 20-23 November, Melbourne, Australia.

Oliveros R., J. & Gonzalez-Mathiesen, M. (2017). Real estate social housing market employing a PPP approach: The Chilean Case. Paper presented at *State of Australian Cities National Conference 2017 (SOAC)*, 28-30 November. Adelaide, Australia.

Oliveros R., & Aibinu, A. (2017). Ex-post impact evaluation of PPP projects – The Project Success evaluation pyramid model. Paper presented at *International Conference on Innovative Production and Construction (IPC) 2017*, 30 November – 1 December. Perth, Australia.

Non-refereed articles

Oliveros R. J.; Aibinu, A. (2018). Ex-post impact evaluation of PPP projects using multiple stakeholder perspectives. Paper presented at *13th European Evaluation Society Biennial Conference*, 1-5 October. Thessaloniki, Greece.

Oliveros R., J.; Cerda, D. (2017). Propuesta avanzada para el Sistema de retribucion del programa Becas Chile. Paper presented at *5th Chilean Graduate Student Conference 2017*. University of New South Wales, Sydney, Australia.

Oliveros, J. (2016). After everything is built and used, did we do it right?, *Atrium Magazine*, Faculty of Architecture, Building and Planning, University of Melbourne. Melbourne, Australia.

TABLE OF CONTENT

Abstract	ii
Declaration	iii
Acknowledgement	iv
List of publications	v
Table of content	vi
List of figures	x
List of tables	xi
List of equations	xii
List of abbreviations	xiii
CHAPTER 1 Introduction	1
1.0. Introduction	1
1.1. Research background	1
1.2. Research problem	2
1.3. Research aim & questions	5
1.4. Research rationale	5
1.5. Research strategy	6
1.6. Significance of the research	8
1.7. Thesis structure	8
CHAPTER 2 The broad context of Public-Private Partnerships	10
2.0. Introduction	10
2.1. PPP characteristics	10
2.1.1. Public provision defects	12
2.1.2. PPP advantages	13
2.1.3. PPP's feature is a better risk allocation	14
2.1.4. PPP success factors	15
2.1.5. PPP problems and risks	15
2.2. PPP procurement issues	16
2.2.1. AHP in PPP projects is not recommended	16
2.2.2. Demand estimation is a key problem for PPP legitimisation	16
2.2.3. Least Present Value of Revenues (LPVR) mitigates the demand risk	16
2.2.4. Non-transport PPP projects (schools, hospitals, and prisons)	17
2.3. Public-Private Partnerships in Chile	17
2.3.1. The Chilean social housing system is a non-traditional PPP model	19
2.4. Public-Private Partnerships in Australia	19

2.5. Summary	21
CHAPTER 3 PPP Evaluation from multiple stakeholder perspectives: A literature review	23
3.0. Introduction	23
3.1. Initial definitions	23
3.2. Summary of the existing evaluation approaches	24
3.3. Research on PPP evaluation and monitoring	27
3.3.1. Value for Money concept	28
3.3.2. PPP ex-post evaluation guidelines	28
3.3.3. PPP performance	28
3.4. Management-based evaluation and monitoring approaches	29
3.4.1. Project Success concept	29
3.4.2. Stakeholder management	30
3.4.3. Performance Measurement Systems	31
3.4.4. Project audit	32
3.5. Program evaluation (public policy) approaches	32
3.6. International agencies' approaches	34
3.7. The gap in knowledge	34
3.8. Summary	35
CHAPTER 4 Research methodology	37
4.0. Introduction	37
4.1. Adopted research paradigm	37
4.2. Design Science approach review	37
4.2.1. Design Science structure	39
4.2.2. Design Science artefact evaluation	40
4.2.3. Design Science contribution to knowledge	40
4.3. Research strategy and methods	41
4.3.1. Stage A - Development	43
4.3.2. Stage B – First testing on test case	43
4.3.3. Stage C – Second testing on test case	44
4.3.4. Stage D – Third testing on test case	44
4.3.5. Artefact delivery	44
4.4. Data collection and analysis	45
4.4.1. Stage A - Development	45
4.4.2. Stages B, C, and D – First, second and third testings	46
4.5. Research project validation as a Design Science study	47
4.6. Summary	48
CHAPTER 5 Conceptual framework development	50
5.0. Introduction	50
5.1. PPPs should be evaluated using multiple perspectives	50
5.2. Project Success as a departing point for developing the conceptual framework	51
5.3. Refinement of the Project Success understanding	53
5.3.1. Success criteria should be based on stakeholder expectations	53

5.3.2. Success criteria are subjective	54
5.3.3. 'Stakeholder satisfaction' is not a success criterion	55
5.3.4. Existing success frameworks can be represented as sets of stakeholders' expectations	56
5.3.5. A successful project is different from a perfect project.....	61
5.3.6. Proposed refined Project Success definition	62
5.4. Stakeholders' expectation framework for Project Success	63
5.5. The Project Success judgement process for multiple perspectives	64
5.5.1. Meaning of success in Project Success concept.....	64
5.5.2. Application of Turing's test approach to the Project Success understanding	65
5.5.3. What successful projects do have in common.....	68
5.6. Project Success Evaluation Pyramid Model (PSEPM)	73
5.7. Summary	76
CHAPTER 6 Presenting the conceptual framework to experts	78
6.0. Introduction.....	78
6.1. Development of prototype evaluation method	78
6.2. Feedback about the context of PPP evaluation	82
6.3. Feedback about the proposed conceptual framework	87
6.4. Prototyping exercises	89
6.5. Summary	92
CHAPTER 7 Testing the evaluation method on three PPP projects	93
7.0. Introduction.....	93
7.1. Development of the final evaluation method.....	93
7.2. Test case 1 – Social housing project	96
7.2.1. Test case 1 overview	96
7.2.2. Test case 1 stakeholder identification	99
7.2.3. Test case 1 stakeholder workshop analysis	100
7.2.4. Test case 1 evaluation conclusion.....	106
7.2.5. Refinement insights from test case 1.....	107
7.3. Test case 2 - Transport PPP project	110
7.3.1. Test case 2 overview	110
7.3.2. Test case 2 stakeholder identification	111
7.3.3. Test case 2 stakeholder workshop analysis	113
7.3.4. Test case 2 evaluation conclusion.....	124
7.3.5. Refinement insights from test case 2.....	126
7.4. Test case 3 – Health PPP project	133
7.4.1. Test case 3 overview	133
7.4.2. Test case 3 stakeholder identification	134
7.4.3. Test case 3 stakeholder workshop analysis	136
7.4.4. Test case 3 evaluation conclusion.....	144
7.4.5. Refinement insights from test case 3.....	146
7.5. Analysis of the testing on the three test cases	149
7.6. Summary	151
CHAPTER 8 The artefact - PPP ex-post impact evaluation method	153

8.0. Introduction	153
8.1. Overview of the PPP ex-post impact evaluation method	153
8.2. Evaluation setup	154
8.2.1. Stakeholders identification	157
8.3. Stakeholder workshop preparation	160
8.4. Stakeholder workshop	161
8.5. Stakeholder workshop analysis	164
8.6. Evaluation conclusion	171
8.7. Feedback from experts regarding the evaluation method	173
8.8. Summary	174
CHAPTER 9 Discussion	175
9.0. Introduction	175
9.1. About evaluation theory	175
9.2. About the PPP context	177
9.3. About the test cases	178
9.4. About the evaluation method	179
9.5. About the design process	180
9.6. Implications outside the scope of the study	181
CHAPTER 10 Conclusion	183
10.0. Introduction	183
10.1. Research problem	183
10.2. Objective 1: Develop a conceptual framework that informs a PPP evaluation	184
10.3. Objective 2: Develop, test and refine a method that can be used to conduct a PPP ex-post impact evaluation based on the conceptual framework	185
10.4. Implications in theory and practice	186
10.5. Study limitations	189
10.6. Recommendations and future research	190
References	192
Appendix A: ethics approval, consent form and plain language declarations	200
Appendix B: evaluation forms templates	205
Appendix C: presentation employed in interviews	208
Appendix D: prototype evaluation method procedure	210

LIST OF FIGURES

Figure 1-1: Research strategy presented as design iterations.....	7
Figure 2-1: Typical PPP project structure.....	12
Figure 4-1: Information System Design Science framework.....	39
Figure 4-2: Design Science framework applied to this research study.....	41
Figure 4-3: Research strategy presented as design iterations.....	42
Figure 4-4: Conceptual flowchart of the research project.....	42
Figure 5-1: Theoretical relationship between semantic concepts of Project Success.....	54
Figure 5-2: Link between successful project implementation and stakeholder's expectations.....	57
Figure 5-3: Link between Project Success, logical framework method, and stakeholder's expectations	57
Figure 5-4: Link between project's micro and macro view, and stakeholder's expectations.....	58
Figure 5-5: Link between The Square root framework, and stakeholder's expectations.....	59
Figure 5-6: Link between multidimensional Strategic Project Success framework, and stakeholders' expectations.....	59
Figure 5-7: Stakeholders' expectations over the performance framework (example)	60
Figure 5-8: Link between Project Status Model and stakeholders' expectations	61
Figure 5-9: Characterisation of projects in terms of stakeholders' expectations.....	62
Figure 5-10: Example of one stakeholder's expectations profile in terms of criteria, project lifecycle, and relevance.....	63
Figure 5-11: Representation of the Turing's approach for reducing the grey/discussion area in machine thinking dilemma	66
Figure 5-12: Project as an endeavour versus project as a system.....	68
Figure 5-13: Representation of the project success dilemma mirroring Turing's approach	69
Figure 5-14: Project Success Evaluation Pyramid Model (PSEPM)	74
Figure 6-1: Prototype evaluation method (details are described in Appendix D)	81
Figure 6-2: Development of the prototype evaluation method	82
Figure 7-1: Representation of how the insights from the test cases inform each other.....	93
Figure 7-2: Final evaluation method (fully described in Chapter 8)	95
Figure 7-3: Test cases' contribution to the final evaluation method development	96
Figure 7-4: Project lifecycle stakeholder diagram (Social housing test case)	99
Figure 7-5: Incentive-based stakeholder map with the identified stakeholders (Social housing test case)	100
Figure 7-6: Evaluation criteria for project stakeholders separated by critical and non-critical (Social housing test case)	101
Figure 7-7: Stakeholder criteria separated by the influence of the PPP model (Social housing test case)	104
Figure 7-8: Incentive-based map with stakeholder judgement over the three judgement principles 'willingness for repletion', 'no excessive negative outcome', 'no evident better alternative' (Social housing test case).	105
Figure 7-9: Project lifecycle stakeholder diagram (Transport test case)	112
Figure 7-10: Incentive-based stakeholder map with the identified stakeholders (Transport test case)	113
Figure 7-11: Example of consistency diagram for the Treasury Ministry (Transport test case).....	114
Figure 7-12: Evaluation criteria for project stakeholders separated by critical and non-critical (Transport test case)	115
Figure 7-13: Stakeholder criteria categorised by theme (Transport test case)	118
Figure 7-14: Descriptive statistical analysis of impacts at the project level (Transport test case).....	120

Figure 7-15: Descriptive statistical analysis of critical impacts at the project level (Transport test case)	120
Figure 7-16: Stakeholder criteria categorised by PPP effect (Transport test case)	121
Figure 7-17: Descriptive statistic analysis of the PPP effect on the project (Transport test case)	122
Figure 7-18: Incentive-based map with stakeholder judgement over the three judgement principles (Transport test case)	124
Figure 7-19: Incentive-based stakeholder map	130
Figure 7-20: Project lifecycle stakeholder diagram (Health test case)	135
Figure 7-21: Incentive-based stakeholder map (Health test case)	136
Figure 7-22: Evaluation criteria for project stakeholders separated by critical and non-critical (Health test case)	137
Figure 7-23: Stakeholder criteria categorised by PPP effect (Health test case)	139
Figure 7-24: Incentive-based map with stakeholder judgement over the three judgement principles (Health test case)	144
Figure 8-1: Final evaluation method	154
Figure 8-2: Possible milestones for assessment application form (to be used by the evaluator)	155
Figure 8-3: Prototype of the rationale summary table (to be used by the evaluator)	156
Figure 8-4: Example of Incentive-based stakeholder map applied to a PPP hospital project	157
Figure 8-5: Example of a lifecycle stakeholder diagram applied to PPP hospital project	159
Figure 8-6: Example of a network diagram to describe stakeholder consistency analysis	165
Figure 8-7: Example of critical criteria analysis applied to a transport project	166
Figure 8-8: Example of theme criteria analysis applied to a transport project	166
Figure 8-9: Example of PPP effect analysis applied to a transport project	169
Figure 10-1: Project Success Evaluation Pyramid Model (PSEPM)	185

LIST OF TABLES

Table 3-1: Summary of evaluation approaches	24
Table 4-1: Design Science evaluation methods	40
Table 4-2: Design Science research contribution types	40
Table 4-3: Summary of data collection and data analysis	45
Table 4-4: Experts and practitioners involved in Stage A – development	46
Table 4-5: Test cases description for research stages B, C, and D	47
Table 5-1: Summary of evaluation approaches (summarised version of Table 3-1)	51
Table 6-1: Design inputs from evaluation context themes	79
Table 6-2: Design inputs from method feedback themes	79
Table 6-3: Design inputs from prototyping exercise observations	80
Table 7-1: Social housing test case evaluation summary	106
Table 7-2: Summary of PPP benefits and PPP problems (Transport test case)	122
Table 7-3: Transport test case evaluation summary	124
Table 7-4: Identified benefits of the PPP model (Health test case)	140
Table 7-5: Identified problems of the PPP model (Health test case)	141
Table 7-6: Identified 'avoided Traditional Procurement problems' (Health test case)	141
Table 7-7: Selected PPP benefits and PPP problems (Health test case)	142
Table 7-8: Health test case evaluation summary	144
Table 7-9: Summary of the three test cases	149
Table 7-10: Summary of data analyses applied to test cases	150
Table 7-11: Design inputs for evaluation method development and refinement	150
Table 8-1: Incentive-based stakeholder map categories	158

Table 8-2: Example of PPP benefits recognised by the stakeholders of a health project	170
Table 8-3: Example of an evaluation summary applied to a health project.....	171

LIST OF EQUATIONS

Equation 5-1: Project Success of stakeholder S	64
Equation 5-2: Project Success expression.....	64
Equation 5-3: A gap defined as the difference between a perception and an expectation.....	74
Equation 5-4: The sum of all stakeholder expectations should allow a feasible project	75
Equation 7-1: PPP problems and benefits	122
Equation 7-2: PPP benefits and PPP problems	127
Equation 7-3: Disaggregation of PPP problems and benefits	127
Equation 7-4: Value of disaggregated PPP benefits and problems	128
Equation 7-5: Value of disaggregated PPP benefits and problems	142
Equation 8-1: PPP problems and benefits	164
Equation 8-2: Value of PPP benefits and problems isolating PPP extra cost.....	171

LIST OF ABBREVIATIONS

PPP context

AHP: Analytical Hierarchy Process
 DB: Design and Build
 DOOT: Design, Own, Operate, and Transfer
 DBMT: Design, Build, Maintain, and Transfer
 DOT: Design, Operate, and Transfer
 EPC: Engineering, Procurement, and Construction
 IRR: Internal Rate of Return
 NPV: Net Present Value
 PPP: Public-Private Partnership
 SPV: Single Purpose Vehicle
 O&M: Operation and Management
 TP: Traditional Procurement
 VfM: Value for Money

Research and design process

AI: Artificial Intelligence
 CSF: Critical Success Factor
 DHHS: Department of Health and Human Services
 DS: Design Science
 DSR: Design Science Research
 DTF: Department of Treasury and Finance
 EC: Evaluation context
 HP: Health project test case
 IS: Information systems
 KPI: Key Performance Indicators
 MDB: Manage Design and Build
 MF: Method feedback
 NPV: Net Present Value
 Public: Public sector stakeholder group (only Table 3-1)
 PE: Prototype exercise
 Private: Private sector stakeholder group (only Table 3-1)
 PSEPM: Project Success Evaluation Pyramid Model
 SH: Social housing project case study
 TP: Transport project test case
 User: User and beneficiary stakeholder group (only Table 3-1)

Spanish translations (for Chilean test case in Chapter 7)

EGIS: Entidad Gestionadora Social (social housing developer)
 Coordinacion de Concesiones de Obras Públicas (coordination of public infrastructure concessions)
 MINVU: Ministerio de Vivienda y Urbanismo (Housing and Urbanism Ministry)
 MOP: Ministerio de Obras Públicas (Public Infrastructure Ministry)
 Techo NGO: Roof NGO (formerly 'Un Techo para Chile')

CHAPTER 1

INTRODUCTION

1.0. Introduction

This thesis presents the research study entitled ‘ex-post impact evaluation method for Public-Private Partnership projects from the perspective of multiple stakeholders’. This evaluation method is grounded in the project management discipline. The method can guide private companies, policy consultants, and government agencies to perform ex-post impact evaluations of existing and future PPP projects. The main feature of the evaluation method is the consideration of multiple stakeholders’ perspectives for reaching an evaluation judgement. The study employs mostly the Australian and Chilean Public-Private Partnership context to develop and test the evaluation method.

This introductory chapter is an overview of the study, starting with the research background in section 1.1. The research problem (section 1.2) leads to the research question (section 1.3). The research rationale is detailed in section 1.4, and a summary of the research strategy is presented in section 1.5. Section 1.6 argues the significance of the research and in section 1.7 the structure of the thesis is presented.

1.1. Research background

The Public-Private Partnership (PPP) model of procurement is a way to deliver public infrastructure using private funding and managing risk for public purposes. PPP projects have been developed in many sectors, such as health, education, water supply, transport, electric power, etc. Many countries like Australia and the UK have been using this model since the 1980s with good results especially in terms of efficiency; delivering on average more projects on time and on budget (Hevner, March, Park, & Ram, 2004; National Audit Office, 2013; Peter Raisbeck, Colin Duffield, & Ming Xu, 2010). This type of partnership between public and private sector has been promoted worldwide especially in low-income countries to improve their economic performance (World Bank, Asian Development Bank, & Inter-American Development Bank, 2014, p. 33).

PPP projects are usually classified based on the private sector’s level of engagement (Dewulf, Blanken, & Bult-Spiering, 2012). The spectrum considers on one end the private company just constructing the public infrastructure and on the other end the private entity overseeing, conceptualising, designing, constructing, owning, operating and then transferring to the public sector the whole project after finishing the operational stage. The involvement of the private entity in the project determines the type of PPP, which can include design & build (D&C); design-build-operate (DBO); design-build-operate-transfer (DBOT); design-build-own-operate-transfer (DBOOT), among others. Concession contracts are also a PPP classification when a private entity operates and invests in a public infrastructure during a fixed period.

Infrastructure projects (developed using a PPP model) can be classified as economic or social (Australian Government, 2016, p. 7). Economic infrastructure projects refer to projects that have essentially economic purposes; like ports, highways, power plants, and water treatment plants. Social infrastructure projects have the aim of creating social welfare in the short term; such as public housing, educational buildings, prisons, and health infrastructure.

Low, middle and high-income countries employ PPPs with special institutional structures. Australia, UK, and Chile have special units that deal with PPPs. They support, promote, award, and develop PPPs with private companies. For example, in Australia, ‘Partnerships Victoria’ has developed infrastructure projects for the State of Victoria with a total investment of AU\$ 12.4 billion since the year 2000 (State Government of Victoria, 2016). In the UK, the Private Finance Initiative (PFI) has awarded 466 PFI

schemes between 1992 and 2008 with a total investment of USD 51.1 billion (Hellowell & Pollock, 2010). And in Chile, the 'Coordinación de Concesiones de Obras Públicas' (Coordination of public infrastructure concessions) has commissioned PPP projects between 1993 and 2014 for USD 15 billion mostly in transport (Ministerio de Obras Públicas, 2016). From a global perspective, the Organisation for Economic Co-operation and Development (OECD) states that in OECD countries PPP projects represent currently around USD 0.8 trillion and an equal amount 'is in the pipeline' (OECD, 2016).

For the taxpayer, the PPP model is an alternative to Traditional Procurement for the provision of public services. In the traditional model, the government has to finance, conceptualise, design, manage, operate, and maintain the facility, outsourcing only the construction. The PPP model makes the private sector to finance, design, construct, and operate (or maintain) that same facility. The private entity recovers its investment through user fees or by direct government transfers. In any case, the taxpayer must ultimately pay for the service.

The added value of PPPs is the main reason for adopting this procurement method. In terms of project management indicators, PPP projects have been delivered in less time, at a lower cost and in some cases at a better quality (Peter Raisbeck et al., 2010) when compared to traditionally procured projects. Innovation has also been improved in PPPs mainly because of the private incentive to achieve better results, especially in terms of cost. Finally, PPPs have had a better response to risk events, because the allocation of risk is made according to the capacity of the entity to manage the risk; in this case the public or private entity (KMPG, 2015, p. 12).

PPPs have several evaluation and decision milestones over their lifecycle. According to the European PPP Expertise Centre (2011a), the most common are:

- Investment assessment;
- Assessment of PPP option;
- Selection of procurement method and bid evaluation criteria;
- Evaluation of tenders and PPP contract award;
- Monitoring of project delivery and service output; and
- Ex-post evaluation after the operation stage has finished.

Each evaluation milestone has a specific purpose, and they can be divided into approvals and reviews. For example, the investment assessment and the procurement method selection are intended for continuing the project cycle, authorising further actions. Conversely, the monitoring review and the ex-post evaluations aim to review, assess, and take lessons from the project that is already operating. This thesis will focus on the ex-post evaluation milestone that, in theory, takes place after the project has finished, but in practice, it can be conducted throughout the operation stage. Even though ex-post evaluations are considered relevant within the PPP context (European PPP Expertise Centre, 2011a; Oliveros-Romero & Aibinu, 2019), the research around its development and application is not extensive. The next sub-section describes the problems that PPPs are facing, to argue why PPP projects should evaluate ex-post, their impacts on the stakeholders.

1.2. Research problem

Currently, PPPs have been criticised in the media, social media and in the academic literature, notably regarding their limited contribution to real public welfare (Heald, 2003). The arguments are related to (1) the amount of profit that is conferred to the private entity, (2) the long-term implications for the public sector in term of service payments (National Audit Office, 2018), and (3) the quality of the service which can be driven by lack of performance measurement (J. Liu, Love, Davis, Smith, & Regan, 2013), among others. For example, the international organisation of work unions (PSI) has launched a report against PPPs stating that they are expensive and inefficient compared to a public alternative

(Hall, 2015). In the UK, the restricted access to actual PPP contracts, called Private Finance Initiative (PFI), due to 'commercial confidentiality' has given rise to criticism about the transparency of PPPs (National Audit Office, 2018). "...show the true cost of PFI", an article in 'The Guardian' states, referring to the unknown cost of PPP projects (Pollock, 2012). In New South Wales, Australia, the bankruptcy of the private company that was building a social housing PPP project opened the discussion about how reliable are private companies that deliver public welfare (Rogers, 2013).

From a taxpayer point of view, PPP projects should be monitored closely like any kind of public investment, especially if a private company is making profit out of it. One relevant instrument to protect the taxpayers' interest is an ex-post impact evaluation. This evaluation should state what was done correctly, what mistakes were made, and what was not achieved (Owen & Alkin, 2007). The problem is that a PPP ex-post evaluation is difficult to perform because usually (1) projects are large and complex, (2) there are different perceptions of the delivery and the impacts (Gao & Zhao, 2020), (3) the time horizon for the evaluation in some cases is more than 20 years, and (4) there are political, economic, and ideological incentives that create biases towards their evaluation (Boyer & Scheller, 2018).

The impact, in terms of welfare, generated by social infrastructure PPP projects is difficult to assess. Economic infrastructure projects, such as ports and power plants, have an economic objective that can be assessed with precise economic criteria such as revenue, price or service efficiency. Conversely, social infrastructure PPP projects pursue immediate social impact, which is harder to measure and is subject to bias. An assessment to determine the real contribution to society of social Infrastructure PPP projects is complex but extremely relevant from the taxpayer's perspective.

A Chilean and an Australian case exemplify the criticism against public welfare achievement of PPPs. Both countries have a mature PPP market and they are recognised as a good environment to develop PPPs (KMPG, 2015; The Economist Intelligence Unit, 2014).

In Chile, social housing projects are developed using PPPs since 2006. Even though the social housing system is not considered formally a PPP, it can be presented as a Manage, Design, and Build (MDB) model which is not Traditional Procurement, nor a privatisation. The private entity called EGIS is a certified company that manages all the process for delivering a social house. This includes engaging the applicants, designing, land searching, construction, and social assistance. Here, private companies have been criticised because they delivered some projects badly located, with building defects and questionable design (Valenzuela, 2009). The argument is that private companies are getting more profit by cutting costs in location and quality. Although these claims might be based on single cases, an ex-post evaluation of the projects could help to understand what went wrong and how to prevent mistakes in the future. Furthermore, whether correct or incorrect, perceptions are often the basis for the rejection of PPP as a procurement method. A new method to assess a PPP project from an overall perspective can provide an answer to the taxpayer's concerns regarding the achievement of public welfare. The evaluation could help to answer questions such as, should private companies continue developing similar projects? is the government capable of dealing with private companies in social housing projects? or, was it better to use a PPP rather than a Traditional Procurement method for the project?

In Australia, the Melbourne Royal Children's Hospital, opened in December 2011, is a social infrastructure PPP project. The hospital was procured with a PPP contract for 25 years of service. As it has been established, the aim of a PPP is to use private resources to create public welfare. In 2014 the media revealed that part of the funding for the hospital was public (Butler & Cook, 2014). The Victorian Funds Management Corporation funded 1 billion dollars to build the hospital using taxpayer's money. Despite the controversy, there are certain questions that seem relevant to answer. Does the source

of funding matter to the project? Is the government capable of managing a private project while having no conflict of interest? Would the Royal Children's Hospital project be a model to follow or to avoid? An assessment that can inform a judgement about how successful the hospital is at the end of its operational stage could support the answers to those questions.

PPPs are judged by multiple positions and sectors, making its assessment very difficult. International organisations like the World Bank promote the correct use of PPPs (World Bank et al., 2014). Countries like Chile and Australia have PPP agendas approving and cancelling projects based on their interpretation of efficiency and public welfare, e.g. Gordon (2016). Users complain about the PPP toll road service fee prices, based on presumable excessive private profits (Sepulveda, 2016). Consultants have positive judgements of PPPs when compared to traditionally procured projects. Some researchers show the historical flaws of them, e.g. excessive non-competitive contract renegotiation (Guasch, 2004). Governmental entities such as the Victorian Government Treasury department have a clear PPP agenda (State Government of Victoria, 2017), while others, such as the National Audit Office questioned the real benefits that PPPs delivered in the UK context (National Audit Office, 2018). Work unions are generally against PPPs presumably by ideological reasons, e.g. Hall (2015). Urban planning research is divided towards mega infrastructure projects. Some authors are against these projects, arguing that they have negative consequences to society and cities (Tarazona Vento, 2017). Others are in favour of PPPs and see them as opportunities to improve public welfare when they are correctly applied (Khasnabis, Dhingra, Mishra, & Safi, 2010). Finally, PPPs can have collateral impacts such as economic growth and employment (European PPP Expertise Centre, 2011b), and in some cases they are able to provide public services that were unlikely to be provided by any other method (Heald, 2003). The multiplicity of actors and views is probably the biggest obstacle for evaluating PPP projects, as the evaluation varies depending on the evaluator and their chosen criteria.

A lack of theory to address the multiplicity of actors in the assessment of PPPs is considered both a problem and an opportunity for research. There is no tailored method for evaluating PPPs that considers the multiple perspectives that need to be part in an evaluation. There are several evaluation methods that include multiple stakeholders, but they are not purposely developed for infrastructure PPP projects.

The problem to be addressed by this research is the lack of a method to evaluate, on an ex-post basis, the impact of an infrastructure PPP project from the perspective of multiple stakeholders. The existing criticism of PPPs regarding their achievement of public welfare supports the decision for exploring new ways to evaluate them. Within the different evaluation methods, ex-post evaluation is considered the most appropriate approach to assess the described PPP problems, especially in regard to the long-term contracts. It would be difficult to forecast ex-ante the PPP problems for the next 10 to 20 years of operation, meaning that an ex-post evaluation is a better option for analysing and revealing the real impact PPPs are having. Moreover, the discussion around PPPs is mainly about their impacts within the society, rather than about their performance in terms of time, cost and quality; therefore, an ex-post impact evaluation is considered a necessary tool for assessing existing PPPs. Finally, the difficulty associated with PPP evaluation due to their multiple perspectives from the different stakeholders sets a constrain for the ex-post impact evaluation. This is considered the problem and also the opportunity for research.

1.3. Research aim & questions

The aim of the research is to improve the development of future infrastructure PPP projects. To pursue the aim, the outcome of the thesis is a developed PPP ex-post impact evaluation method. The research question is framed as:

RQ: How can ex-post impact evaluation of Public-Private Partnership projects be conducted employing the perspective of its multiple stakeholders?

Ex-post evaluation refers to the evaluation that accounts for past actions, providing information for accountability, and learning purposes. Impact refers here to the criterion in which a stakeholder is affected by a project. PPPs are usually mega infrastructure projects with many stakeholders from the public sector, the private sector, the community, and the society (taxpayers). Each of them is affected by one PPP project in many ways, from profit, to urban segregation, to geopolitical risk. All stakeholders have a role and they are impacted by a project according to their own judgement (Copestake, 2014).

The process of designing an evaluation method consists of a theoretical process that involves thinking about the way PPPs should be evaluated, and a practical process in which a real evaluation method is developed and tested for providing a tangible outcome. Therefore, two objectives are formulated to address these processes. They are:

- *Objective 1: Develop a conceptual framework that informs a PPP evaluation*
- *Objective 2: Develop, test and refine a method that can be used to conduct a PPP ex-post impact evaluation based on the conceptual framework*

1.4. Research rationale

Four reasons drive this research: (1) The existing criticism towards PPPs, which is of relevance due to its use of public funds which requires a high level of probity from the governmental entities. (2) There is no tailored (fit-for-purpose) method for ex-post evaluating PPPs. (3) There is no comprehensive analysis of the existing bodies of knowledge capable of addressing a PPP evaluation; they are all fragmented. And (4) there is a need for capturing lesson learned from existing projects and their key insights for the strategic discussion about PPPs. Each reason is discussed in detail.

First, the existing criticism towards PPPs is of great relevance, as it involves public funding, and could suggest a lack of governmental probity that affects the taxpayer. Most of the PPP questioning refers to the impacts that they have over society and over specific issues such as working conditions, taxpayers' money, public service reliability, and political long-term implications. Therefore, the analysis of these impacts is considered the point of departure for this research. Moreover, there are multiple views in favour and against PPPs, which makes the impact assessment even more complex. The issues that were pointed out about PPPs could only be analysed from an ex-post perspective, because the first question to answer is whether they are true. Ex-ante evaluation for new PPP projects could provide no insight on whether existing PPPs are impacting society in the multiple ways it is being argued. This supports the selection of an ex-post impact evaluation as a tool for addressing, exploring, and assessing the existing questioning and counter argumentation of the PPP model.

Second, there is no tailored method for evaluating the impact of PPP projects from multiple stakeholders. Thus, the development of an evaluation purposely designed for the PPP context should have a better performance compared to the existing methods that are not fully applicable to the research problem. This means that the proposed evaluation should deliver more value for legitimising PPPs, capturing lesson learned, and increasing the discussion over the employment of PPP as a procurement method for future public projects.

Third, the existing evaluation approaches are not comprehensively addressed, and knowledge is fragmented in this area. PPP is a multidisciplinary concept comprising of knowledge from economics, management, project management, construction management, public policy, social work, etc. Therefore, this study analyses the different approaches in terms of their capacity to address the singularities of PPPs.

Finally, there is no structured procedure to capture lessons learned from existing projects. Moreover, the discussion about the use of PPPs in the long term from a philosophical and political view has its support on specific technicalities (i.e. cost and Value for Money), unstructured experience, and some ideological statements. More research in this area should help to improve the theory to understand PPPs better. The evaluation from multiple stakeholders can help to address this need.

1.5. Research strategy

To answer the research question, the Critical Realism (CR) philosophical assumption (Carlsson, 2006) and the Design Science (DS) approach (Hevner et al., 2004) are adopted. Critical Realism stands as a bridge between the traditional Positivism and Constructivist approaches, allowing the existence of real objects (positivist view) and subjective experiences (constructivist) to be together in the same reality (Hoddy, 2019). This ontology is consistent with the methodology of Design Science, that generates knowledge through the creation of designed artefacts that solves problems in the real world.

The reason for choosing the CR philosophical assumption and the DS research approach is that the main outcome of the research is an evaluation method designed to solve an existing problem in the real world, in addition to its contribution to existing knowledge. The problem is solved by designing an 'artefact', which has a specific purpose, it changes the surrounding nature in which it is applied, and it can have different levels of abstraction such as constructs, theories, models, instantiations, frameworks, procedures, tools, physical objects, etc. (Hevner et al., 2004) Design Science differs from the traditional behavioural approach that has the aim of describing nature but without changing it (Simon, 1969). In this case, the PPP evaluation method is the artefact to be designed.

The research structure follows the Design Science approach, generating the artefact through several iterations between develop/build processes and evaluate/justify processes (Figure 1-1). Every iteration will refine the artefact until it can be delivered. This research considers four iterations for producing the PPP ex-post impact evaluation method (the artefact).

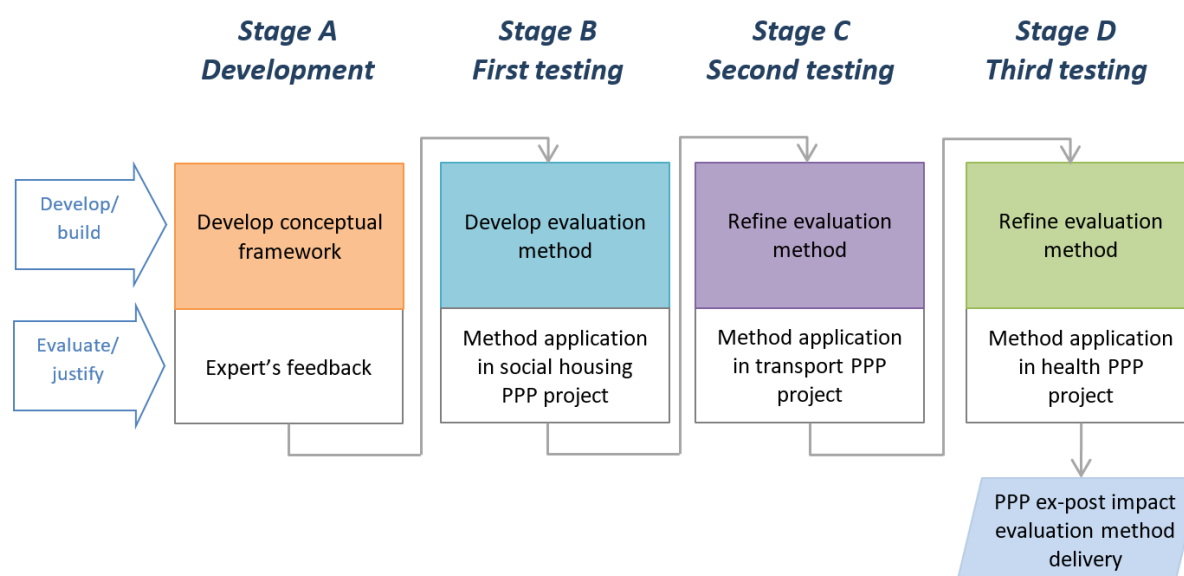


Figure 1-1: Research strategy presented as design iterations

[Source: Author's]

The first iteration has two processes: the develop/ build is for developing the conceptual framework and the evaluate/justify is for assessing the framework by receiving feedback from PPP experts and practitioners. The second Iteration involves the development of the method and its test on a PPP social housing test case. The third iteration is for refining and testing the method on a PPP urban toll road test case. The fourth iteration involves refining the method and testing it again on a PPP hospital test case.

The three test cases were purposely selected for testing and improving the artefact using a diversity of scenarios. They all present a different array of stakeholders and different sets of contractual arrangements that provide enough diversity to stress the developed artefact. Moreover, the selected PPP projects are social infrastructure projects which are even more difficult to assess because they do not follow clear economic objectives (such as power plants, water treatment plants, etc.). The evaluation method is designed for a broad range of PPP projects in many areas, regions, and industrial sectors; therefore, the three test cases were selected to emulate that range by including social housing, health, and urban transport. The commonality between the test cases is their complex stakeholder interaction, which is one of the obstacles for assessing the impact of PPPs. The testing and refining process incorporates all the lessons learned (refinement insights) into the final artefact delivery.

Chile and Australia were chosen for testing and refining the developed evaluation method. They are considered mature economies for PPP development and they provide two additional dimensions to the pursued diversity for stressing the evaluation method. First, they differ in terms of wealth, being Australia a high-income country with a GDP per capita of approximately USD 60,000, whereas Chile is a middle high-income country with and GDP per capita of only USD 15.000 (World Bank national accounts data, 2020). Second, they are culturally different, Australia with an Anglo Saxon background, and Chile with a Latin cultural background that includes, for example, a Napoleonic legal system (contrasting Australia's common law system)(Saint, 2016). The social housing project and the urban toll road project are located in Chile, and the hospital is located in Australia.

Consistent with Design Science approach, the outcomes of this study are (1) the artefact (ex-post impact evaluation method) that solves a problem in real-world; (2) the theory behind it, which

contributes to knowledge; and (3) the methodology for designing an artefact in the Project Management discipline, which is also a contribution to knowledge.

1.6. Significance of the research

The research contributes to the project management body of knowledge, to public policy evaluation, and to project management practice. The study also contributes to the contexts in which the evaluation method was tested: social housing, transport, and health.

The research contributes to the Project Management body of knowledge because the conceptual framework is based on the Project Success concept that has been researched and discussed by many authors during the last 70 years (Atkinson, 1999). The measurement of success is considered difficult because it is related to human perceptions (de Wit, 1988), so to advance the discourse in this topic is a theoretical contribution. Moreover, the research contributes to the topic of stakeholder management that has increased its relevance in the project management discipline (Davis, 2013, 2016).

From a public interest perspective, the research contributes to the public policy evaluation body of knowledge by providing an evaluation method that can assess past projects; improving the mechanism for controlling public expenditure. The use of this method can promote big changes in the long-term strategy for infrastructure investment. The relevance of these types of evaluation is discussed in the evaluation discipline by several evaluation societies around the world (American Society of Evaluation, European Evaluation Society, Australasian Evaluation Society, etc.).

The study has practical implication, as the main outcome of the research is an evaluation method that can be used beyond the academic domain. It provides a procedure that can be applied in real-world projects by practitioners from the public and private sector. The evaluation method contributes to knowledge and practice because it helps the development of better projects and better long-term strategic decisions.

Finally, the research is also significant in the contexts where the evaluation method was tested. Through the design process, the evaluation method was applied to three real PPP projects: a social housing project, a transport project, and a health project. Each application generated an evaluation that provides relevant information about how those projects are impacting their stakeholders. These are valuable insights for decision-making in the context of each project.

1.7. Thesis structure

This thesis has 10 chapters. This introductory chapter presents the research background, the research problem, the research question, the research rationale, the research design, the significance of the research, and the thesis structure.

Chapter 2 introduces the context of PPP projects in which the evaluation method is developed; including the types of PPPs, the procurement systems, the success factors, the advantages, the problems, the current market in the world, among others.

Chapter 3 is the literature review chapter that confirms the gap in knowledge by presenting a broad and critical review of the different existing bodies of knowledge that could be employed for performing an ex-post impact evaluation of PPP projects.

Chapter 4 describes the research strategy and methods; it reviews the Design Science concept and shows in detail the data collection and analysis.

Chapter 5 is the answer to the first objective by developing the conceptual framework 'Project Success Evaluation Pyramid Model'.

Chapter 6 presents the feedback provided by PPP experts towards the developed framework (Project Success Evaluation Pyramid Model). The chapter also includes the results of the first prototype exercises that are part of the evaluation method development.

Chapters 7 answers to the second objective of developing, testing and refining a PPP evaluation method. Three sections present the results of the design process through three test cases: (1) a PPP social housing project in Chile, (2) a PPP toll road project in Chile, and (3) a PPP hospital in Australia.

Chapter 8 is the main deliverable of the thesis: the procedure for performing a PPP ex-post impact evaluation with steps, instructions, and diagrams.

Chapter 9 discusses the results of the research focusing on the evaluation theory, the PPP context, the test cases (assessed PPP projects), the evaluation method, the design process, and the implications outside the scope of the research.

Chapter 10 presents the conclusion of the study that includes the implications to theory and practice, the study limitations, and the recommendations for future research.

CHAPTER 2

THE BROAD CONTEXT OF PUBLIC-PRIVATE PARTNERSHIPS

2.0. Introduction

This chapter aims to broadly introduce the PPP topic, including their characteristics, problems, and the contexts in which they operate. It contributes to the understanding of the PPP complexities, describing issues that will be part of the development of the proposed evaluation method. It also includes a description of how the PPP model works in the two countries that are used for testing and refining the evaluation method: Chile and Australia.

The chapter reviews PPP characteristics in section 2.1, PPP procurement issues in 2.2, PPPs in the Chilean context (section 2.3), and PPPs in the Australian context (section 2.4).

2.1. PPP characteristics

Public-Private Partnerships are a specific kind of projects in which the public sector delivers a public service (infrastructure, service, or products) employing private funding and expertise. The private sector joins the partnership to get profit through future service fees and/or public transfers. The PPP procurement model is a middle step between a pure public provision system and a complete privatisation of the service provision. There are different types of PPPs and they can be classified by the level of interaction and involvement that the private sector has in the project (Dewulf et al., 2012), named:

- **Traditional Procurement**: Contract in which the public authority hires a private company to build a specific facility or provide a product or service. This is not considered a PPP.
- **Design-Build (D&C)**: Contract in which the public authority makes a bidding process that includes the design of the infrastructure and the subsequent construction, following a pre-set design or requirements presented by the public authority.
- **Design-Built-Operate (DBO)**: Contract in which the public-sector tenders to a consortium (a group of companies) the function of designing, building, and then operating the service for a fixed period. In the operation stage, the private entity is entitled to recover its investment in designing and building.
- **Built-Operate-Own-Transfer (BOOT)**: Like the DBO, the asset is operated and owned by the private consortium. This option is considered when a clear fix public service provision is required, allowing the private company to adjust its lifecycle investment for providing that service in the most efficient way.
- **Operation-Maintenance and Service Contracts**: A contract in which the public sector hires a private company for operating an existing infrastructure or service. Mayor investments are public sector's responsibility.

There are other types of PPPs in which the private sector and public sector cooperate or have a formal relation, named:

- **Passive private investment**: To make private infrastructure or service, available for a public purpose (government operation)
- **Joint Ventures**: To create a company owned by the government and a private entity, taking both responsibilities for the service or product delivery.
- **Concessions**: The public sector awards a private company the full responsibility for the operation, maintenance and investments of a specific infrastructure or service for a fixed period. Usually, concessions are made for natural monopolies, such as sewerage services, or ports. The concession period is around 25 to 30 years.

- Passive public investment: Inserting public funding into private companies that meet the requirements of a specific public policy, such as innovation, environmental improvement, geographic zone development program, etc.
- Community-based provision: Employing any possible arrangement to meet communities' need, and in which the government cannot deliver it such as water management in rural areas, specific infrastructure, etc.

Infrastructure projects can be classified as economic or social. Broadly, economic infrastructure projects are projects that have essentially the economic purposes of increasing economic wealth, while social infrastructure projects aim to deliver social welfare. Within the context of PPPs, the Australian government makes a clearer distinction between these two types of infrastructure. According to the Australian National PPP policy, economic infrastructure refers to the facilities that are not supposed to be delivered with public funds, and instead, they are to be economically sustainable through services fees such as ports, power plants, and highways. Conversely, infrastructure that needs to be paid by the government using direct transfers or via facility availability fees is considered social infrastructure such as prisons, schools, hospitals, public buildings, etc. (Australian Government, 2016, p. 7). For this research, the urban toll road test case is considered a social infrastructure project because even though it works with service fees, its main purpose is to provide transportation within a city, which is much more closer to the broad understanding of social infrastructure. This classification is not essential for the research but is important to clarify that an urban highway is intended to connect citizens regardless of the payment method, and not just to increase economic wealth.

PPPs are usually financed using a Project Finance scheme in which a Single Purpose Vehicle (SPV) is created for delivering the project. Project Finance provides financing for projects that can be recognised as individual business units, meaning that are able to provide cash flow to its investors regardless of the sponsor company. It is employed in big projects in mining, oil and gas, real estate, and public services. Contracts are created with the government and between the different agents of the systems, including contractor, lenders, insurance, consultants, facility operators, etc. The objective is to create a long-term arrangement mitigating risks and optimising the design, building and operation.

Within a Project Finance structure, many stakeholders interact with clear roles. Figure 2-1 shows the relationship that exists between the entities for providing the desired public service; the project company is structured as an SPV and the service provision is separated between the Engineering, Procurement and Construction (EPC), and the Operation and Manage (O&M). At this point is possible to argue that the different stakeholders can differ in their motivations and, consequently, their assessment of the PPP project (which is the aim of this research). Moreover, this structure does not consider the extended stakeholder network that includes the political sphere, the community, and the users (of the public service); they are not part of the diagram. The multiplicity of stakeholders is considered by this research as the main difficulty for evaluating PPP projects; each stakeholder in the structure can have a different motivation and information for assessing the project.

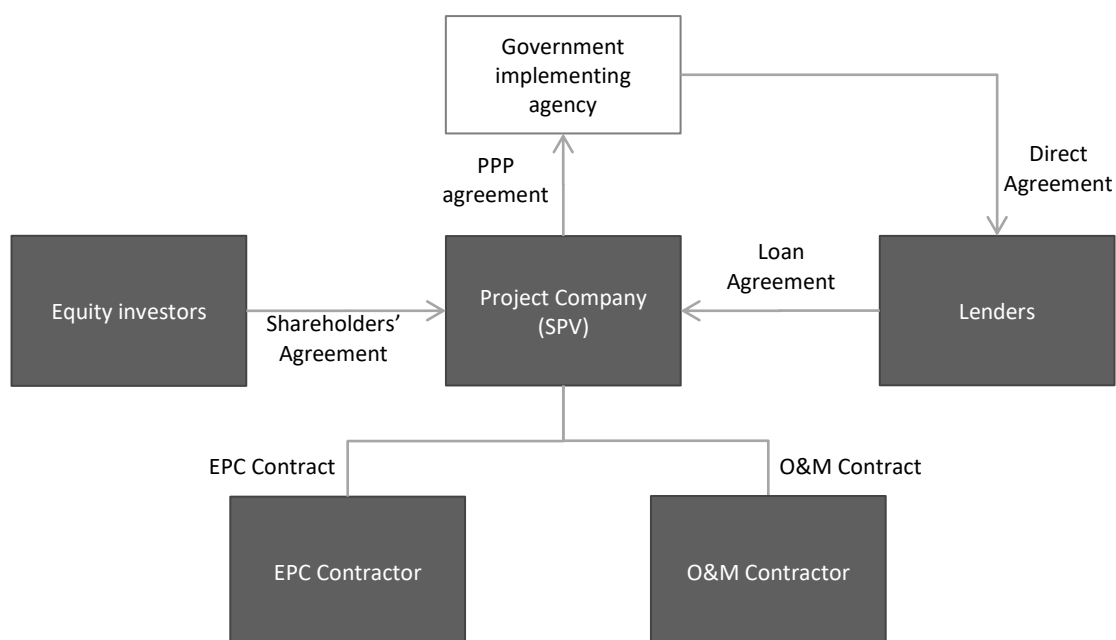


Figure 2-1: Typical PPP project structure

[Source: Adapted from World Bank et al. (2014, p. 51)]

The PPP procurement model is employed in low, middle, and high-income countries. From a global perspective, the Organisation for Economic Co-operation and Development (OECD) states that in OECD countries PPP projects represent currently around USD 0.8 trillion and an equal amount ‘is in the pipeline’ (OECD, 2016). According to the World Bank, the investment in PPP projects has declined in the developing world, with a strong trend towards removable energy projects, and the number of projects is stable (World Bank, 2015). In the long run, it looks that the period 2010-2015 had a higher average annual investment (between USD 100 and USD 150 billion) compared to the period 2015-2019 that has less than USD 100 billion (World Bank, 2019). The OECD, the World Bank, Investment Banks, and other international organisation put great effort into researching, monitoring, and developing knowledge to use the PPP model in the most efficient way. However, reports such as the National Audit Office (2018) from the UK suggest that PPPs are not worth it in every case.

The following sub-sections will describe three aspects of the PPP model that allows a better understanding of the real value PPP projects can potentially generate for their stakeholders: (1) the existing problems of public infrastructure provision, which is the standard by which PPPs should be analysed, (2) the potential advantages or benefits that the use of the PPP model could generate, and (3) the potential problems of employing a PPP model.

2.1.1. Public provision defects

According to Engel, Fischer, and Galetovic (2014), the traditional public provision of infrastructure has many defects. The performance of public projects is the benchmark for evaluating PPPs, as they are the agreed opportunity-cost option. The defects of public provision are (1) deficient selection of projects, (2) low infrastructure maintenance, (3) inefficient price regulation, (4) lobby and corruption, (5) bad institutional design, and (6) renegotiations.

- Deficient selection of projects: Public projects (different from businesses) do not have an economic constraint, meaning that public authorities can select a specific project even if it is not economically feasible. Moreover, they can adjust the social benefit measurement if there are strong political drivers. PPPs that have a pay-per-use method and demand-based

payments make private companies to put risk over the demand estimation and bid only in projects which have financial sustainability. It is theoretically not possible to build ‘white elephants’ with well-designed PPPs (white elephants refer to unnecessary or ineffective mayor infrastructure projects that are built for political reasons).

- Low maintenance: Public projects have low maintenance because the incentive of political authorities is to spend public funding on building new infrastructure, rather than on a strict and efficient maintenance plan. Conversely, PPPs have, by contract, specific quality requirements to operate. The private entity needs to optimise the maintenance cost to improve its bid, combining good design, material, and early maintenance; which is proven to be more efficient than reactive maintenance. Even a mayor infrastructure repairing has better political revenue compare to keeping it in good quality using fewer resources, because it creates milestones and media coverage.
- Inefficient price regulation: For political reasons, public infrastructure does not have the incentive for increasing the fee of its use according to financial sustainability. Moreover, an easy political gain could be to reduce or not increase its fee. This is not consistent with a financial evaluation that should adapt the fees to match the costs of the infrastructure, generating a pay-per-use financial equilibrium.
- Lobby and corruption: Big construction companies can have a big influence over construction policies, advocating for regulations to avoid a competitive market. For example, in Brazil, the bidding process was restricted only to Brazilian companies (Engel et al., 2014, p. 26), which can suggest a lobby of the construction sector to keep an advantage over possibly more efficient international companies.
- Bad institutional design: In some countries, the institutional design does not allow independent monitoring of PPPs. Some PPP units do not have enough budget and they have more responsibilities apart from PPP management. In fact, according to World Economic Forum (2013, p. 9), one of the success factors for PPPs is the negotiation capacity of the public sector, which can be easily surpassed by big international consortiums with vast negotiation skills and expertise.
- Renegotiations: Renegotiations are a problem in big construction projects because they are a hidden source of profit for companies with political ‘assets’. Efficient companies cannot compete with companies that can lower the bid ensuring future renegotiation with the public authority. This is a way of corruption and inefficient use of taxpayer’s money.

These identified defects of public provision are important to acknowledge because they must be considered under any type of procurement method. A PPP evaluation method should consider that some problems could occur with any type of procurement method, not only with the evaluated one.

2.1.2. PPP advantages

Several authors such as Engel et al. (2014); European PPP Expertise Centre (2011b); Public-Private-Partnership in Infrastructure Resource Center (2016) recognise possible PPP benefits. The financial benefits can be (1) budget off-balance sheet, (2) free resources for other social requirements, (3) operational efficiency, (4) introduce competency, (5) set the correct price for users, and (6) distributive aspects. They are elaborated below.

- Budget off-balance sheet: In many countries, the way PPPs are registered in the public accounting is discussed, because they do not affect public debt, but they compromise (in certain cases) future payment from the public sector, which should be restricted and monitored. The discussion in England is that taxpayers do not know how much they should pay in the future to the private PPP consortiums for the existing PPP projects (Heald, 2003).

- Free resources for other social requirements: PPPs that are funded with user fees (pay-per-use) allows the possibility to allocate the original public resources in social programmes that are not economically sustainable or are not possible to achieve with service fee strategies such as inner-city infrastructure and maintenance.
- Operational efficiency: In traditionally procured infrastructure projects, the builders have the incentive to minimise the construction cost, to the detriment of possible savings in the maintenance and operation phase. Conversely, the PPP option employs a unique contract for financing, building and operating/maintaining an infrastructure, which generates the incentive to be efficient and reduce the lifecycle cost expressed in the bidding price; this concept is known as bundling.
- Introducing competency: PPPs provide market prices for the use of public services; creating an incentive for new international investments. They generate more internal competitive markets in countries that can ensure stability for these long-term businesses.
- Set the correct price for users: Publicly managed infrastructure has the incentive to have inefficient (lower) fee prices due to political drivers. Conversely, PPPs could have efficient prices consistent with the maintenance of the infrastructure and investment recovery.
- Distributive aspects: One criticism is that PPPs are not fair with low-income users, as they do not have any differentiated fee. However, in the case that PPPs provide new infrastructure, this argument is questioned, because low-income people benefit from having more infrastructure with less congestion, due to the voluntary payment of some users to use the newly available infrastructure. This is under the assumption that the infrastructure cannot be built with public funding or if it is not a priority.

In addition, non-financial benefits of PPPs are (1) accelerated delivery, (2) enhanced delivery, and (3) wider social impacts

- Accelerated delivery: In the case a PPP option produces an accelerated project delivery compared to its public alternative, the cost of time can be (as it should be) considered under the lifecycle cost comparison. Here, a social discount rate can be applied when comparing, for example, two projects of 20 years of operation in which one (the PPP) starts two years earlier.
- Enhanced delivery: When more effective maintenance is done, other impacts can be assumed. In the case of a highroad, better road maintenance can lead to fewer accidents, less carbon emission, and/or better quality of the vehicle fleet.
- Wider social impacts: They can include innovation brought by including international competition in the bidding process, information on public services' real prices, fiscal programming and certainly, employment, etc.

The PPP advantages should be confirmed or questioned by a PPP ex-post evaluation. The National Audit Office (UK) suggests that in the UK the theoretical PPP benefits are difficult to prove in practice and that there are no suggestions that those benefits are real in the existing projects (National Audit Office, 2018).

2.1.3. PPP's feature is a better risk allocation

In every project, the many risks need to be established, allocated between the parties that are part of the project, and stated in the contracts. In the case of PPPs, the type and amount of risk are allocated between the private sector (consortium), the public sector, and the taxpayer. Bing, Akintoye, Edwards, and Hardcastle (2005) recognised the existing risks for a PPP and they can be classified under (1) construction risk, including defects in design, over costs and delays, (2) O&M risks, (3) performance risk, including service disruption, (4) residual cost risk, cost of asset disposal, (5) political risk, including

macroeconomic uncertainty, and government actions that affects the project (expropriations, public infrastructure that compete with the PPP, etc.), (6) demand estimation risk, and (7) Financial risk, including interest rates and exchange rates.

Risk allocation for PPPs follows the principle ‘each risk should be allocated to whoever can manage it best’ (Irwin, 2007, p. 56). For example, the construction risk in Traditional Procurement is handled by the taxpayers, whereas in PPPs this risk is handled by the consortium. One of the PPP advantages is the bundling, which makes the consortium to handle better the constructions risk, as it is the one that influences the design, construction and operation. In public provision, taxpayers are the ones handling the risk, having almost no influence on any of the decisions that can be made to mitigate that risk (Engel et al., 2014).

2.1.4. PPP success factors

Hardcastle, Edwards, Akintoye, and Li (2005) describe the critical factors for developing PPP from a construction management perspective. Literature review in the discipline shows that there are 19 success factors, including stable macroeconomic environment, political support, transparency in the procurement method, available financial market, a strong private consortium, and competitive procurement process. Success factors can be divided into five groups: effective procurement, project implementability, government guarantee, favourable economic conditions, available financial market. Even though most of these studies use questionnaire methods and expert consultation, instead of experimentation or project-data, they do present a big picture on what PPP should comply with for increasing its success rate, regardless of a clear measurement of ‘successful PPP’.

Government expertise is also considered a success factor, as PPPs are meant to be developed in countries where the public sector can be a feasible counterpart for the PPP negotiation and providing economic macro stability inside the country (World Economic Forum, 2013).

2.1.5. PPP problems and risks

PPPs theoretically provide benefits compared to traditionally procured projects. However, research shows that they are not free of problems (or potential problems). Engel et al. (2014); Heald (2003); National Audit Office (2018); Public-Private-Partnership in Infrastructure Resource Center (2016), among others, have comprehensively addressed the PPP risks and problems. Some of them are:

- High cost of PPP contracts: PPPs have a big bidding cost, and for that reason, PPPs are usually restricted to high investment projects.
- Renegotiations: Guasch (2004) recognises that renegotiation of PPP contracts in Latin America is 54% in road projects and more than 70% in water plants. The average time for the first renegotiation was three years after signing the contract. Most of the renegotiation processes were originated by the companies. This is a problem because renegotiation is an instrument where the company has big power over public authority. Large components or modifications in megaprojects are difficult to procure again in an open process.
- Public budget control: The consequence of having PPPs free from the public budget is that they can avoid public budget control. PPPs are a way to provide infrastructure without restricting public budget; therefore, future payments associated with the PPP can be underestimated in the long term (Heald, 2003).
- Fixed service provision: PPP contracts are long-term obligations for all the parties involved. Then, any change in the scope of the service (for example, in a hospital) is difficult to implement and could have a higher cost.

The technical problems and risks can be linked to the existing criticism that was shown in the research problem section (1.2). The presumably excessive profit conferred to private entities can have multiple

reasons including a deficient risk allocation, a change in the project's nature, a political campaign against PPPs, or a corruption issue in the form of renegotiation. The long-term implications for the public sector in term of service payments is another criticism towards PPPs and could also be a consequence of bad planning, rather than the PPP model. The low quality of the PPP service provision is a criticism that is also difficult to assess, because the quality should be compared with a reasonably comparable public option, and it should depend on unique project conditions.

The PPP problems should be addressed by the PPP evaluation to confirm the relevance of their impacts. Existing research generally presents the problems for the overall PPP system, but a PPP evaluation should try to confirm each of them for every evaluated project.

2.2. PPP procurement issues

A PPP project can be affected by the auction mechanism and by the contract structure, among other factors. This section reviews four issues that can affect the way PPP are developed and, consequently, evaluated: Analytical hierarchy process, demand estimation in fixed-term contracts, the Least Present Value of Revenues model (RPVR), and contracts for non-transport PPP projects (schools, hospitals, and prisons). These factors that affect PPP development are important to review, as they could lead to the recognition of PPP impacts over their stakeholders.

2.2.1. AHP in PPP projects is not recommended

Analytical hierarchy process (AHP) (Saaty, 1987) is a decision-making mechanism that can be used for bidding awarding by evaluating and weighting different criteria. According to Engel, Fischer, and Galetovic (1997), this scoring system is not recommended for PPP projects. In many countries, they are employed assuming that all criteria are relevant and that it is possible to quantify them objectively. However, the problem is in the weighting of the different criteria, which depends on the bureaucrats' judgement rather than on an objective process. In highways, where the scope of the project and the operational stage are well known (compare to hospital, airports, etc.), simple awarding systems (e.g. lowest bid) should generate better outcomes and better biddings. Special conditions, such as profit-sharing with the government above a certain threshold, can create extra complexity to the bid development, which can affect the quality of the proposals.

2.2.2. Demand estimation is a key problem for PPP legitimisation

Design-build-operate PPP fixed-term contracts must deal with the demand estimation risk. In a 20-year contract, demand estimation is very difficult to forecast and, therefore, companies need to allocate a high-risk factor for this issue, increasing the bidding price to mitigate this risk (Vassallo, 2006). This has created one the main controversies in the public opinion towards PPPs; private consortiums are making 'too much profit' (Sepulveda, 2016). The excessive private profit can be considered the trigger for the public opposition to PPPs; which is the non-achievement of public welfare. The problem is that with fix terms contracts, the demand estimation risk needs to be owned by someone. The other option is that the government addresses the demand risk by ensuring to the private consortium a minimum demand for the project; reducing the bidding price. The counterfactor is that the market is not capable anymore to create an efficient selection of projects, because the demand is endured. Under this condition, it is possible to build projects that have no economic sustainability, because they will be profitable with any scenario.

2.2.3. Least Present Value of Revenues (LPVR) mitigates the demand risk

Least Present Value of Revenues (LPVR) auction mechanism, presented by Engel et al. (1997), is a good option for dealing with the demand estimation risk problem. Traditionally, PPP contracts are set by fixing the length of the contract and awarding the project to the consortium that offers fewer toll prices, or alternatively, the toll price is fixed, and the project is awarded to the consortium that offers the least concession length. LPVR auction mechanism awards the project to the company that offers

the least present value of revenues. This means that the income to be pursued by the consortium is fixed, and the project will have a flexible term. This auction mechanism reduces the demand estimation risk because now is not necessary to estimate the demand in terms of time, but only the total demand needed to get the project margin. Additionally, the total cost of the project (and consequently, the total cost to be paid by the users) gets reduced, because there are no 'good or extremely good scenarios' in which the demand is over the expectations; this is a gain in terms of welfare. The use of 'total income' as a standard to measure the awarding of PPPs is a much clearer and simpler view from the perspective of the taxpayer. This auction model makes the project easier to adjust for policy changes and toll prices. The renegotiations are also simpler to address because the income and the construction costs are fixed and the beginning of the contract. Finally, if the government wants the asset back to the public sector prior to the end of the contract, then, the price is easy to estimate, as the difference between the total revenue and actual revenue in present value, plus specific and minor corrections (apportioned of maintenance schedules, or periodic upgrades).

2.2.4. Non-transport PPP projects (schools, hospitals, and prisons)

Health and education PPP models are more complex than transport because the provided service is more difficult to define, measure, and monitor. The operation of a hospital including different department, processes, databases, services, etc. is completely different from a highway that has a unique objective providing a road from point to point, with clear technical standards in term of road quality and availability, and a simple monitoring system of tolls and visual inspections.

There are different models for health infrastructure PPPs. But the main difference is between a Design-Build-Finance-Operate (DBFO) and a Design-Build-Finance-Maintain (DBFM) (McKee, Edwards, & Atun, 2006). Usually, for hospitals, DBFM is the preferred method, as the core health service is provided by the public sector. In this model (DBFM) the maintenance involves the services that need to be done to infrastructure (electrical, floors, lifts, etc.), plus in some cases, it can be also the peripheral services that are not related to the health service (food provision, cleaning, parking management, etc.).

The benefits of using a PPP model for non-transport projects are mainly bundling and innovation. Bundling, which is a result of a single contract for the design, construction and operation; creates efficiency incentives in terms of design and construction. For example, for a PPP hospital, the awarded bid could have designed cheaper and simpler maintenance for walls, floor materials, lighting, windows, etc. Additionally, the use of PPP model can promote business innovation to incorporate commercial areas, supporting services (hotels and restaurants) or fostering the use of parking spaces for the extended area of the infrastructure (neighbourhood).

PPP attributes are important to review because they might affect the way PPPs are assessed. The contract arrangement, the risk allocation, and the business opportunities are examples of factors that can be managed for achieving better PPP results. However, as presented earlier, still the question remains on how to assess those results.

As this study employs the Australian and Chilean contexts, these countries' PPP systems will be reviewed showing the specific conditions by which PPPs are developed.

2.3. Public-Private Partnerships in Chile

The developed PPP ex-post impact evaluation method is tested in Australia and Chile. Therefore, this chapter includes the general review of how the PPP model works in these two countries to provide a broad background on how the ex-post evaluation method could be applied and tested.

Chile has been using the PPP model since the 1990s with a stable law and investment agenda, especially in transport projects. Within the Latin-American and Caribbean context, Chile is first ranked

in terms of the environment to develop PPPs (The Economist Intelligence Unit, 2014), when considering aspects such as legal and regulatory frameworks (bidding processes, PPP regulation, etc.), Institutional framework (Institutional design and PPP contracts), operational maturity (public capacity and experience, risk allocation and awarding processes, etc.), investment climate (political will and distortion, business environment), among others. The Chilean context is relevant for the development of the evaluation method because specific contextual issues could inform the design process and help in the evaluation method testing.

Within the public sector, the PPP model is employed by the different government agencies that need to provide infrastructure for their own objectives. Most of the PPP projects in Chile are developed by the Ministerio de Obras Publicas (Public Infrastructure Ministry), who has a PPP unit called 'Coordinacion de Concesiones de Obras Publicas' (Concession's coordination of public infrastructure). They have most of the expertise in PPP development, mostly on transport projects. Some prisons and hospitals are also under PPP model, managed by Justice and Health ministries, respectively). Additionally, the Housing and Urbanism Ministry employs a PPP approach for delivering social housing projects; delivering more than 3 million dwellings in the last 26 years (MINVU, 2013). However, this approach is not formally recognised as a traditional PPP project and they are not part of the general statistics.

For a country that has a population of 17 million and a GDP per capita of USD 20.000, Chile has a large investment in infrastructure using the PPP model. According to Abedrapo (2016), the PPP model has been regularly used for 20 years, with a total investment of nearly USD 20 billion; 81% in highways, 11% in public building, 7% in airports, and 1% in Dams. The highways under concession system have in total 3575 km, from which 1991 km is part of the 12 concession routes of the main highway that connects Chile from north to south. In terms of health infrastructure, from the total 190 hospitals, three are DBFM PPPs and they are discussing the new agenda for building 18 out of 25 using a PPP model (Chavez, 2018). For the justice system, in 2001, a plan for building 10 new PPP prisons was approved with a total investment of USD 260 million. In the end, only eight were built and, due to cost overruns that were paid by the government, the plan had no subsequent stage (Vedoya & Fernández, 2018).

In terms of evaluation and approval, three governmental entities work together. First, the developer Ministry (Public Infrastructure in most of the cases) should develop the business case of the project, including the PPP model. Second, the Social Development Ministry is the government agency that evaluates any investment that involves public funding (including PPPs) from a social point of view. The Social Development Ministry has procedures and methodologies to assess public investments and approve their development. However, for the case of PPPs, they do not have a role as relevant as with the traditional public investment, and a PPP evaluation is 'unofficially' out of their scope. Third, The Treasury Ministry has the final approval of a PPP, deciding whether they can compromise the warranties, overcome the exposure to risks, and keep the budget control. Additionally, another entity to consider is the 'Contraloría General de la República' (general auditor), who can audit PPP projects. Each PPP contract is ruled by a mandate agreement, a prequalification term, bidding term and its clarifying circulars, the technical and economic bids of the successful tender, and the supreme decree of award.

The bid tendering process is open and non-discriminatory, sufficient asset demonstration from the bidders is required. The reception of bid offers is public and transparent; the evaluation consists of two stages: technical filter and economic valuation. In terms of project finance, there is a free choice of financing (currency, country, term), and no currency insurance is provided by the State. There is also freedom to form consortiums, but a Special Purpose Vehicle (SPV) must be created by the winner. According to Abedrapo (2016) from the Public Infrastructure Ministry, the current form of PPP

includes warranties for the financial gearing by a Minimum Income Guarantee and a variable term of contract (Total Income of Concession), in addition to a special pledge of public work. The income sources come from user payments; mostly tolls which are determined by tender documents. State payments are employed only when necessary. In terms of transparency, every PPP project has available to the public the original prequalification terms, the actual contract, the bidding terms, and in some cases the periodic performance reports of the public services (airport, highways); all at the Public Infrastructure Ministry webpage.

The PPP procurement model is also part of the political agenda and they are affected by political drivers. During the last government period (2014-2018), in which a centre-left coalition was governing, the PPP procurement method was questioned and the PPP agenda with several projects in the pipeline was delayed; specifically, the construction of PPP hospitals. In the parliament, the discussion for stopping the PPP hospitals agenda was due to the claim of higher costs compare to traditionally procured hospitals. Two ministries (Health Ministry and Public Infrastructure Ministry) had differences in the calculation of the construction cost per square meter. For the case of the PPP hospitals, the Health Ministry stated that building the hospital with a PPP procurement method increased the cost by 77% (Sandoval & Villegas, 2014). The Public Infrastructure Ministry questioned that calculation, arguing that the data was not trustable. This confrontation was understood as an ideological position against PPP projects from the Health Ministry; which was not shared by the Public Infrastructure Ministry. This conflict shows how the discussion about PPPs can affect long-term strategies and real welfare at a national level. Moreover, now the actual government (2018-2022) is centre-right and a PPP agenda was presented, which includes investments beyond their period. There is no certainty that the next government after 2022 will continue the PPP agenda.

2.3.1. The Chilean social housing system is a non-traditional PPP model

The Chilean social housing provision has a non-traditional PPP structure (could be described as Manage, Design and Build) due to the large deficit that exists in terms of social housing. More than 30,000 families live in slums concentrated mostly in the three biggest Chilean cities. The actual house policy is focused on giving these families (in the lowest income level) a 45 sqm houses without any debt. The total investment, the size of the house and amount of debt can vary depending on the poverty level. To accelerate the building process, a PPP structure was developed. In 2006, the new legislation for developing social housing projects created the concept 'Entidad gestidora inmobiliaria social EGIS' (social housing development entity). Profit, non-profit, and local government entities can apply to become certified by the central government to be EGIS. The government pays for every eligible family that starts living in a social house. Thus, these entities are in charge of engaging the families with a conceptual project, designing the project, searching for the land, hiring the contractor, be accountable for the house quality, and performing the social assistance to the families. An EGIS makes a profit depending on how efficient they perform the overall process considering a fixed subsidy per house.

The current regulation, the government entities involved, the auction mechanisms, the political discussion, and the specific strategies such as the social housing provision provide an understanding of how PPPs work in the Chilean context. They could inform the way a PPP evaluation should address contextual issues in real projects beyond the general theory of PPP development.

2.4. Public-Private Partnerships in Australia

The evaluation method development includes a testing and refining process on a PPP Hospital in Australia. Therefore, the Australian context is broadly reviewed in this chapter to provide background on how the Australian government manages the PPP model. Projects in this country are usually developed at a state level, having their own PPP units such as Partnership Victoria, NSW Public-Private Partnerships, or Queensland Treasury. The National PPP policy is developed by the joint of different

national and state entities, mainly from the different treasury and finance departments. Their aim is to ensure Value for Money, improve the PPP national policy, increase cooperation between jurisdictions, and promote the use of PPP as a procurement method. All Australian states and territories will consider the PPP option with every project that has a capital cost above AUD 50 million. Australia is ranked as 'mature growth PPP market' with a history of innovation and good practices in PPP development (KMPG, 2015). For example, one innovation is the Capital Contribution Model, that promotes a better risk allocation and Value for Money achievement using partial government payments during construction or after completion. Australia has a forecasted population growth for the following decades, suggesting the need for a strong agenda on infrastructure development that certainly includes PPPs. At a federal level, Infrastructure Australia is an independent body that advice and prioritise the nationally significant infrastructure usually proposed by the different governmental entities (states). This pipeline of projects is also developed with the collaboration between Australian state, territory and local governments.

The Victorian state is considered a leading one with its PPP unit 'Partnership Victoria' under the Department of Finance and Treasury. It manages PPP projects during the conceptualisation, design, and bidding process. It supports and monitors PPP projects or initiatives together with the government agencies that are requiring the infrastructure (health, education, transport, infrastructure, etc.). Partnership Victoria has a portfolio of 29 PPP projects worth \$15.9 billion for roads, hospitals, schools, prisons, and water initiatives. Two projects are completed; 23 projects are in operation, and four are under construction (State Government of Victoria, 2017). Partnership Victoria provides public access to review PPP contracts, specifications and audit reports. For example, the Victorian Auditor-General audited the Royal Women Hospital (RWH) redevelopment project, procured with a PPP. The audit had the objective of assessing the adequacy of the state's planning, procurement, and management of the RWH redevelopment. This objective was addressed by assessing whether: (1) planning for the redevelopment project was adequate and included soundly based advice and recommendations to government on project financing and procurement methods, (2) procurement for the redevelopment project was conducted in line with relevant policy and guidance, (3) the state has been adequately managing its involvement in the construction phase of the project, and (4) correct arrangements established by the state, to manage and supervise the long-term role of the private sector which has manage a total investment of AUD 12.4 billion since year 2000 (Frost, 2008). The audit found that the investment plan was adequate, the private bids (Value for Money) were adequately assessed, and the construction phase was effectively managed. This audit report was developed at the end of the construction phase, leaving out the operation phase.

Politics is also a driver for PPP developments in Australia. For example, the East-West Link toll road project in Melbourne was developed as a PPP during a centre-right state government. In response to protests that occurred against the construction of the highway, the opposition stated that if they were elected, the project was going to be cancelled. The centre-left (opposition) became government and the cancellations of the project cost \$339 million, plus other costs of preliminary work estimated by the (new) opposition to be minimum of \$400 million. The Victorian Auditor-General estimates a \$1.16 billion cost for cancelling the project. The cancellation included the purchase of the consortium that was already established, the assets, the project design, and the early construction (Gordon, 2016). The controversy is between one political party saying that the cancellation of the project is a taxpayer's money waste; while the other party claims that the negotiation and the commitment to major projects cannot be done so close to elections (ABC News, 2015). This kind of issues in PPP projects damage the legitimacy of them, the potential benefits of employing them, and also supports the argument for developing a method to evaluate existing projects in order to assess the procurement of new ones without having this kind of backfires.

The review of the PPP context provides a basis for discussing the research topic; PPP ex-post impact evaluation from multiple stakeholders perspectives. The existing literature about PPP evaluation will be reviewed in the next chapter.

2.5. Summary

This chapter reviews the PPP general context. The purpose is to provide enough contextual knowledge to inform and apply real-world experience into the evaluation method development. It helps to understand PPPs from an economic and management perspective, as well as from a political and governmental perspective.

PPPs are classified by the level of engagement that the private entity has over the project: Design-Build; Design-Build-Operate-Transfer; Design-Built-Own-Transfer; among others. They are financed using a Project Finance scheme; creating a Single Purpose Vehicle (SPV) to develop the project.

To review the PPP characteristics, three aspects are listed: (1) the existing problems of public infrastructure provision, which is the standard by which PPPs should be analysed, (2) the potential advantages or benefits that the use of the PPP model could generate, and (3) the potential problems of employing a PPP model.

- Main public provision problems: (1) deficient selection of projects, (2) low incentive to maintain public infrastructure due to high incentive in building new one, (3) inefficient price regulation due to political drivers, (4) lobby and corruption due to high concentration of big construction companies, (5) bad institutional design to, for example, monitor PPP developments, and (6) renegotiations of projects for companies' advantage.
- Most common (potential) PPP advantages: (1) budget free from the public record, (2) free resources for other social (not economically sustainable) requirement, (3) lifecycle cost efficiency incentive, (4) competency improvement, (5) efficient price selection, and (6) distributive capacity for low-income users (for new infrastructure). Non-financial benefits are (1) accelerated delivery, (2) enhanced delivery, and (3) wider social impacts. Even though PPPs should provide a better risk allocation between the private and public sector, the demand estimation risk (traffic risk) is still a non-solved issue and it creates some of the main critics about PPPs; the excessive profit of private consortiums.
- Potential problems associated to PPPs: (1) high costs of PPP contracts, (2) excessive renegotiations, (3) incentive to avoid public budget control, (4) inflexible service provision, and (5) non-competitive tendering processes.

The way PPP projects are procured is also important to review, as this is a factor that can affect the impacts PPPs generate on their stakeholders. In terms of the auction mechanism, the Analytical Hierarchy Process (AHP) is considered a complex and non-transparent auction option as the weighting process is vulnerable to bureaucratic bias. Some authors recommend a simple 'lowest bid option' when quality and scope can be accurately defined. PPP fix-term contracts have the problem of creating a demand estimation risk, that is covered by a higher bidding cost. Lease Present Value of Revenues (LPVR) is presented as an alternative that creates a flexible-term contract, mitigating the demand estimation risk. Hospital and more complex infrastructure projects (compare to roads) that are procured using the PPP model are differentiated between (1) operate the infrastructure (which include the core service), and (2) maintain the infrastructure (excluding the core services); being the core service usually provide by the public sector (i.e. medical service in hospitals).

For developing the evaluation method, this research employs three PPP projects that are located in two countries: Chile and Australia. Therefore, their PPP contexts are reviewed to explore specific insights that could help the evaluation method design process.

Chile is considered a leading country for PPP development in the Latin-American context, with stability, expertise, and a PPP unit in the Public Infrastructure Ministry. It has near USD 20 billion in investment, mainly in transport projects. The use of PPPs for the hospital provision was questioned by the last government, cancelling the PPP agenda and creating a parliamentary discussion.

Australia is considered a mature PPP market, it has National Policy and state PPP units. Partnership Victoria is one the leading PPP units, it has procured 29 PPP projects with AUD 15.9 billion in investment. The East-West link project in Melbourne is an example of a controversial issue in PPPs, in which the toll road project was cancelled by the newly elected government, costing more than AUD 1 billion.

The next chapter will review the literature about PPP ex-post evaluation, which is the basis for confirming the gap in knowledge and for developing a PPP ex-post impact evaluation method.

CHAPTER 3

PPP EVALUATION FROM MULTIPLE STAKEHOLDER PERSPECTIVES: A LITERATURE REVIEW

3.0. Introduction

In the existing literature, there is no tailored (fit-for-purpose) method for evaluating PPP delivery and impact. Therefore, Chapter 3 provides an in-depth literature review about the different approaches to evaluate a PPP. A broad review was conducted, including many disciplines, and establishing the strengths and weaknesses of each approach for addressing the PPP ex-post evaluation. The gap in knowledge is confirmed with no method capable of conducting an ex-post impact evaluation from the perspective of multiple stakeholders. However, some methods are highlighted as they are useful contributions for developing an evaluation method that can answer the research question.

The chapter is separated into four parts. First, the initial definitions of the main concepts for this thesis are presented in section 3.1. Second, section 3.2 summarises 30 evaluation approaches that could be employed for conducting a PPP ex-post impact evaluation. The third part is the detailed review of the 30 approaches separated into four sections (3.3, 3.4, 3.5, and 3.6). Finally, the gap in knowledge and the usefulness of certain approaches are discussed in the last section 3.7.

3.1. Initial definitions

For clarity, a brief description of the essential concepts employed in this thesis is presented. Nevertheless, the concepts will be discussed throughout the thesis, questioning and refining its meanings.

- **Project:** Traditionally, 'project' refers to a temporary endeavour undertaken to create a unique product, service or result (Project Management Institute, 2013, p. 1). However, this research discusses and proposed a refinement of the concept, understanding it as a system of stakeholders; which will be called a project-system.
- **Evaluation:** For this research, evaluation is broadly understood from the perspective of public policy. Evaluation is the application of empirical social science research methods to the process of judging the effectiveness of public policies, programs, or projects.
- **Impact:** Within the public policy domain, impact can be defined in quantitative or qualitative terms. The quantitative approach refers to the change in a 'Y indicator of interest' due to an intervention. Consequently, the cause for that change will be a matter of a quantitative impact evaluation. Conversely, this thesis employs a qualitative approach (White, 2010) in which the impact refers to the effect that a project has on every stakeholder, measured in terms of one or more preferred criteria relevant to each of them.
- **PPP:** Public-Private Partnership project is a model of procurement in which the private entity employs resources and expertise to provide a public good. In most of the cases, the private entities will design and build an infrastructure, operate or maintain it for a specific period (+25 years), and then transfer it to the public sector. The revenues for the private entity can come from direct transfers or from user service fees.

3.2. Summary of the existing evaluation approaches

After presenting the initial definitions of this thesis, the rest of the chapter focuses on reviewing existing evaluation approaches for PPP projects. They are summarised in this section and then reviewed in detail in the subsequent sections.

30 evaluation approaches were gathered from four disciplines and bodies of knowledge, namely: research on PPP evaluation and monitoring, management-based evaluation and monitoring approaches, program evaluation (public policy) approaches, and international agencies' approaches. Table 3-1 summarises the 30 evaluation approaches including description, the stakeholders involved, and a critic. How the approaches involve their stakeholder is an important variable of analysis because the multiplicity of views regarding PPPs is what makes this kind of projects difficult to assess. The researcher employed four stakeholder group to shows which of them were included in each approach. The stakeholder groups are: private sector (abbreviated as 'private'), public sector (abbreviated as 'public'), users/beneficiaries (abbreviated as 'user'), and community (not abbreviated). It is important to highlight that the groups are just referential, and in real-world, many stakeholders can be part of one group/category (for example, the different government agencies are part of the public sector).

Table 3-1: Summary of evaluation approaches

Reference	Evaluation	Description	Involved stakeholders	Critic/Gap
Research on PPP evaluation and monitoring (section 3.3)				
Heald (2003)	Value for Money (VfM) evaluation	Compare a PPP with Public Sector Comparator (PSC) to determine which gives more VfM	Private, public	It is ex-ante and the PSC can be biased due to the public funding availability
European PPP Expertise Centre (2011a)	PPP ex-post evaluation guidelines	General guidelines for creating an institutional framework and a technical framework	Not specified	The guidelines are very generic with no specific evaluation method
Love and Liu (2014)	PPP performance (whole life assessment)	List of performance measurements that need to be considered during the life cycle of a PPP; based on a Performance Measurement System (PMS)	Private, public, user, community (based on the PMS Performance Prism)	It is a broad approach with general KPIs and does not focus on stakeholders' impact
Management-based evaluation and monitoring approaches (section 3.4)				
De Wit. (1988)	Project success evaluation	Successful project is a positive condition of a project and it includes all the stakeholders	Private, public, user, community	There is no accepted approach for assessing success in practice.
Barclay and Osei-Bryson (2010)	Stakeholder management	A set of tools for managing stakeholders within a project	Private, public, user, community	It is not a proper evaluation approach, but it understands the diversity of stakeholders
Kaplan, Norton (1995) Neely,	Performance Measurement System (PMS)	A set of metrics used to quantify the efficiency and	Private, public, user, community (holistic view of Performance Prism PMS)	Different from an evaluation, it does not provide clear guidance on the decision-making process

Adams, & Crowe (2001)			effectiveness of actions	
Nalewaik and Mills (2017)	Project audit		Audit approach that extends the traditional financial audit	Private, public, user, community Take stakeholders' opinion, but it is based on the auditor
Program evaluation (public policy) approaches (section 3.5)				
Quade and Carter (1989)	Before/after program		Compare the performance before the program started and after it finishes	User, community Applicable for behavioural policies (e.g. drug control), but not for new infrastructure (performance easily will be improved)
	Forecast/actual performance of the program		Compare the forecasted performance without the program and the actual performance with the program	User, community The forecast would be the PSC which has been the subject of criticism in PPP literature and practice because it is unreliable (Heald, 2003)
	With program/without program		Compare results between program target and not target	User, community Not applicable to a PPP because usually, PPPs add new infrastructure to the system
	Experimentation with users		Isolate the program variables to evaluate their effectiveness against a control group	User, community Not applicable to a PPP because usually PPPs add new infrastructure to the system; there is no control group
	Planned/actual results		Compare the planned performance and the actual performance	User, community It does not consider unforeseen impacts in the public authority domain. It is based on the original contract
Langbein and Felbinger (2006)	Descriptive – implementation	–	Review of how the program was implemented usually in terms of time and budget	Private (contractor) It is a partial evaluation related to the project management part. PPPs impact goes beyond the project. It also includes service delivery which is the original intent for using PPPs
	Descriptive output	-	What were the outputs of the program	User, community The outputs refer to mostly quantitative indicators in terms of the benefits. Perspectives about non-quantitative impacts are also critical in PPPs
	Causal implementation	–	Is the implementation working, why? Where is it not working	Private, user A PPP can be implemented in accordance with what was agreed upon, but the impacts to each stakeholder are not addressed
	Causal – output		Is the problem been solved? what else needs to be done?	Private, user The PPP is just the model of procurement, not the solution in terms of provision of public service
Owen and Alkin (2007)	Objective-based evaluation		Judge if the objectives of the program were achieved	Private, user Achieving objectives generate impacts on each stakeholder. It is a different measurement

	Needs-based evaluation	Judge if the program fulfilled the original needs	User	The needs can be related to impacts, but it is a narrow perspective focusing only on the users, and not on the other stakeholders
	Goal-free evaluation	Judge if the program is doing something positive, considering foreseen and unforeseen outcomes.	User, community	The judgement of what is positive and negative is done only by one stakeholder; the evaluator
	Process-outcome evaluation	How was the program implemented? (time, budget, etc.)	Private, user	It is a narrow perspective focusing on the private entity or the contractor. It does not take account of other stakeholders
	Realistic evaluation	Judge what the program is doing correctly, in what context, when and where	User	Multiple stakeholders can have different views about what is correct and what is not correct. It is narrowed to one perspective; the evaluators
	Performance audit	Measure efficiency and effectiveness in financial and non-financial terms	Private	Related to implementation, it is an indicator to control the contractor
State Government of Victoria (2008)	Gateway process	Project review to check business case validity and capture of lessons learned	Public, user	Very complete, but do not explicitly assess all the stakeholders
Ministerio de Desarrollo Social (2017)	Ex-post evaluation: short term and long-term	Essentially a project performance audit	Public, user	It does not consider lessons learned, and has a single stakeholder perspective
International agencies' approaches (section 3.6)				
World Bank (2016)	Beneficiary assessment	Understands the perspective of the beneficiary from shared experience and from observation	User	It is focused just on understanding one stakeholder, with no systematic procedure.
	Social impact assessment	Analysing, monitoring and managing the intended and unintended social consequences of interventions	User, community	It does not consider all the project management part with the public authority and private contractor
World Bank (2012)	Environmental assessment	Provide protection to the environment for sustainable development	Public, community	The environmental impacts can be on the interest of one or more stakeholders. The metrics can be useful, but the aim of the assessment is different
Zimring and Reizenstein (1980)	Post-occupancy evaluation	For any facility, this evaluation is intended for solving existing problems and also for lessons learned at the design level	Private, user	It has a focus on the design and utilisation of a facility, but not in the process or the accountability of other stakeholders

			and the strategic level		
United Nations (Lund-Thomsen, 2007)	Qualitative impact evaluation	PPP	Assess PPP in terms of effectiveness, relevance, efficiency, impact, and sustainability	User	It is a single perspective approach for the UN internal decisions
Jackson & Kassam, 1998	Participatory evaluation		In-depth evaluation employing mainly qualitative data from the users, beneficiaries, or community	User	Focuses on lessons learned, extracting one specific perspective from a single stakeholder

[Source: Author's]

The summary table shows the comprehensiveness of the review, it allows a broad view of the different bodies of knowledge and emphasises the stakeholder inclusion as the main variable of analysis (represented in the second last column). The rest of the chapter reviews in details each of the four areas of knowledge (represented in Table 3-1), arguing why the approaches could be suitable/unsuitable for conducting a PPP ex-post impact evaluation. Additionally, the review also discusses how some approaches could contribute to develop an evaluation method for PPP projects.

3.3. Research on PPP evaluation and monitoring

The first body of knowledge described in Table 3-1 is the research within the PPP context, specifically about evaluation and monitoring. Overall, PPPs have several evaluation and decision milestones over their lifecycle. According to the European PPP Expertise Centre (2011a), the most common are:

- Investment assessment: This first milestone in the PPP project lifecycle in when the idea and the purpose of a project are accepted according to the government's objectives or the public strategy. This approval includes different levels of details and does not necessarily mean it is financially possible.
- Assessment of PPP option: From the multiple procurement options, this assessment should demonstrate that the PPP model is the best fit for the project. Usually, the PPP option is selected if it provides better value for money including the value of financing.
- Selection of procurement method and bid evaluation criteria: If the project was selected as a PPP, then further arrangements need to be set up. This includes the type of PPP option, the bidding criteria, the project team, etc. This milestone allows the project to be tendered.
- Evaluation of tenders and PPP contract award: This assessment has the goal of selecting the consortium that will develop the PPP project in the agreed form.
- Monitoring of project delivery and service output: This evaluation tracks the operational and financial indicators that were established in the PPP contract. It is a constant process that goes from the construction to the operational stage.
- Ex-post evaluation after the operation stage has finished: It should include the assessment of key criteria and appoint a specific team to conduct the evaluation.

The assessments can be divided into approvals and reviews. The formers are not considered relevant to include in the literature review because they only aim to approve the project stages for continuing its lifecycle. Here, the review assessments are the monitoring and the ex-post evaluation. Employing Jingfeng, Yajun Zeng, Skibniewski, and Qiming (2009) map of PPP research, here, three approaches for evaluating PPP are reviewed: the Value for Money concept, the PPP evaluation guidelines, and PPP performance measurement.

3.3.1. Value for Money concept

Value for Money (VfM) is considered the main evaluation approach for PPP projects. The concept is used in many disciplines usually referring to the efficient use of resources. In a PPP context, VfM can be defined as “the optimum combination of whole life cost and quality to meet the user's requirement” (J. Liu, Love, Smith, Regan, & Palaneeswaran, 2015). VfM is applied to PPP projects in many countries like the UK, who has developed a specific framework to measure it. Although VfM is used empirically, there are still concerns about the subjectivity of some criteria (Bauld & McGuinness, 2006).

The Value for Money assessment differs by country and there are several ways to measure it. However, the most common way is the comparison between the selected private bid and a Public Sector Comparator (PSC). The PSC estimates the cost of having the same project built and operated in a traditional way without a PPP. The allocation of risk, the tendering process cost, the future service payments, among other components are considered to provide an estimation of how much ‘less’ money would cost the PPP for the taxpayers (Grimsey & Lewis, 2005). However, research has suggested that the availability of public funding can influence the result of the PSC. If there is a certainty that the government cannot afford the project with public funding, then the PSC probably will be worse (more expensive) than the PPP option, in order to approve the PPP project (Heald, 2003).

Burger and Hawkesworth (2011) studied the ex-post assessment of Value for Money in OECD countries by surveying PPP unit directors and public authorities. 11 countries (Including Chile and Australia) answered that they do conduct ex-post evaluations of Value for Money. However, it was confirmed that those reports were not available; only ex-ante Value for Money assessments were conducted. Consequently, it becomes difficult to rely on secondary information regarding Value for Money assessment.

Value for Money could help the development of an evaluation method, assuming that is an important criterion for PPP projects. However, it is not clear how stakeholders could be involved, nor how to evaluate the impact of the project.

3.3.2. PPP ex-post evaluation guidelines

The European PPP Expertise Centre (2011a) from the European Investment Bank (EIB) is an organisation that strengthens the ability of the public sector to engage in PPP transactions by helping the members to share their experience and expertise. In their PPP guidelines, they include the application of a PPP ex-post assessment as part of the process of procuring and delivering a PPP, validating the relevance of the research topic and the development of the evaluation method. However, the approach is presented as a general guideline with no clear instruction on how to assess a PPP in practice.

3.3.3. PPP performance

PPP performance has been studied by many researchers in construction management (Hardcastle et al., 2005; Jingfeng et al., 2009; J. Liu, Love, Davis, Smith, & Regan, 2015; J. Liu, Love, Smith, et al., 2015; Mladenovic, Vajdic, Wündsch, & Temeljotov-Salaj, 2013; Peter Raisbeck et al., 2010; Xiong, Yuan, Li, & Skibniewski, 2015). The goals of the studies can be characterised as (1) comparison between the KPIs of PPPs and Traditional Procurement; (2) whole-life assessment of a PPP from a project management perspective, (3) success factors for PPPs; and (4) mathematical models to manage stakeholder's satisfaction.

PPP performance differs from what this research understands as PPP evaluation because the former does not include a clear decision-making stage. It is about how to measure performance using indicators, rather than how to judge those measurements. The PPP performance framework

developed by J. Liu, Love, Davis, et al. (2015) has a detailed description of the stages and processes of PPP projects. The framework contains the different KPIs to measure performance at every stage and the subsequent actions to take in each of the stages. However, it aims to improve the PPP performance of an ongoing project, without explicitly addressing the impacts of the project on the stakeholders.

PPP critical success factors is a subtopic of PPP performance. Hardcastle et al. (2005) evaluate the importance of 18 CSF for PPP/PFI in UK construction projects grouping them in 5 categories: effective procurement, project implementability, government guarantee, favourable economic conditions and available financial market. J. Liu, Love, Smith, et al. (2015) developed a framework to describe the difference CSF by the stages of the projects. The main stages are planning, procurement, and partnership. In each stage, success factors are described which should enhance the successfulness of an infrastructure PPP projects. Even though PPP critical success factors can help indirectly in the evaluation of PPPs, a factor is defined as a key area for improvement (Lim & Mohamed, 1999), and not criteria for measuring that improvement.

3.4. Management-based evaluation and monitoring approaches

The second body of knowledge (area of knowledge) presented in Table 3-1 (section 3.2) is the management-related disciplines, from which four approaches are reviewed: Project Success, stakeholder management, Performance Measurement System, and project audit. As stated in the introduction of the chapter, the aim is to search for possible evaluation approaches that could evaluate a PPP from the perspective of multiple stakeholders.

3.4.1. Project Success concept

To assess the impact of a PPP on stakeholders, Project Success concept from the project management discipline can be employed. A PPP is a project with a specific procurement strategy; this means that the concept of Project Success is applicable and useful to the research aim. Existing literature has focused on (1) developing frameworks to analyse Project Success; (2) success factors, which are key areas that can improve the possibility of having a successful project; and (3) success criteria, which are the measurements that can be assessed to determine if a project is successful or not. However, existing literature in this area lacks detail (R. A. Müller & Jugdev, 2006). There is no accepted method to assess the success of a project and the existing research is all within the academic domain. Moreover, the definition of Project Success is difficult to apply because it considers the judgement of different stakeholders (Davis, 2013).

The definition of Project Success has been largely discussed and yet there is no consensus on its definition. Originally, it referred to a project that finished on time, on budget and on specifications (R. A. Müller & Jugdev, 2006). However, this is now more appropriately called project management success. Currently, the literature argues about the relevant role of stakeholders in the definition of Project Success (Davis, 2013) as a project is successful depending on the judgement and expectations of their stakeholders. However, a formal definition of Project Success that considers the stakeholders' viewpoints remains unclear.

Measuring Project Success is even more difficult. de Wit (1988) concluded after analysing different methods for measuring Project Success that the objectivity of the measurement is an illusion, mainly because the perception of success can vary within the different stakeholders and also it can change over time. He also states that any criteria-weighting for Project Success is incorrect, as it will only give an artificial answer. What is implicit is de Wit's conclusion is that the measurement of Project Success must be subjective and based on perception, which is how it is currently in practice; any stating of success is usually a judgement. Shenhar, Dvir, Levy, and Maltz (2001) also suggest that literature is divided in term of the definition of Project Success and currently there is still no accepted framework for assessing Project Success.

Critical success factors and success criteria are important components in the Project Success research. Lim and Mohamed (1999) make a clear distinction between criteria and factor. A criterion is “a principle or standard by which anything is or can be judged”, and a factor is “any circumstance, fact, or influence which contributes to a result”. For example, a success factor can be the ‘team expertise’ (necessary to achieve a successful project) while a success criterion can be the ‘end-user satisfaction’ (unsatisfied end users can mean the project was a failure). The search for success factors and criteria in context-specific projects (including PPPs) is an important research branch in this area (Ika, 2009).

3.4.2. Stakeholder management

Even though it is not a formal evaluation approach, the review also includes the stakeholder management concept, which is closely related to the multiple stakeholder perspective attribute. A stakeholder is defined as an individual or group of people that can affect or be affected by the processes, contents or outcomes of a project (Eskerod, Huemann, & Savage, 2015). This definition refers to the project manager, the project team, the employees, the shareholders, the public authorities, the clients, the consumers, the end-users, the community, etc.

Methods to identify project stakeholders have been developed. For example, Barclay and Osei-Bryson (2010) states a way to identify the stakeholders using 8 steps: (1) identify active roles in the project, (2) impacted by the project, (3) involved in project management processes, (4) impacted by the product, (5) reduce to common set, (6) prioritise by influence and power, (7) categorise, and (8) review. Another overall framework for stakeholder management is present by Yang and Shen (2015) in which the identification and assessment of stakeholders are described as in a general way: (1) define the stakeholder management objective, (2) identify the full list of stakeholders, (3) collect stakeholders information, (4) assess power, urgency and proximity; (5) prioritise stakeholders, (6) analyse interrelationships, and (7) assess behaviour. Although the methods for stakeholder identification are based on academic research, they are considered a challenge in real-world projects for their lack of applicability. Project managers have commented that current frameworks for stakeholder management are just conceptual ideas with no practical instructions (Jepsen & Eskerod, 2009).

One important attribute of stakeholders is the relationship between their expectations and the overall success of the project. The stakeholders’ expectations are what every stakeholder involved in a project expects from the outcomes of the project or from the processes during the lifecycle of it. Stakeholders’ expectations can also be related to stakeholder satisfaction; Xiong et al. (2015) define stakeholder satisfaction as the “consumer’s response to the evaluation of the perceived discrepancy between prior expectations and the actual performance of the product”.

A possible way to measure the expectations and the satisfaction of stakeholders is by comparing Performance Objective Indicators (POI), and Key Performance Indicators (KPI) (Xiong et al., 2015). The satisfaction of each of the stakeholders can be measured as the difference between the expected level of POI and the resulting level of POI measured in a KPI. This kind of measurement uses quantitative metrics for some criteria (time, budget, profit, etc.), and Likert scale metric for the ones that are related to perception. This approach does not consider the qualitative information regarding the expectations, and it cannot deal with unforeseen events that are not considered in the first place. Nevertheless, it is a clear framework for understanding and managing stakeholder expectations.

In other disciplines, apart from project management, the concept of stakeholders’ expectations is also defined and analysed. In business management discipline the customer satisfaction concept is very close to the stakeholder expectation concept. SERVQUAL is one of the most used models to determine and measure consumer satisfaction for service quality. It was developed by Parasuraman, Zeithaml, and Berry (1988), and it is based on the gap that exists between the expectation of the customer about

the service (how it would/should be the service) and the perception of the service quality (how was the service). This comparison is measured in five dimensions, named: (1) tangibles, (2) reliability, (3) responsiveness, (4) assurance, and (5) empathy. This approach could be applicable to the project management field.

Stakeholders' expectations from the project management perspective could be applicable to the evaluation of PPP projects. However, there is not enough theory for establishing an ex-post evaluation method based on stakeholder expectations. Specific tools, from other disciplines, like the SERVQUAL method, can be a good support for developing practical tools and procedures in the project management field. The topic of stakeholders' expectation is linked with the Project Success concept, which is linked to PPP success. The measurement of stakeholders' expectations should be a core component for any attempt to evaluate the impact of a PPP project over multiple stakeholders.

3.4.3. Performance Measurement Systems

Performance Measurement Systems (PMS) such as the Balanced Scorecard (Kaplan & Norton, 2007) and the Performance Prism (Neely, Adams, & Crowe, 2001) could also be an approach to perform a PPP ex-post evaluation. They have the aim of measuring performance using specific metrics and indicators. However, they are not capable of performing an evaluation which consists of measuring, capturing data, and then deciding. Conversely, Performance Measurement Systems are frameworks to capture quantitative data, but with no decision process included. Nevertheless, an ex-post impact evaluation could use the PMS structure at the measurement stage.

Performance measurement is defined as the process of quantifying action (Neely, Gregory, & Platts, 1995). It has the aim to control and monitor the performance of any process or project with indicators called key Performance Indicators (KPI). PPPs monitor their performance using (partially or completely) a Performance Measurement System, usually at the construction and operation stage. Key Performance Indicators are included in the contracts as the way to measure compliance to the PPP objectives.

A Performance Measurement System (PMS) is a number of individual performance measures that are grouped in a certain way to monitor the ongoing process within an overall strategy (Neely et al., 1995). The first performance measurement frameworks were criticised due to their disconnection between their different measurements. This means that the KPIs were employed individually, but the group of KPIs had no meaning, goal or overall objective. This was partially solved by Fitzgerald, Brignall, Silvestro, Voss, and Robert (1991) making the difference between measuring results (the goal) and measuring factors that determine results (necessary to achieve the goal). Later, Kaplan and Norton also integrated this attribute using a cause/effect relationship between the KPIs, being one of the key features of the Balanced Scorecard (Neely et al., 2000). When employing PMSs it must be noticed that measuring action also generates action, creating a virtuous or vicious cycle. Additionally, the way of measuring action influence how the action is performed, affecting the desired goal (Neely et al., 1995).

The Balanced Scorecard and the Performance Prism are reviewed in detail, confirming the difference between a PMS and an evaluation; but also providing insights on how a PMS can be useful for developing an evaluation.

The Balance Scorecard

The Balance Scorecard is one of the most popular management tools and probably the most popular Performance Measurement System that has been implemented by thousands of firms around the world (Balanced Scorecard Institute, 2016). The value of this system is by supplementing the traditional finance measurements with three other perspectives: customer, internal processes, and

learning and growth. Each perspective is connected to each other with a cause-effect relation for achieving a financial goal. The monitoring of the different measures that have a financial effect on the three other perspectives will help to achieve the long-term goals that the business is pursuing (Kaplan & Norton, 2007).

The implementation of a Balanced Scorecard is presented by Kaplan and Norton (2007) as a cycle within a timeframe of five years. The four steps of the cycle are: translating the vision, communicating and linking, business and planning, and feedback and learning. By using the implementation as a continuum process, the scorecard provides a better link between the company's long-term strategy and their short-term actions.

The Performance Prism

Neely et al. (2001) developed the PMS Performance Prism which is considered a new generation of measurement frameworks for performance measurement. It has a stakeholder orientation and differs from the Balance Scorecard specifically in this matter, making a broader view of the stakeholders. The Balanced Scorecard considers only shareholders and customers as the relevant stakeholders, whereas the Performance Prism considers shareholders, customers, authorities, employees, community, suppliers, alliance partners, and/or intermediaries.

One of the paradigms that the Performance Prism breaks is the relationship between the strategy and the stakeholders. Neely et al. (2001) recognise that the strategy does not influence the stakeholder, but the other way around. The starting point should be the stakeholders' satisfaction triggering the generation of the strategy, processes and capabilities.

The structure of the Performance Prism is based on five facets. The first two facets are the satisfaction and contribution of the stakeholders, which are the top and bottom of the prism, respectively. The three other facets are the strategies, processes and capabilities of the organisation that need to be measured and are conceptually the three sides of the prism. The aim of the structure is that the stakeholders are the start and the end of the system and that the organisation should analyse three aspects from a macro to a micro view.

The review shows that the structure of the Performance Prism is a multiple stakeholder approach and could be employed for developing a PPP ex-post impact evaluation method. However, the task of monitoring and improving action is different from evaluating and judging the impact of a past action.

3.4.4. Project audit

Project Audit is another body of knowledge to address a PPP ex-post evaluation. Originally an audit is conceptualised as a financial tool to assess compliance, rather than to assess the quality and the impacts of the outcomes. Nevertheless, there are approaches such as Nalewaik and Mills (2017) that addresses the question 'was the right thing done?' In contrast to the traditional 'was the thing done right?'. The problem with this approach is that it is mainly based on the auditor's single perspective, complemented with inputs from the rest of the stakeholders. It is not considered a multiple stakeholder perspective approach.

3.5. Program evaluation (public policy) approaches

The third body of knowledge summarised in Table 3-1 (section 3.2) is the program evaluation and public policy-related research. A PPP is a model of procurement for building infrastructure and providing a public service. In fact, PPP contracts include the operational stage in which a public service will be provided by a private entity. Therefore, program evaluation could be used as an approach for evaluating PPPs using, what is called, an impact evaluation (Langbein & Felbinger, 2006; Owen & Alkin, 2007; Quade & Carter, 1989). The problem with this approach is that all the different types of impact

evaluation are mostly performed by a team of experts using their unique perspective without a major focus on stakeholders. Thus, the impact of the project on each stakeholder (the aim of this research) is difficult to achieve with the current program evaluation approaches.

From the public policy analysis and evaluation discipline, Langbein and Felbinger (2006, p. 3) define program evaluation as *“The application of empirical social science research methods to the process of judging the effectiveness of public policies, programs, or projects”*. Owen and Alkin (2007) explain the logic of evaluation using four processes: (1) establishing criteria of worth, (2) construct standards, (3) measuring performance and comparing with standards, and (4) synthesising and integrating evidence into a judgement of merit or worth. They distinguish the process of judging as a very difficult task in which all the information is condensed into one word, ‘good or bad’. PPP projects can be certainly evaluated using program evaluation as they provide a public service and their aim is to increase public welfare. This is important to notice because any attempt to evaluate a PPP should consider partially the program evaluation perspective.

Within a set of different kinds of program evaluations, the impact evaluation can be compared to an ex-post assessment, as it pursues the same goal of measuring the effect of a project over a situation, organisation, society, stakeholder, etc. There are different ways of classifying impact evaluations. For example, Owen and Alkin (2007) identify impact evaluations as: (1) objective-based evaluation, made by judging if the objectives were achieved; (2) needs-based evaluation, made by judging if the program meets participants’ needs which might differ from the objectives; (3) goal-free evaluation, that reviews the stated goals, as well as unforeseen outcomes; (4) process-outcome studies, which evaluate how the implementation was performed, apart from the final outcomes; (5) realistic evaluation, that states in what context and to whom the program is really effective; and (6) performance audit, which measures efficiency and effectiveness from a financial and a non-financial perspective. Other approaches such as Quade and Carter (1989) or Langbein and Felbinger (2006) classify impact evaluation by focusing on the results, on the method, or on the sample. The general overview of the different approaches of impact evaluations does not consider explicitly the perception of the stakeholders about a certain project. This can be either a gap in knowledge (a new way of measuring impact) or an approach that is not considered relevant from the perspective of program evaluation.

Within program evaluation, two governmental evaluation methods are reviewed below: The Gateway Review Process from the Australian Government, and ex-post assessment of the Social Development Ministry of Chile. Both are fit-for-purpose tools that are employed for evaluating public investments. The Chilean and Australian cases are employed for developing the PPP ex-post impact evaluation, and therefore their current evaluation methods could also be useful for obtaining applicable contextual insights.

The Gateway Review Process (Australian Government)

The Gateway Review Process from the Victorian Government of Australia has a standardised procedure for evaluating and monitoring public investments; examining projects and programs at key decision points. It has six gates (stages): strategic assessment, business case, readiness for market, tender decision, readiness for service, and benefit evaluation (Department of Treasury and Finance, 2018). It is applied by Partnership Victoria, which is the PPP unit for the state government, part of the Victorian Department of Treasury and Finance.

The aim of the last gate of the process is to confirm that the business case was justified, and the lessons learned were captured. This procedure can be used at different stages of the project; shortly after starting a long-term contract, at the mid-stage of the contract, and/or at the end of the contract. It is divided into sections: operational phase, business case, ongoing improvement (Value for Money), performance and innovation, organisational learning and maturity targets. However, as explained by

the auditor general, the gateway is not an audit, nor a technical review, it is a review to identify where the project team needs to put the efforts for achieving a successful project. It differs from an explicit ex-post impact evaluation, but the topics that are covered could be incorporated.

Social Development Ministry's ex-post evaluation procedure (Chilean Government)

The Social Development Ministry oversees the evaluation, approval, and subsequent monitoring of every public investment from a social perspective. It has its own procedure for performing ex-post evaluations. Two kinds of evaluations are documented: short-term ex-post evaluation (simplified), and medium-term ex-post evaluation (complete). The simplified version contains key information relevant from a project management perspective: cost, time, modifications, contracts compliance, expenses, efficiency indicators, social impact comparison (with ex-ante), among others. The long-term ex-post evaluation includes service level analysis, demand analysis, initial problem identification, level of solution, funding schedules, user profile, problem-solving indicator, user satisfaction, cost-benefit analysis (Ministerio de Desarrollo Social, 2017). Employing program evaluation taxonomy, this ex-post evaluation procedure is considered a performance audit in which only the public authority and the user/beneficiary perspectives are included.

Program evaluation provides a robust body of knowledge for understanding an ex-post impact evaluation. The theory describes many alternatives with their specific objectives, and in practice, two governmental procedures were reviewed. However, program evaluation does not provide an explicit impact evaluation method that considers the complexity of the many stakeholder perspectives that surround a PPP project.

3.6. International agencies' approaches

The final category summarised in Table 3-1 (section 3.2) describes the evaluation approaches from international organisations. They are constantly monitoring their social projects that have the purpose of increasing welfare and life quality in lower-income countries.

An article from the United Nations Research Institute for Social Development (UNRISD) programme (Lund-Thomsen, 2007) explains the method for assessing the impact of PPPs in the global south. The report assesses a PPP in terms of effectiveness, relevance, efficiency, impact, and sustainability. Even though each criterion is useful for decision-making purposed, it does not inform the impact of a PPP over the multiple stakeholders. This kind of evaluation can be classified as a specific case of a realistic evaluation (Owen & Alkin, 2007) with fixed evaluation criteria.

Participatory evaluation (Jackson & Kassam, 1998) is another approach that focuses on the capacity of the users/beneficiaries, and the community to address the evaluation of social programmes; capturing in-depth-knowledge about its performance. It employs qualitative and some quantitative methods to generate data that can be used for improving future social interventions. The multiple stakeholder perspective is not evidently present in this approach as it is mainly a feedback process rather than an impact evaluation.

The World Bank employs different methods for evaluating projects, including environmental evaluation, beneficiary evaluation, social evaluation, among others. They follow the same pattern of using a specific set of criteria, without considering the multiple views within a project.

3.7. The gap in knowledge

Chapter 2 reviewed the broad topic of PPPs including their characteristics, their problems, and their current applications in Chile and Australia. The PPP questioning in terms of efficiency, effectiveness, and probity supports the argument that PPPs should be assessed with a proper evaluation method. The selected approach for addressing these issues is an ex-post evaluation, which should (in principle)

confirm whether a PPP is facing all the claimed problems. The research's point of departure is that PPPs should be better assessed, that the different perspectives over PPPs make their assessment even harder, and that there should be a proper tool to conduct this assessment.

Chapter 3 focused on the approaches that could ex-post evaluate a PPP project. The result is that there is no explicit method to do so, which in turn, validates the need for developing an evaluation approach which is considered a gap in knowledge. In this chapter, 30 evaluation approaches were reviewed from four different disciplines and bodies of knowledge (summarised in section 3.2, Table 3-1): research on PPP evaluation and monitoring, management-based evaluation and monitoring approaches, program evaluation (public policy) approaches, and international agencies' approaches. The main variable of analysis was their capability to conduct a PPP ex-post impact evaluation from multiple stakeholder perspectives. Therefore, this attribute was discussed in every reviewed approach (or area of knowledge).

The conclusion of the review is that there is no tailored (fit-for-purpose) method that can answer the research question. Only a subset of the approaches is explicitly ex-post and/or impact evaluations. An even smaller portion has the multiplicity of stakeholders incorporated. From the research on PPP evaluation and monitoring (section 3.3) the whole of life PPP assessment considers multiple stakeholders, but they are not intended for evaluating impacts; their main role is for monitoring and improving project management performance. From the management-based approaches (section 3.4) the Project Success concept (together with the stakeholder management) understands the multiplicity of stakeholders, but it is only conceptually developed and there is no application for an impact evaluation. Performance Prism also considers multiple stakeholders, but it is a PMS which differs from an evaluation in their decision-making process. Finally, project audit is formally an ex-post evaluation, but it considers the multiple stakeholders as just an input to the main evaluation process.

The gap in knowledge is confirmed as there is no method for explicitly evaluating ex-post the impact of a PPP project over multiple stakeholders. However, there are many insights that can be captured for informing a ex-post PPP impact evaluation in terms of structure, data presentation, logic, etc. The main ones are (1) the structure of the PMS using KPIs, (2) the understanding of a successful project from multiple perspectives, (3) the Gateway procedure and structure, (4) the impact evaluation (from public policy) decision-making and judgement understanding, and (5) the Value for Money indicator as measurement of PPP worthiness.

3.8. Summary

Chapters 1 and 2 argued that PPPs have problems, that they are being criticised and that they should be properly assessed. The research's point of departure is that the impacts of PPPs should be ex-post assessed. Chapter 3 reviews the existing literature and confirms the gap in knowledge by finding no method capable of conducting a PPP ex-post impact evaluation from the perspective of multiple stakeholders. A total of 30 evaluation approaches were gathered from four disciplines (bodies of knowledge): research on PPP evaluation and monitoring, management-based evaluation and monitoring approaches, program evaluation (public policy) approaches, and international agencies' approaches. From the review, four approaches are highlighted as they could be a useful contribution for developing a PPP ex-post impact evaluation: (1) the PPP whole life assessment, (2) the Project Success concept, (3) the PMS Performance Prism, and (4) the project audit body of knowledge. Insights from these approaches can be useful in terms of structure, logic, presentation, etc.

The four disciplines (bodies of knowledge) were reviewed in detail.

From the PPP research, three topics were discussed: The Value for Money indicator, the PPP guidelines, and PPP performance. (1) The Value for Money concept is the common indicator for

approving a PPP project; it compares the PPP with a traditionally procured version of the same project, measuring which one provides more value for the same whole of life cost. However, it is not a proper evaluation approach and it does not explicitly involve the project stakeholders. (2) PPP guidelines (good practices) from the PPP expertise centre (EPEC) recommend conducting a PPP ex-post evaluation, which supports the relevance of the research topic. However, the guidelines for the ex-post evaluation are very broad with no clear instruction. Finally, (4) PPP performance research topic is mostly related to project management performance that has a strong focus in project cost, time and quality; rather than on stakeholders.

Within the management-based evaluations, four approaches are reviewed. (1) the Project Success concept involves multiple stakeholders; however, it is only conceptual with rare application in practice. (2) Stakeholder management relates to stakeholder expectations, which is a relevant topic when evaluating a PPP from a multiple stakeholder perspective. Even though it is not an evaluation approach, it can contribute to the development of one. (3) Performance Measurement Systems allow a project to be measured and monitored with Key Performance Indicators (KPIs). However, it differs from an evaluation, because it does not involve the decision-making process. (4) Project audit is a body of knowledge that traditionally evaluates from a financial perspective and ex-post. Even though there are more holistic approaches (beyond financial), the 'auditor' is considered the main perspective, which manages the rest of the stakeholder perspectives.

Program evaluation from public policy is a strong theoretical support for PPP evaluation, explicitly describing an impact evaluation. 17 approaches are under this discipline. However, they do not involve multiple stakeholders and they focus mostly on user/beneficiary stakeholders. This same characteristic can be seen in governmental approaches such as the Gateway review from Australia and the medium-term ex-post evaluation from Chile.

The World Bank and other international organisations have evaluation methods to evaluate PPPs, but they are mostly applications and cases from the program evaluation theory.

The literature review (chapters 2 and 3) argues that PPPs are facing some presumable problems, that these issues need to be assessed properly, that the best assessment approach for this case is an ex-post evaluation, and that there is no approach for ex-post evaluating a PPP. Therefore, the gap in knowledge and the opportunity for research is the development of a PPP ex-post impact evaluation method. The next chapter describes the research strategy for answering the research question.

CHAPTER 4

RESEARCH METHODOLOGY

4.0. Introduction

This chapter presents the research design and the methods employed for conducting the research project. The Critical Realism (CR) philosophical paradigm is adopted which supports a Design Science methodological approach.

The research strategy is an iterative design process consisting of four stages; one for development, and three for testing and refinement. For the development stage, data is collected from experts and practitioners through interviews, mainly for feedback purposes after presenting them the conceptual understanding of the evaluation method. In the three testing and refining stages, the evaluation method is applied to three PPP projects (one per stage), which are called 'test cases'. For each test case, the data is collected from documents and stakeholder workshops that took place following the procedure of the proposed evaluation method.

The structure of this chapter is as follow. Section 4.1 reviews the Critical Realism philosophical assumption and section 4.2 the Design Science concept. The research strategy is explained in section 4.3. The data collection and analysis are detailed in section 4.4. Finally, the validation of the research project as a suitable Design Science research is presented in section 4.5.

4.1. Adopted research paradigm

A research paradigm is defined as a set of beliefs on how the surrounding reality can be understood. Usually, it is defined in ontological (what is reality, single or multiple realities), epistemological (how to know reality), and methodological (how to get knowledge) terms (Guba & Lincoln, 1994). Two paradigms are commonly employed: Positivism and Constructivism. Positivism believes in a single reality that can be measured. Hence, its methodology is mostly for confirming cause-effects relationships through experimental and survey research. Constructivism believes in multiple realities that people can create (ontology) and its epistemology is through interpretation. Methodologically, it employs ethnography, grounded theory, discourse analysis, etc. A third research paradigm is Critical Realism, which is considered a bridge between Positivism and Constructivism.

This research adopted a Critical Realism (CR) paradigm as its believed reality can be presented in three layers called the domain of the real, the domain of the actual, and the domain of the empirical (Carlsson, 2006). In these layers, mechanisms (present only in real), events (present in real and actual), and experiences (present in all layers) reside. CR argues that (generally) 'mechanisms' in real-world generate 'events' which subsequently generate 'experiences'. In this structure, objects can solve a problem, changing the empirical domain in which experiences exist (Hoddy, 2019). This is consistent with a Design Science methodology that generated knowledge through the creation of objects called artefacts. Moreover, the CR research aim is to critique, transform, and emancipate reality (Badewi, 2016, p. 83). The reasons for adopting CR paradigm are (1) the creation of an artefact (ex-post impact evaluation method) that deals with social world perceptions, and (2) the research aim of solving a real problem (unknown overall impact of PPPs).

4.2. Design Science approach review

The Design Science approach is based on Simon (1969) 'the science of the artificial'. Simon argues that science can be separated in two main forms: understand and describe nature's behaviour, and change nature through artefacts that can solve real problems in the world. Physics and biology, among others, are disciplines that are usually called 'natural sciences'; the reason for that is that their main concern is about 'truth' and 'necessity' (Hill, 2009). In contrast, artificial sciences concern about 'usefulness' and 'contingency'. Many design questions can be answered from disciplines such as engineering,

medicine, law, architecture, and business. All of them must have a common notion of what constitutes an 'artefact'; which needs to be designed to solve a problem in the world. According to Simon, the behavioural sciences (natural sciences) conquered the way of doing research since after the Second World War. He explains that within the universities, the engineering faculties had transformed into physics faculties, the medicine faculties into biology faculties, and so on. According to him, there is a need for changing the trend and explore this alternative way of understanding research (Simon, 1996, p. 112). Design Science has developed from Simon's seminal work, but still is considered a new approach, at least within some disciplines (Hill, 2009, p. 18). Design Science creates prescriptive knowledge, opposite to the descriptive knowledge (Gregor & Hevner, 2013). Prescriptive knowledge shows how to do something, while descriptive knowledge explains how things are naturally done. To evaluate a PPP in a different way changes how things are done.

The term 'artefact' refers to a creation that is made by humans and therefore it is considered artificial; an artefact is something that does not occur naturally (Simon, 1996). The artefact should provide a solution to an existing problem in the real world, but also it should solve the problem in an efficient manner providing utility to the user (Hevner & Chatterjee, 2010, p. 2). March and Smith (1995) defined artefacts as:

1. Constructs (vocabulary and symbols);
2. Models (abstractions and representations);
3. Methods (algorithms and practices); and
4. Instantiations (implemented and prototype systems).

Hevner et al. (2004), specifically, for IT context added a fifth option:

5. Better design theories.

The previous definition is based on, and consistent with, the original definition of Simon regarding what constitutes an artefact. According to Simon (1996), an artefact has several characteristics that differentiate it from a natural phenomenon. An artefact is synthesised by a human being, it may imitate the appearance of something natural, and can be characterised in terms of function, goal and adaptation. Example of artefacts (even though not designed with a Design Science approach) exist in all its definitions; Porter's 5 forces model is a theoretical model to understand strategy in the business discipline, C++ language is an artefact for programming software in a computer, simplex is a method to solve a linear programming problem, and Microsoft Word is an instantiation (software) for writing text. An evaluation method is an artefact to be designed in this research employing Design Science approach.

Even though the big development for Design Science Research (DSR) has been done in the Information System (IS) research field; DSR has been employed in other fields such as engineering and management (Hevner & Chatterjee, 2010, p. 196). From March and Smith (1995) it is possible to distinguish three insights regarding the use of DSR out of the IT field. First, that technology means 'practical implementation of intelligence', making it possible to define an ex-post impact evaluation method as a technological solution. Second, it states that the discussion about the legitimacy of scientific research between behavioural (natural) science and Design Science supporters is not only present in the IT field, but also between engineering and physical sciences. Third, March and Smith (1995) explained that there is a need for a combination of natural science and artificial science to create a designed artefact in the management field; a cycle for the creation of useful and trustful knowledge for organisations. To design a PPP evaluation, PPPs must also be researched using a descriptive approach.

The Design Science research structure, the way to evaluate Design Science outcomes, and the contribution to knowledge of this methodological approach are reviewed in the following subsections.

4.2.1. Design Science structure

The Information System (IS) discipline has developed knowledge to apply the Design Science Research (DSR) in practice. According to Hevner et al. (2004), the basic process of Design Science is the iteration of a develop/build process and a justify/evaluate process (Figure 4-1). After a number of iterations, the research should have two main outcomes: solve a real problem and contribute to the existing knowledge.

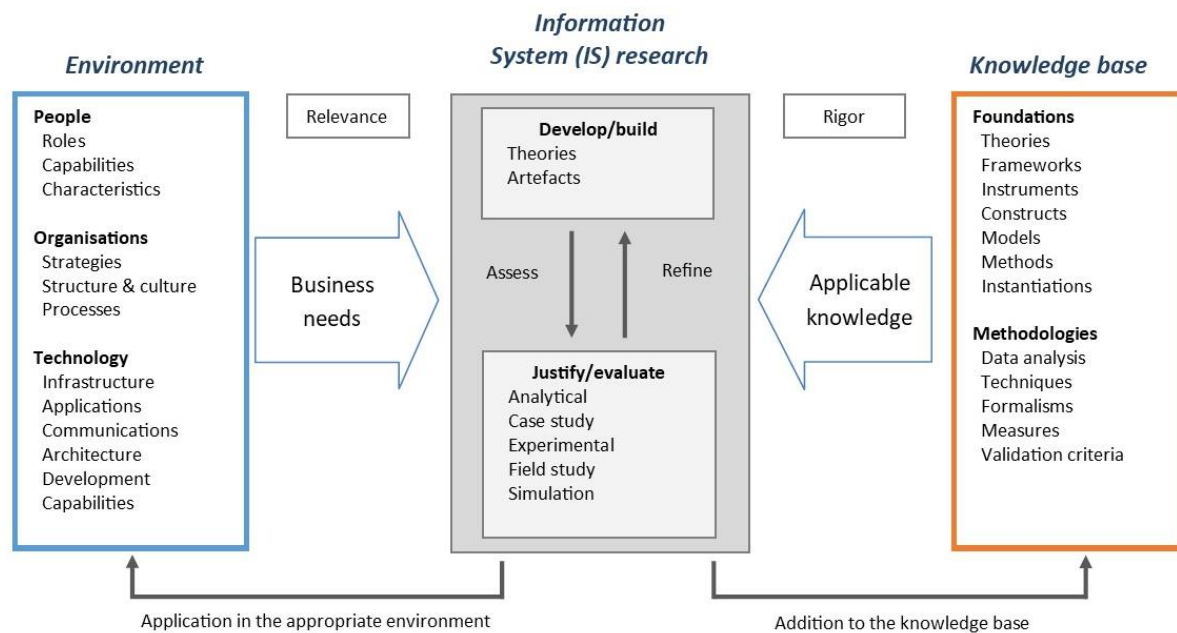


Figure 4-1: Information System Design Science framework

[Source: Adapted from Hevner et al. (2004)]

The 'develop/build' process has the aim of developing an artefact in the form of a construct, model, method or instantiation. The design of the artefact is based (1) on the connection that exists with the environment (people, organisations and technology) as a business need, and (2) on the connection with a knowledge base (existing foundations and methodologies) that can be applied to the new design. The surrounding environment of the designed artefact provides relevance to the research. The knowledge base provides rigour to the designed artefact. This rigour is needed in research and science (different from a pure design process).

The 'justify/evaluate' process gives the artefact the assessment to be refined in an iterative process. After the final iteration the artefact has the validity to be applicable in the environment, according to the business needs, and with sufficient support to be part of the existing knowledge base.

Peffer, Tuunanen, Rothenberger, and Chatterjee (2007) proposed a structured methodology for performing DSR in the IS field. It includes 6 steps: (1) problem identification and motivation; (2) definition of the objectives for a solution, (3) design, and development; (4) demonstration; (5) evaluation; and (6) communication. To start a DSR the first question to answer is what do we know already? From what existing knowledge can we draw? Prescriptive and descriptive knowledge bases are necessary to start the research process. By having a baseline of knowledge it is possible to establish and evaluate the novelty of the new artefact that is going to be designed to answer the research

question (Gregor & Hevner, 2013). A relevant component in the Design Science approach is the evaluation process, described in the next section.

4.2.2. Design Science artefact evaluation

The evaluation of the artefact is important for demonstrating its usefulness and necessity. The aim is to test the artefact in terms of efficiency, utility, and performance (Hevner & Chatterjee, 2010, p. 109). The available methods for evaluating the artefact are presented in Table 4-1. For this research, the PPP ex-post impact evaluation is tested using case study under observational methods. According to Hevner, an artefact is complete when it satisfies the requirements and constraints of the problem that it is intending to solve. The PPP ex-post impact evaluation method is analysed using these 7 steps in section 4.5.

Table 4-1: Design Science evaluation methods

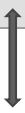
ID	Description
Observational	Case Study: Study artefact in-depth in a business environment
Analytical	Field Study: Monitor use of artefact in multiple projects
	Static Analysis: Examine structure of artefact for static qualities (e.g., complexity)
	Architecture Analysis: Study fit of artefact into technical IS architecture
	Optimization: Demonstrate inherent optimal properties of the artefact or provide optimality bounds on artefact behaviour
	Dynamic Analysis: Study artefact in use for dynamic qualities (e.g., performance)
Experimental	Controlled Experiment: Study artefact in a controlled environment for qualities (e.g., usability)
	Simulation – Execute artefact with artificial data
Testing	Functional (Black Box) Testing: Execute artefact interfaces to discover failures and identify defects
	Structural (White Box) Testing: Perform coverage testing of some metric (e.g., execution paths) in the artefact implementation
Descriptive	Informed Argument: Use information from the knowledge base (e.g., relevant research) to build a convincing argument for the artefact's utility
	Scenarios: Construct detailed scenarios around the artefact to demonstrate its utility

[Source: Based on Hevner et al. (2004)]

4.2.3. Design Science contribution to knowledge

The contribution to knowledge from a DSR depends on the type of artefact that the research process can have as an output. Gregor and Hevner (2013) state that the DSR should have artefacts in different levels, from the instrument that solves a specific problem, to the abstract knowledge that supports the development of the instrument. These theories, known as design theories, are the abstract laws and conceptualisations that allow the development of other future artefacts capable of solving similar problems. Table 4-2 analyses the type of artefact and the contribution.

Table 4-2: Design Science research contribution types

	Contribution Types	Example Artefacts
More abstract, complete, and mature knowledge	Level 3. Well-developed design theory about embedded phenomena	Design theories (mid-range and grand theories)
	Level 2. Nascent design theory—knowledge as operational principles/architecture	Constructs, methods, models, design principles, technological rules.
More specific, limited, and less mature knowledge	Level 1. Situated implementation of artefact	Instantiations (software products or implemented processes)

[Source: Based on Gregor and Hevner (2013)]

The PPP ex-post impact evaluation has a research contribution at level 2 and 1. The theory for PPP evaluation is considered knowledge as an operational principle. And the procedure to evaluate a single

PPP project (main outcome on chapter 8) is classified under level 1 as an implementation of the artefact.

4.3. Research strategy and methods

Based on the Design Science approach (reviewed above), the research strategy is presented differently from a traditional behavioural science research that can include hypothesis and methods to confirm a cause-effect relationship; or the methods to explore a phenomenon, describing its behaviour. The research question of this study is:

How can ex-post impact evaluation of Public-Private Partnership projects be conducted employing the perspective of its multiple stakeholders?

The strategy for designing the ex-post impact evaluation method consists of several iterations between develop/build and justify/evaluate processes. Figure 4-2 presents the application of the Design Science diagram (presented in Figure 4-1) to this research. It shows the research process contextualised to the development of the PPP ex-post impact evaluation method. The environment, the knowledge base, and the design cycle now include the context-specific concepts for this research project. For example, the ‘business need’ presented by Hevner is now replaced with ‘society need’ as a more appropriate way to support the relevance of the research problem.

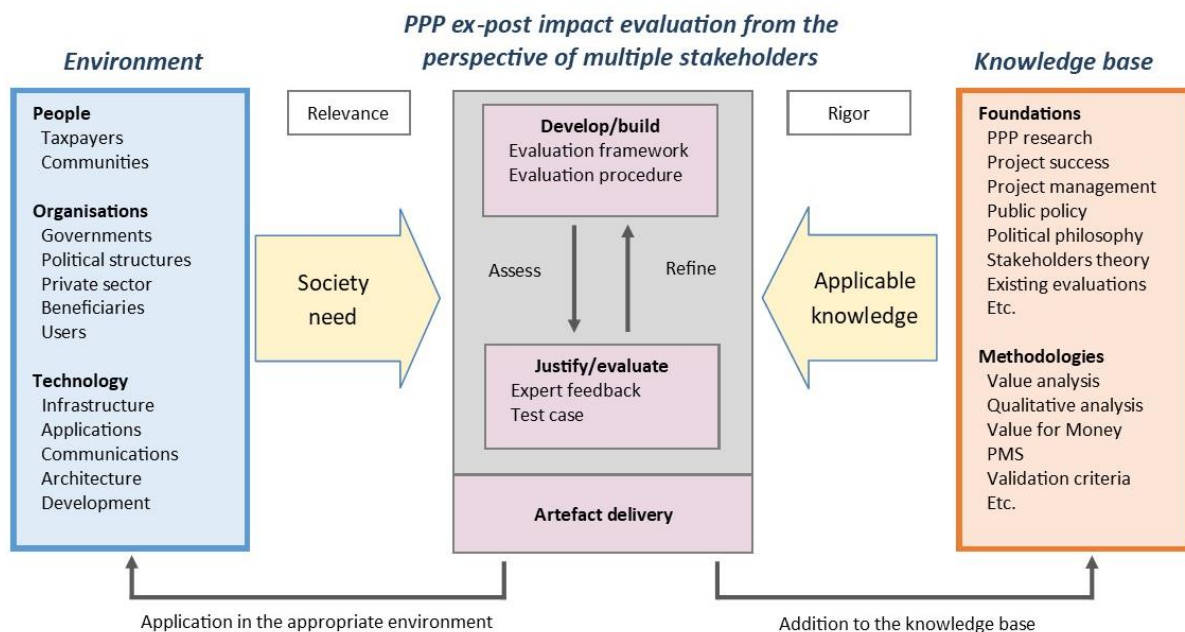


Figure 4-2: Design Science framework applied to this research study

[Source: Based on Hevner et al. (2004)]

Figure 4-3 describes the research strategy with four main stages: (A) development, (B) first testing, (C) second testing, and (D) third testing. Each stage is an iteration of develop/build and justify/evaluate processes. The main theoretical work is done in the first stage, and then the practical development of the evaluation method (procedure) is done throughout the three remaining testing stages. The artefact delivery (shown at the end of the design process) is for packaging and finalising the outcome.

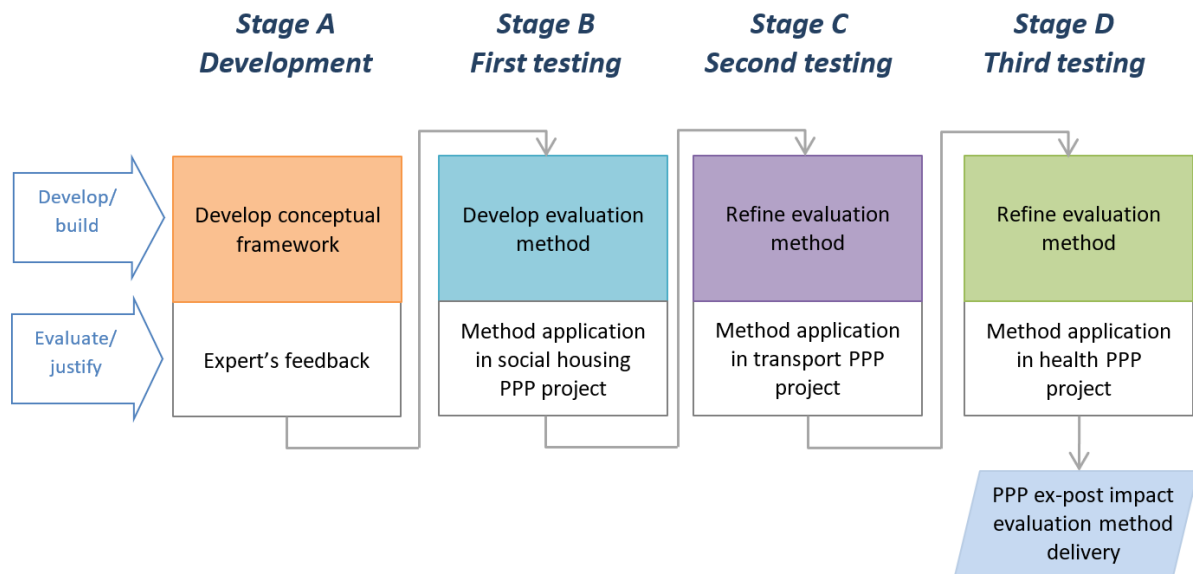


Figure 4-3: Research strategy presented as design iterations

[Source: Author's]

Figure 4-4 shows the detailed design process using an 'input – method/action – output' diagram. The chapters in which the different processes and outputs are presented within this thesis are included in the last column.

Process	Inputs	Method/action	Outputs	Chapters
Stage A - Development				
Develop/build	Literature review	Critical analysis	Gap in knowledge/ real problem	2 and 3
	Gap in knowledge/ real problem	Theoretical thinking	PSEPM - evaluation method ideas	5
Justify/evaluate	PSEPM - evaluation method ideas	Experts interviews	Suitability of the PSEPM	6
			Insights for evaluation method development	6
Stage B - First testing				
Develop/build	PSEPM - evaluation method ideas	Design -prototyping exercises	Prototype evaluation method	Appx. D
	Insights for evaluation method development			
Justify/evaluate	Prototype evaluation method	Test case 1 (social housing)	TC1** refinement insights	7.1
Stage C - Second testing				
Develop/build	TC1** refinement insights	Design	Prototype evaluation method with modification	Not inc.
Justify/evaluate	Prototype evaluation method with modification	Test case 2 (transport)	TC2** refinement insights	7.2
Stage D - Third testing				
Develop/build	TC2** refinement insights	design	Prototype evaluation method with modification	Not inc.
Justify/evaluate	Prototype evaluation method with modification	Test case 3 (health)	TC3** refinement insights	7.3
Artefact delivery				
	TC3** refinement insights	Procedure development	PPP Ex-post impact evaluation method (final)	8
	modified evaluation method	Re-analysis of test cases	Test cases evaluation reports	7

*PSEPM: Project Success Evaluation Pyramid Model / **TC: test case

Figure 4-4: Conceptual flowchart of the research project

[Source: Author's]

Stage A (development) answers to the first research objective.

- *Objective 1: Develop a conceptual framework that informs a PPP evaluation*

Stages B, C, and D (testing process) answer to the second research objective.

- *Objective 2: Develop, test and refine a method that can be used to conduct a PPP ex-post impact evaluation based on the conceptual framework*

Each stage is described in detail in the following subsections.

4.3.1. Stage A - Development

The development stage is to (1) understand the relevance of the problem, (2) explore the context of PPP evaluation in academia and in practice, (3) support the need for a solution, and (4) design an initial solution (PPP evaluation). The outcome of this stage is a conceptual framework that supports the development of the evaluation method, and the initial design of the evaluation method.

Develop/build process

To understand the problem and its relevance, a literature review was conducted. This confirmed that in the 'environment' (employing Design Science taxonomy) there is a problem that can be solved. The Literature review was also for exploring the different existing approaches that can be applied for evaluating a PPP.

The theoretical development, that involved critical thinking (Bowell & Kemp, 2014) and literature review, delivered as an outcome the conceptual framework called 'Project Success Evaluation Pyramid Model' (PSEPM) described in Chapter 5.

Justify/evaluate process

The justify/evaluate process at this stage takes an exploratory and feedback approach through in-depth interviews with PPP experts and practitioners. The aim was to (1) explore and discuss the challenges for designing a PPP ex-post impact evaluation, and (2) gather feedback from the presentation of the PSEPM. Some information about the evaluation method draft was also included in the presentation to the experts, and their feedback considered partially this development. The data collection and analysis section (4.4.1) provides more details about the interviewers and the interview structure.

The outcome of the justify/evaluate process is presented in Chapter 6 as a set of refinement insights divided into PPP general context and PPP evaluation method feedback (refer to Figure 4-4, development stage).

4.3.2. Stage B – First testing on test case

Develop/build process

In Stage B, the prototype evaluation method was designed. The inputs for designing it were three. First, the feedback from the experts regarding the context of PPP evaluation. Second, the feedback from the experts regarding the proposed conceptual framework. Third, two prototyping exercises of the stakeholder workshop which is one specific components of the evaluation. The workshop is the interaction between the evaluator and the stakeholder to gather the stakeholder's perspective towards a project. These two prototype exercises were conducted with a contractor's representative and a university authority respectively. The three inputs for designing the prototype evaluation are presented in Chapter 6.

Justify/evaluate process

Stage B was the first testing of the evaluation method in a real PPP project (social housing project in Chile). The outcome of the testing can be typified as: (1) ex-post evaluation results of the social housing project from the perspective of their multiple stakeholders; and (2) refinement insights proposed by the researcher through observation during the process of testing the method on the test cases. The project contexts, the project evaluations (the result of applying the evaluation method to a project), and the refinement insights to improve and develop the evaluation method are presented in section 7.1.

According to Design Science taxonomy, a case study refers to a real situation in which the artefact is tested for refining and improving the artefact's performance (Hevner et al., 2004). This is different from the case study concept employed in traditional research that refers to an in-depth enquiry into a specific and complex phenomenon (Yin, 2013). To avoid any confusion, this research employs the concept of 'test case' when referring to the three PPP projects in which the evaluation method was applied and then refined with the observations that were gathered from the artefact application.

4.3.3. Stage C – Second testing on test case

Stage C is the second testing of the developed evaluation method on another PPP project. Again, employing Design Science taxonomy, the build/develop process is the refinement of the evaluation method using the insights of stage B. Thus, the justify /evaluate process is the second testing on the test case, a toll road PPP project in Chile. The outcomes of this stage are (1) ex-post evaluation results of the PPP toll road from the perspective of their multiple stakeholders; and (2) refinement insights from applying the method in the test case. The project context, the results of the evaluation, and the refinement insights (lessons learned) that informed the subsequent method refinement are presented in section 7.2.

4.3.4. Stage D – Third testing on test case

Stage D is the third testing of the developed evaluation method on a final PPP project. The build/develop process is the refinement of the evaluation method using the insights of stage C. The justify/evaluate process is the third testing on the test case, a PPP hospital in Australia. The outcomes of the stage are (1) ex-post impact evaluation results of the PPP hospital from the perspective of their multiple stakeholders; and (2) lesson learned by the researcher through observation during the process of testing the method in the test case. The project context, the results of the evaluation, and the refinement insights (lessons learned) that informs the final refinement of the method are presented in section 7.3.

Even though stages B, C, and D are presented as a sequence of developing, testing, refining, and then testing again; in practice, this process is not completely sequential as the different stages had some components that were developed simultaneously. This means that the test cases inform each other.

4.3.5. Artefact delivery

The design process and the second objective of the research finishes by delivering the artefact that was developed and refined. This delivery involves two activities. First, packaging the procedure for conducting a PPP ex-post impact evaluation method presented in Chapter 8. This includes the theory, the steps, the explanation of activities, the instruction for data collection and analysis, the forms, the templates, and every piece of information that is relevant to conduct the evaluation. This procedure (Chapter 8) is self-contained, meaning that it should be able to be understood without reading any other part of the thesis. It summarises part of the theory and provides brief explanations of some topics that were developed in detail in other chapters. The second activity for delivering the artefact is a final feedback process with some of the experts that were interviewed in Stage A. They were asked

to provide feedback about the proposed evaluation method after presenting them the major results that were obtained from the test cases. The aim is to analyse whether the final evaluation method is effective and efficient.

After presenting the research strategy, the data collection and data analysis are described for each of the stages (A, B, C, and D).

4.4. Data collection and analysis

the data collection and analysis for each of the four described stages is summarised in Table 4-3. In total 13 experts were interviewed, and from the three test cases, a total of 25 stakeholder's representative were contacted to participate in the evaluation method testings. Each stage is described in detail in the following subsections.

Table 4-3: Summary of data collection and data analysis

Stage	n	Unit	Type of collection	Data analysis
Stage A	13	Experts	Interviews	Content analysis
Stage B	6	Stakeholder's representatives	Stakeholder workshops (1-hour meeting)	Functional testing
Stage C	11	Stakeholder's representatives	Stakeholder workshops (1-hour meeting)	Functional testing
Stage D	8	Stakeholder's representatives	Stakeholder workshops (1-hour meeting)	Functional testing

[Source: Author's]

4.4.1. Stage A - Development

Interviews

Interviews with PPP experts and practitioners had two objectives: (1) gather feedback from the presentation of the PSEPM, and (2) explore and discuss the challenges for designing a PPP ex-post impact evaluation. The first objective can be considered action research (Järvinen, 2007; Rapoport, 1970) because their feedback about the artefact's design is considered for its development; thus, experts are part of the design process. For the second objective about the challenges of ex-post evaluation, expert interview for exploratory purposes can be more efficient than other forms of research; experts can provide a broad and practical view that can lead to better results (Bogner, Littig, & Menz, 2009). Experts, in this case, are recruited for their technical knowledge, which differs from interpretive knowledge and process knowledge (Bogner et al., 2009, p. 52). Technical knowledge, specific in a certain topic, is gathered through a systematising-expert interview (Bogner et al., 2009, p. 43), that complements the explorative nature of the interview, searching for specific knowledge.

The experts are from Australia and Chile, two countries in which the PPP environment is considered 'mature', leading in their region (South Pacific and Latin America respectively) (KMPG, 2015; The Economist Intelligence Unit, 2014). The experts were purposefully selected because of their firsthand experience and in-depth understanding of PPP projects. They are from the public and private sector, involved in many PPP projects from conception to completion. Some have also been involved in policy formulation around PPPs. Prior to the interview, they were asked to give consent. For confidentiality reasons, it is not possible to reveal their identity.

The selection of the experts was done through an Expert Sampling approach. The criteria for the selection of the experts (as a diverse group) are:

- 1) Work experience in (1) PPPs (2) public evaluation, (3) social housing, and (4) project management
- 2) Level of authority that includes strategic and management decisions and thinking
- 3) From the two contexts in which the study is mostly based (Chile and Australia)
- 4) Public and private sectors

Table 4-4: Experts and practitioners involved in Stage A – development

Interviewee	n	Description	Country	Sector
Interviewee A	1	Former regional minister of social development (entity in charge of performing any evaluation ex-ante and ex-post for any kind of public investment)	Chile	Public
Interviewee B	1	Former director of infrastructure at the Public Infrastructure Ministry	Chile	Public
Interviewee C	1	Former minister of public infrastructure	Chile	Public
Interviewee D	1	Chief office director, regional department of the Social Development Ministry, Biobio Region	Chile	Public
Interviewees E	4	Social housing department team (4 people), social housing developer	Chile	Private
Interviewee F	1	PPP director, consulting company	Australia	Private
Interviewee G	1	Former authority for major projects	Australia	Public
Interviewee H	1	Principal policy adviser, social housing, Department of Health & human services, Victoria	Australia	Public
Interviewee I	1	Construction manager at a mega-contractor company	Australia	Private

[Source: Author's]

The structure of the interviews was as follow: A 20-minute presentation of the proposed conceptual framework for evaluating PPP projects was presented (1) to frame the interview, and (2) for feedback purposes in the design process (Venable, Pries-Heje, & Baskerville, 2012); then, a 30-minutes discussion regarding (1) the conceptual framework developed in Chapter 5 (Project Success Evaluation Pyramid model) and (2) the relevance and challenges of designing and conducting ex-post impact evaluations. They were also asked about the relevance of the different stakeholder views in an ex-post evaluation.

Data analysis

In this stage (development), the data from the experts' interviews and focus group was analysed using thematic analysis (Nowell, Norris, White, & Moules, 2017). The objective is to conceptualise in a logic form the main ideas and insights from the experts to use it as feedback for the conceptual framework and the evaluation method. The interviews were transcribed verbatim (word by word) and a thematic analysis was conducted employing MS Excel. The analysis generated 170 pieces of information coded with more than 40 key concepts. Each piece of information is an idea presented by the interviews in the form of a sentence or paragraph from the interview transcription. Chapter 6 presents the feedback categorised in themes, and a summary of the design inputs is presented in section 6.4.

4.4.2. Stages B, C, and D – First, second and third testings

Test case

To test and refine the evaluation method, three test cases were selected employing a convenient sampling: a social housing PPP project, a toll road PPP project, and a PPP hospital. The rationale behind the project selection is threefold. First, a requirement was to have access to the project documentation and to the stakeholders. This requirement reduced largely the spectrum of available options but it is considered essential for studying in detail the impacts of a PPP and for testing the developed evaluation method thoroughly. The second requirement is that projects should represent the diversity of options that are currently employed in the real world. According to KMPG (2015), the PPP model is applied to a variety of projects, sectors, regions, and business models. Therefore, the three test cases aim to represent this diversity. Third, the Chilean and Australian contexts were chosen to add another aspect of diversity for the testing of the evaluation method. The social housing and the toll road projects are located in Chile, and the hospital project is located in Australia.

To select three different PPP projects was considered a better option than selecting three similar ones (e.g. three PPP toll roads) because the method could be tested with a higher range of variability and contexts. However, the three projects have similar characteristics that are important for the way the evaluation method was developed. First, the three projects present a similar power/interest structure; the public sector and the private entity are both strong stakeholders with some conflicting interests. Second, in all projects, the user/beneficiary is an important stakeholder (as all projects provide a public service). Third, they all have a similar level of exposure to the public, different from an economical infrastructure PPP project such as a port, or a power plant, in which the users/beneficiaries are mostly companies that work strictly for business purposes. Finally, the three projects employed a PPP procurement strategy for the achievement of short-term public welfare, which is different from the indirect and long-term public welfare of economic infrastructure PPP projects.

For each project, the selection of the stakeholders was based on the generic characterisation of (1) public authority, (2) private entity, (3) end-user, and (4) society. However, the inclusion of them was re-evaluated, as the identification of stakeholders is one of the evaluation method's steps.

Table 4-5: Test cases description for research stages B, C, and D

Description	Social Housing Project	Toll Road Project	Hospital Project
Scope	97-units social housing new project	East-west urban highway in a Chilean city	Main Australian hospital relocation
Completion year	2010	2006	2011
Budget	MMUSD 1.5 (estimation)	MMUSD 1.200,5	MMAUD 421.5
Procurement method	Manage-Design-Build	Design-Build-Operate	Design-Build-Maintain

[Source: Author's]

Analysis

The design process involves applying the (proposed) evaluation method to the described PPP projects, obtaining as a result, a project evaluation (evaluation result). The information collected for each test case is mainly from stakeholder's representatives as part of the evaluation method application. The data collected from the test cases (stages B, C, and D) presented as 'project evaluation' was analysed using *Functional (Black box) Testing*, which is part of the Design Science Research (Hevner, 2004). The researcher analyses the data (project evaluations) searching for problems that can be fixed by modifying the designed artefact. For each test case, the researcher identified problems that can affect the efficiency and effectiveness of the artefact. These problems are presented as refinement insights that can modify (or develop in detail) some component of the evaluation method. The project evaluations results and the refinement insights are presented for each test case in sections 7.1, 7.2, and 7.3.

4.5. Research project validation as a Design Science study

This research project is a valid Design Science research because it meets Hevner's validity guidelines that include seven principles:

- Design as an artefact: An evaluation method based on a Project Success concept is a valid artefact as a construct and a method (section 5.6).
- Problem relevance: As explained in Chapter 1, PPPs have been questioned by the media and academia. There is a lack of a method to ex-post evaluate the impact that PPPs have over their stakeholders. The literature recognises ex-post evaluation as a relevant component for successful PPP procurement. At a government level, decisions are being made approving or cancelling PPP projects without full information about how past PPP projects are performing

or impacting society. This supports the claim that the problem is relevant and the solution is valuable.

- Design evaluation: The artefact (ex-post impact evaluation method) was tested using real PPP projects and then presented to experts for a final feedback process. The artefact was tested and refined in every iteration of the research process (section 4.3).
- Research contribution: The study contributes to the theory of Project Success, to the development of PPP projects, and to the use of Design Science in the project management discipline.
- Research rigour: (1) In-depth literature review in a broad spectrum of disciplines, (2) a multidisciplinary approach to address the problem, and (3) a process that includes a real connection with the industry, gives the research the rigour Design Science approach expects.
- Design as search: Within the management field, the researcher understands the actual problem that leads to the lack of ex-post assessments and uses all the knowledge from literature and practice to design a solution.
- Communication of research: Employing Design Science definition of audience (Hevner et al., 2004), the research will be in contact with a technology-oriented audience from the academic sector, and with a management-oriented audience from the public and private entities during the development and testing stages. The procedure (Chapter 8) will be re-formatted into a professional report for a full distribution beyond this thesis.

4.6. Summary

Chapter 4 presents the research strategy to undertake the research project. It employs the philosophical paradigm of Critical Realism and the Design Science methodology.

Critical Realism (CR) paradigm is considered a bridge between the traditional paradigms of Positivism and Constructionism. Positivism believes in a single reality that can be studied and measured, and the latter believes knowledge can be socially constructed. CR stands for multiple layers of reality, in which objective mechanisms create events that subsequently create subjective experiences, allowing the possibility for creating (constructing) knowledge without denying the existence of objectivity and single reality objects. The adopted methodology under the meta-theory of CR is Design Science. DS differs from the traditional behavioural science approach because it does not have the aim of describing nature, but to change nature with a designed 'artefact'. The structure of a Design Science approach is to develop an artefact through a series of iterations of development, testing and refinement. The final version of the artefact should solve a problem in the real world, and it should also contribute with knowledge to science.

To answer the main research question *How can ex-post impact evaluation of Public-Private Partnership projects be conducted employing the perspective of its multiple stakeholders?* The research strategy consists of four main stages that represent the iterations of the design process: (A) development stage, (B) first testing, (C) second testing, and (D) third testing. Additionally, the artefact delivery is included as a final process. In the first stage (development) a conceptual framework was developed and then tested with feedback from PPP experts and practitioners. In the three remaining testing stages, the evaluation method was fully developed, tested on real PPP projects, and then refined with the refinement insights of each testing. After the third testing process, the artefact (ex-post impact evaluation method for PPP projects) was presented as a procedure (Chapter 8) and the results of the testings were presented to PPP experts for final feedback.

The research methods are interviews for the development stage; and test cases for the testing stages. The test cases are three: social housing PPP project in Chile; A PPP toll road project in Chile; and a PPP hospital in Australia. This research adopted the concept of 'test case' instead of 'case study' to avoid

the confusion between the Design Science understanding (a real situation for testing an artefact) with the traditional understanding (a specific researched case to infer a larger phenomenon).

The next two chapters present the first stage of the research (development). Chapter 5 describes the conceptual framework called 'Project Success Evaluation Pyramid Model', which is the first deliverable in the design process, and Chapter 6 shows the feedback process obtained through experts' interviews.

CHAPTER 5

CONCEPTUAL FRAMEWORK DEVELOPMENT

5.0. Introduction

This chapter describes and explains the proposed conceptual framework called ‘Project Success Evaluation Pyramid Model’ (PSEPM). This model conceptualises the process for reaching a Project Success judgement, identifying the success criteria from all the involved stakeholders and then assessing their validity using a set of principles.

This chapter is a sequential theoretical argument that ends with the conceptual framework PSEPM. First, it will argue why multiple-perspectives evaluations are applicable to PPPs; second, it will present the reasons for selecting the Project Success concept as a basis for the conceptual framework; then, it will discuss the theoretical development that leads to the proposed conceptual framework that is presented at the end. This chapter and the PSEPM answers to the first research objective:

Objective 1: Develop a conceptual framework that informs a PPP evaluation

The structure of the chapter is as follow. Sections 5.1 and 5.2 discuss the reasons for employing the project success concept as a basis for the conceptual framework. Sections 5.3 to 5.5 address all the theoretical components that were developed as a support for the conceptual framework. The PSEPM is presented in section 5.6.

5.1. PPPs should be evaluated using multiple perspectives

The departing point of the theoretical development is the argument that PPPs need to be evaluated from multiple perspectives. This is one of the problems for evaluating PPPs. Thus, single-perspective evaluations are analysed in this section, describing a flaw that makes them incompatible with PPP ex-post impact evaluation. This is not an underestimation of their capacity; they are perfectly suitable for many other evaluations but arguably not for a PPP ex-post evaluation.

Single-perspective evaluations are here conceptualised as the approaches that include only the criteria that concerns the evaluation team or the person who is conducting (driving) the evaluation. They generally do not address in a systematic way the complexity of having multiple relevant perspectives. Single-perspective evaluations and audits have a common strategy that can be synthesised as a two-step procedure: criteria selection, and criteria assessment. The criteria selection should be done by an expert team and can have a big spectrum of possibilities, from the most technical (cost, time, budget, profit, waste, CO2 emission, etc.) to the most social/perceptive criterion (compliance, political revenue, VfM, social impact, satisfaction, good or bad, etc.). The criteria selection depends on the type of evaluation and on the client/user/requester of the evaluation. The second step is to assess the selected criteria employing one or more of the following strategies (based on Gertler, Martinez, Premand, Rawlings, & Vermeersch, 2016, p. 9; Owen & Alkin, 2007):

- Comparison before – after (e.g. for social interventions)
- Comparison expected - actual (e.g. contract compliance)
- Comparison project A - project B (e.g. performance comparison)
- Comparison project A – standard (e.g. minimum requirements)
- Comparison project A – expert judgement (e.g. deep qualitative evaluation)
- Etc.

The problem with single-perspective evaluation is that it only addresses the perspective of the evaluation team, which leads to a legitimacy problem. Legitimacy refers to the acceptance of an authority for taking decisions. Using Max Weber's classification, the rational-legal legitimacy is the one involved in this context; where the rationality is assumed to be supported by the public interest and represented by a legal authority (Ringer, 2004, p. 183). A PPP is a public project and therefore, it should be legitimate for its rationality. Legitimacy is not an issue when the power is concentrated by only one stakeholder. However, a PPP is different because it works with a series of stakeholders that have power and interests that can generate conflict. A PPP has (1) a broad public sector (different agencies and ministries), (2) a private consortium conformed by several companies, (3) an aggregated private sector (lobbies and guild groups), (4) a significant role of the users/beneficiary, and (5) an even more determinant extended community (citizens) that can vote for an overall strategy in favour or against PPPs. Legitimacy in a PPP is difficult to achieve if the criteria are defined only by one group of interest, even if they pursue the most objective and rational possible choice.

Another problem is that usually existing evaluations address technical issues to the detriment of 'non-technical' or 'perceptive' issues, which in some cases are the real decision-making triggers (i.e. political discussion). Perceptions are critical for the development of PPPs. Whether they are correct or incorrect, stakeholders' perceptions are often the basis of acceptance or rejection of PPP use and outcomes. Heald (2003) and Bauld and McGuinness (2006) refer to the subjectivity of the Value for Money indicator, suggesting that the concept 'needs' a non-technical component.

This section has argued the limitations of single-perspective evaluation if applied to PPP projects, making a case for preferring a multiple-stakeholder perceptive approach. Thus, the next section analyses the existing approaches that considered this attribute.

5.2. Project Success as a departing point for developing the conceptual framework

A broad review of existing evaluations was described in the literature review chapter (section 3.1) where 30 evaluation approaches were gathered (summarised in Table 5-1). As stated in Chapter 3, the reviewed approaches are not suitable for directly evaluating a PPP project. Therefore, this section discusses why the Project Success concept was selected as the most suitable approach for evaluating PPPs from multiple stakeholder perspectives (aspect included in the last column of Table 5-1).

Table 5-1: Summary of evaluation approaches (summarised version of Table 3-1)

Reference	Evaluation	Involved stakeholders	Multiple stakeholder
Research on PPP evaluation and monitoring (section 3.3)			
Heald (2003)	Value for Money evaluation	Private, public	NO
European PPP Expertise Centre (2011a)	PPP ex post evaluation guidelines	Not specified	NO
Love and Liu (2014)	Whole life assessment	Private, public, user, community (based on the PMS Performance Prism)	YES
Management-based evaluation and monitoring approaches (section 3.4)			
De Wit. (1988)	Project Success evaluation	Private, public, user, community	YES
Barclay and Osei-Bryson (2010)	Stakeholder management	Private, public, user, community	YES
Kaplan, Norton (1995) Neely, Adams, & Crowe (2001)	Performance measurement system (PMS)	Private, public, user, community (holistic view of Performance Prism PMS)	YES
Nalewaik and Mills (2017)	Project audit	Private, public, user, community	YES

Program evaluation (public policy) approaches (section 3.5)			
Quade and Carter (1989)	Before/after Program	User, community	NO
	Forecast/actual performance of the program	User, community	NO
	With program/without program	User, community	NO
	Experimentation with users	User, community	NO
	Planned/actual results	User, community	NO
Langbein and Felbinger (2006)	Descriptive – implementation	Private (contractor)	NO
	Descriptive – output	User, community	NO
	Causal – implementation	Private, user	NO
	Causal – output	Private, user	NO
Owen and Alkin (2007)	Objective-based evaluation	Private, user	NO
	Needs-based evaluation	User	NO
	Goal-free evaluation	User, community	NO
	Process- outcome evaluation	Private, user	NO
	Realistic Evaluation	User	NO
	Performance audit	Private	NO
State Government of Victoria (2008)	Gateway Process	Public, user	NO
Ministerio de Desarrollo Social (2017)	Ex-post evaluation: short term and long-term	Public, user	NO
International agencies' approaches (section 3.6)			
World Bank (2016)	Beneficiary assessment	User	NO
	Social impact assessment	User, community	NO
World Bank (2012)	Environmental assessment	Public, community	NO
Zimring and Reizenstein (1980)	Post-occupancy evaluation	Private, user	NO
United Nations (Lund-Thomsen, 2007)	Qualitative PPP impact evaluation	User	NO
Jackson & Kassam, 1998	Participatory evaluation	User	NO

[Source: Author's]

From the 30 approaches, four approaches were recognised by the researcher as multiple stakeholder perspectives: the Performance Prism, the whole life PPP evaluation approach, the project audit approach, and the Project Success concept.

The three first approaches present flaws and they are not considered good departing points. The Performance Prism is a Performance Measurement System (PMS), developed By Neely et al. (2001). It is considered a more comprehensive PMS than the well-known Kaplan and Norton's PMS Balance Scorecard as it includes all the stakeholders. However, the Performance Prism is not considered suitable for developing a PPP ex-post impact evaluation method because a PMS is not an evaluation, as it does not consider a decision-making phase. A PMS might contribute in monitoring and for performing formative evaluations, but not for an ex-post impact evaluation. The second approach is the whole life PPP evaluation developed by J. Liu, Love, Davis, et al. (2015), which is based on the Performance Prism and it involves multiple stakeholders. However, it aims to monitor and improve a PPP performance during its lifecycle and the overall impact is not its focus. The third approach is the Project Audit body of knowledge that is traditionally related to financial and compliance objectives. Nalewaik and Mills (2015) audit approach expands the scope of the audit, considering multiple stakeholders. However, their approach still has the auditor as the main party that connects with the

other stakeholders for specific purposes. Therefore, (without underestimating its capacity) it is not fully compatible for a multiple stakeholder perspective as defined in this research.

The selected approach for developing the conceptual framework is Project Success, which provides a straight-forward view for evaluating a PPP from multiple stakeholder perspectives; a project will be successful if the expectations of all stakeholders are met (Davis, 2016). Even though the Project Success concept is very theoretical with rare application in practice, this approach is considered significantly better for designing a PPP ex-post impact evaluation approach from the perspective of multiple stakeholders. The aim and the purpose of the concept involves, explicitly, the perspective of the project stakeholders.

The selected evaluation approach is the starting point for developing the conceptual framework. The following sections are based mostly on the Project Success concept.

5.3. Refinement of the Project Success understanding

As reviewed in Chapter 3, in the literature there is no agreement about the definition and measurement of Project Success. However, it is considered the best choice for developing a conceptual framework to evaluate PPP projects. Therefore, this section will discuss inconsistencies in the literature and it will propose a new perspective of Project Success that has not been addressed from a theoretical viewpoint.

Employing as a starting point the understanding of Project Success as ‘meeting stakeholders’ expectations’; five issues will be discussed in detail to contribute to the conceptual framework.

- Success criteria should be based on stakeholder expectations;
- Success criteria are subjective;
- Stakeholder satisfaction is not a success criterion;
- Existing success frameworks can be represented as sets of stakeholders’ expectations; and
- A successful project is different from a perfect project.

The outcome of the five analysed issues is a proposed refinement of the Project Success definition as ‘meeting a substantial portion of stakeholders’ expectations’ which addressed most of the issues that are conceptually developed in this section.

5.3.1. Success criteria should be based on stakeholder expectations

Two important concepts in the research of Project Success are ‘critical success factors’ and ‘success criteria’. By definition, ‘critical success factors’ (CSF) are the key factors necessary to achieve success (Rockart, 1980); and the existing literature has focussed on them to increase the possibilities of achieving successful projects (Shenhar et al., 2001). Conversely, ‘success criteria’ are defined as the measurements to judge if the project was successful or a failure (Ika, 2009). Lim and Mohamed (1999, p. 243) make a clear distinction of the differences between criteria and factors; a criterion is “a principle or standard by which anything is or can be judged”, and a factor is “any circumstance, fact, or influence which contributes to a result”. Using the same example of section 3.4, a success factor can be ‘team expertise’ (necessary to achieve a successful project) while a success criterion can be ‘time of project delivery’ (a delay in the delivery can mean that the project is a failure). Hence, logically, if a success factor does not affect any success criteria, it cannot be considered a factor. The only way to prove that a success factor increases the possibilities of getting a successful project is by defining a success criterion that is affected by that factor; because the only way to measure success is with success criteria. The problem of developing success factors without a clear understanding of success criteria was pointed out in 1994 by Rodney Turner (Ika, 2009).

A third concept that is related to the overall success of the project is stakeholders' expectations. The literature highlights the relevant influence that stakeholders have over the Project Success judgement (Davis, 2013). The stakeholder expectation is what every stakeholder involved in a project expects from the outcomes of the project or from the processes that occur during its lifecycle.

The theoretical connection between success factors, success criteria and stakeholder's expectations is presented in Figure 5-1. Within a single project, *success factors* are actions that increase the probability of obtaining a successful project. Yet, the success of the project is measured using a set of *success criteria*. Consequently, every success factor is related to a set of one or more *success criteria*. The problem is that there is no agreement regarding which are the *success criteria* that can be employed to measure success. However, if a successful project is the one that 'meets stakeholders' expectations', then success can be measured by stating stakeholders' expectations and comparing them with what they perceive. In other words, *success criteria* should be the criteria in which stakeholders expect something from the project. For example, if a stakeholder expects the project to be completed before a certain due date, then 'time' should be considered a success criterion. This logical way to engage the three concepts has not been addressed by the current literature, except partially by Westerveld (2003) 'project excellence model®'. Connection between success factors and criteria is more common (Ika, 2009) but the explicit inclusion of the third concept (stakeholder expectations) was not found in the reviewed literature.

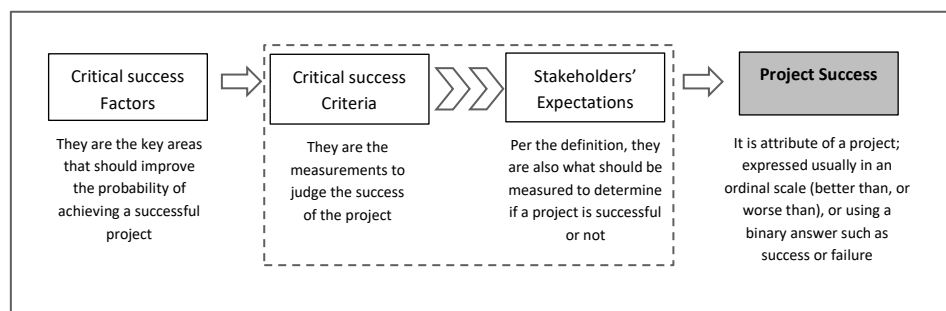


Figure 5-1: Theoretical relationship between semantic concepts of Project Success

Turner and Zolin (2012), McLeod, Doolin, and MacDonell (2012) among others, have empirically connected stakeholder expectations with Project Success through a forecasting model and a case study respectively. However, the theoretical implications were not explicitly addressed. Figure 5-1 synthesises that success criteria refers unavoidably to specific stakeholder expectations; even when a 'objectivist' (positivist) approach of Project Success is taken. Further discussion is needed over this last point; the next section addresses the subjectivity of success criteria.

5.3.2. Success criteria are subjective

A successful project has been defined as one meeting all stakeholders' expectations. Hence, any criteria that can be defined to measure success is 'subjective' to the stakeholders' interests in which they have an expectation. Consequently, a success criterion needs to be connected to at least one stakeholder in order to be valid. This conclusion underlines that it is not possible to establish a set of criteria that can work in every project, unless a stakeholder wants to use the same criteria in all their projects (e.g. owner, project team). This argument differs from existing success frameworks such as those by Atkinson (1999); Bronte-Stewart (2015); Pinto and Slevin (1988). These authors established that criteria such as time, cost and quality (among others) should be assessed in order to obtain a Project Success measurement. Implicitly, they are suggesting certain criteria without considering the involved stakeholders. The researcher argues that a success framework to measure Project Success must first identify the stakeholders and then their success criteria. It is undeniable that there are success criteria that are mostly inevitable (such as time and cost), but their validity should depend on

at least one stakeholders' interest. The subjectivity of success criteria does not mean that the current success frameworks are incorrect or invalid; it means that they are possible suggestions that need to be confirmed with the specific stakeholders of a project.

The statement that success criteria are subjective can contrast a legitimate perspective that there are some universal criteria, such as world peace, or human surviving (extreme examples) that are not subjective, but objective. The answer to this concern is that in theory, these 'objective' criteria are only relevant to most humans in the world today. However, similar to be in the 13th Century and believe in witch-hunts, nothing can prove that actual beliefs are correct and that they will not change in the future. This argument supports the statement that any success criterion is subjective to the stakeholders, even considering that in practice this might not make any difference to essential criteria such as environmental preservation.

Ika (2009) describes the difference between a objectivist and a subjectivist approach of Project Success. While the former refers to a fix set of procedures and knowledge to achieve Project Success, the latter understands the concept as a social construct. The issues presented here supports the 'subjectivist' approach as the correct overarching theory, in which 'objectivists' tools and methods could be applied. Empirically, McLeod et al. (2012) study the different stakeholders perspectives toward an office renovation project. His approach is consistent with this issue and supports the argument about the subjectivity of the Project Success judgement.

5.3.3. 'Stakeholder satisfaction' is not a success criterion

Success frameworks, e.g. those proposed by Atkinson (1999); Olsen (1971); Toor and Ogunlana (2010) are structured sets of success criteria proposed for assessing the success of a specific project. Some of these criteria are cost, time, quality (iron triangle), safety, risk, profit, welfare, etc.

An additional criterion, mentioned directly or indirectly in the literature, is 'satisfaction of stakeholders', 'client satisfaction', or 'meeting stakeholders' expectations' (Davis, 2016). This means that success can be assessed by measuring how satisfied are the different stakeholders. However, satisfaction should not be a feeling; it should be a judgement about the performance of an action (A. Liu & Walker, 1998) which is expressed in terms of certain criteria. In other words, satisfaction should not be considered a success criterion, it should be the results of a judgement process in which a stakeholder contrasts the expectations in several criteria against the actual outcomes of the project. A stakeholder can be satisfied with the time of the project, with the quality of the project, with participation in the process, etc. but cannot be only satisfied, as being 'generally satisfy' suggest a feeling rather than a judgement.

The following example demonstrates the flaw of employing 'stakeholders' satisfaction' as a success criterion. The project manager (PM) is considered one of the main internal stakeholders of a project and probably his/her satisfaction and expectations are part of any success discussion. Satisfaction is based mostly on what the PM perceives about cost, time and specification (Turner, 2009). Then, it is tempting to declare four criteria in the discussion: time, cost, specification, and PM satisfaction. However, there are only three success criteria (time, cost, specification) that support the 'perceived success' of the PM. Depending on how is the PM judgement regarding the three criteria, he/she will be more or less satisfied. The inclusion of this 'satisfaction criterion' in any success framework makes the Project Success definition redundant, because it measures success using the exact same criterion that defines it. This argument against the inclusion of satisfaction as a success criterion is not new and it has been identified by other authors such as Wateridge (1998) when referring to the iron triangle: "project manager have to aim for satisfaction for these three criteria to achieve Project Success". Another argument that can contribute to this confusion in that authors such as Toor and Ogunlana (2010) indirectly employ 'satisfaction of stakeholders' (conform to stakeholder expectations) for

referring only to the stakeholder that are outside the project, such as the user, the client, or the community; which is not precise, as all stakeholders (inners and outsider) have expectations to be conformed with.

5.3.4. Existing success frameworks can be represented as sets of stakeholders' expectations

Even though the definition of Project Success is unclear, many frameworks have been created to describe and analyse this concept.

To show the connection between the definition of Project Success and the expectations of the stakeholders, seven existing success frameworks are analysed:

1. Key factors in successful project implementations by Pinto and Slevin (1988),
2. Logical Framework Method by Baccarini (1999),
3. Micro an Macro view by Lim and Mohamed (1999),
4. the square root model by Atkinson (1999),
5. multidimensional strategic Project Success framework by Shenhar et al. (2001),
6. Performance framework for mega-projects by Toor and Ogunlana (2010), and
7. project status model by Bronte-Stewart (2015).

The analysis of the seven framework shows that the components of each framework can be expressed as a combination of one or more stakeholders' expectations. The analysis of frameworks classifies the stakeholders into four groups, based on the European Quality Excellence model (European foundation for quality management, 2015): (1) owner, (2) project management, (3) end user, and (4) society. This classification aggregates stakeholders such as the project manager, the contractor, and the project team into 'project management' group. Each framework is analysed by showing how the expectations of the stakeholders are connected to each of the frameworks' components. For example, Lim and Mohammed (1999) analysed Project Success with a macro and micro view, which can be expressed as the expectations of the project team (micro view) and expectations of the rest of the stakeholders (macro view). The proposed connections between the frameworks' components and the stakeholder expectations are only referential in each case. The aim of the analysis is to support the existence of a connection and not to prove a perfect homology.

(1) Key factors in successful project implementations

Pinto and Slevin (1988) divide Project Success into measurements related to the project and measurements related to the client (Figure 5-2). The project's criteria are time, cost, and performance; and the client's criteria are use, satisfaction, and effectiveness. What Pinto and Slevin recognise as 'the project' can be interpreted as the expectations of the project team, and what is recognised as 'the client' can be split between the user's expectations and the owner's expectations.

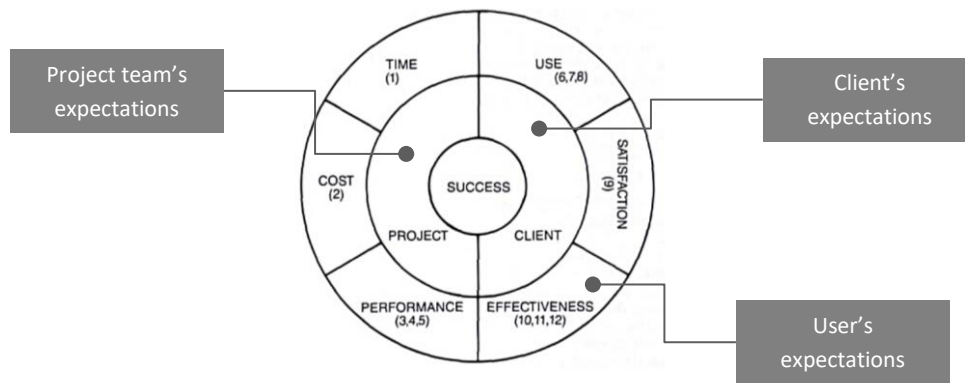


Figure 5-2: Link between successful project implementation and stakeholder's expectations

[Source: Adapted from Pinto and Slevin (1988)]

(2) Logical framework method

Baccarini (1999) employed the logical framework method (LFM) to analyse product success and project management success (Figure 5-3). LFM considers four concepts: inputs, outputs, purpose, and goals. Baccarini states that project management success is related to the inputs and the outputs of the project; and the product success refers to the purpose and the goal of the project. In terms of stakeholder' expectations, the input and the output from the LFM are what the project team expects. The purpose is the expectation of the end user and part of the society. Finally, the goal can be related to the expectations of the owner of the project. Figure 5-3 represents the connection between (1) the Project Success definition, (2) the logical framework method employed, and (3) the relationship with the 4 conceptual stakeholders. Baccarini explains that “project management can be a failure and the product can be successful”. This can be expressed as: the expectations of some stakeholders regarding the construction/implementation stage cannot be met, while expectations of some other stakeholders around the final product can be met.

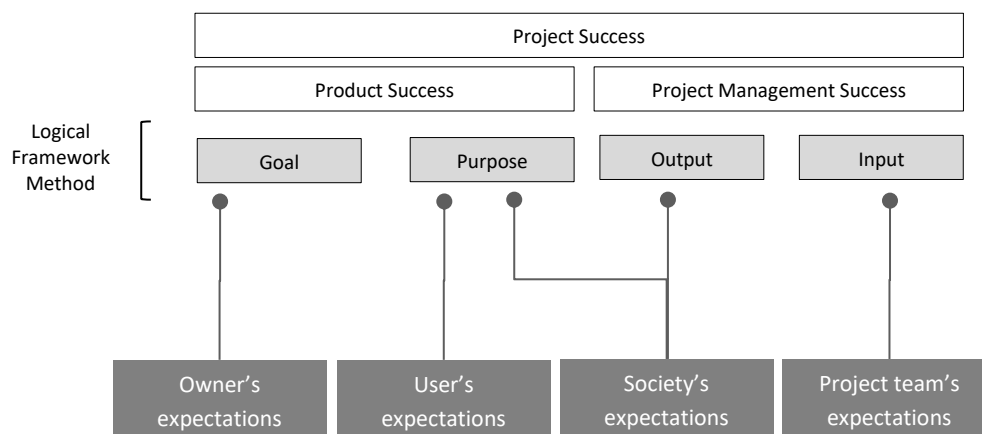


Figure 5-3: Link between Project Success, logical framework method, and stakeholder's expectations

[Source: Adapted from Baccarini (1999)]

(3) Micro and macro view

Lim and Mohamed (1999) disaggregate Project Success into a micro and a macro view (Figure 5-4). The micro view is usually related to the project management part, which is the construction and implementation stage. Conversely, the *macro* view is related to the conceptualisation of the idea, the tendering process, and the final product in terms of operational stage or impacts. In terms of stakeholders' expectations, the *micro* view relates to the project team expectations, while the *macro* view concerns mostly the owner, the user and the society. In the diagram, the owner could be attached to all phases.

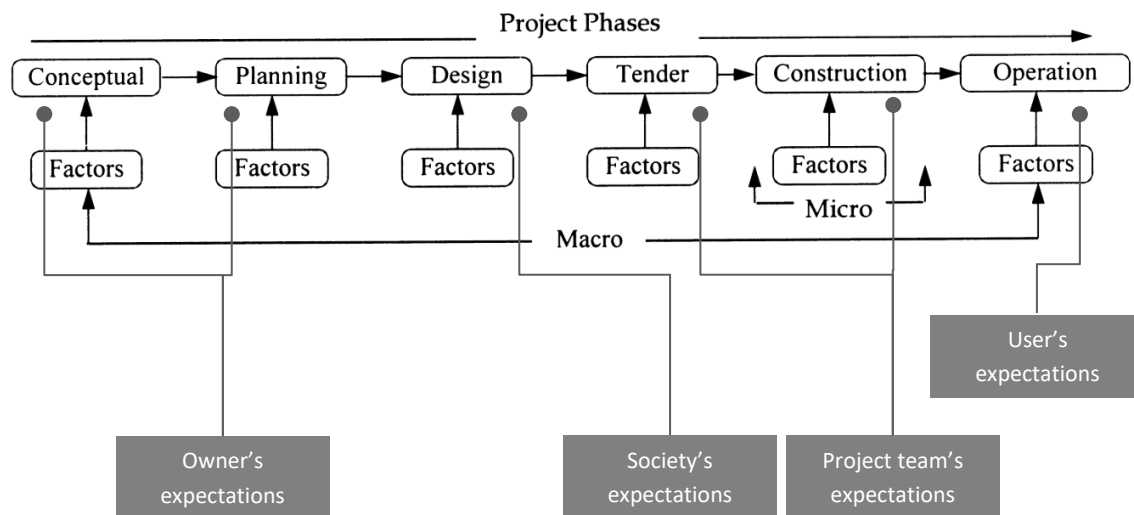


Figure 5-4: Link between project's micro and macro view, and stakeholder's expectations

[Source: Adapted from Lim and Mohamed (1999)]

(4) The square root model

Atkinson (1999) proposes a broader view of Project Success in the field of information system and information technology (IS-IT) beyond the iron triangle (cost, time, quality), including three additional components: (1) the information system, (2) the benefits for the organisation, and (3) the benefits for the community (Figure 5-5). The information system refers to the operational stage of a project and how to measure its performance; the benefits to the organisation refers to the impact inside the organisation in which the IT project was implemented; and the benefits to the community refers to the users, the environmental and social impacts, professional learning, etc. The analysis in terms of stakeholders' expectations shows that the project team's expectations are related to the iron triangle and the information system. It is also possible to connect the owner's expectations with the information system and the benefits to the organisation. Finally, the user and the society's expectations are closely related to the benefits to the community.

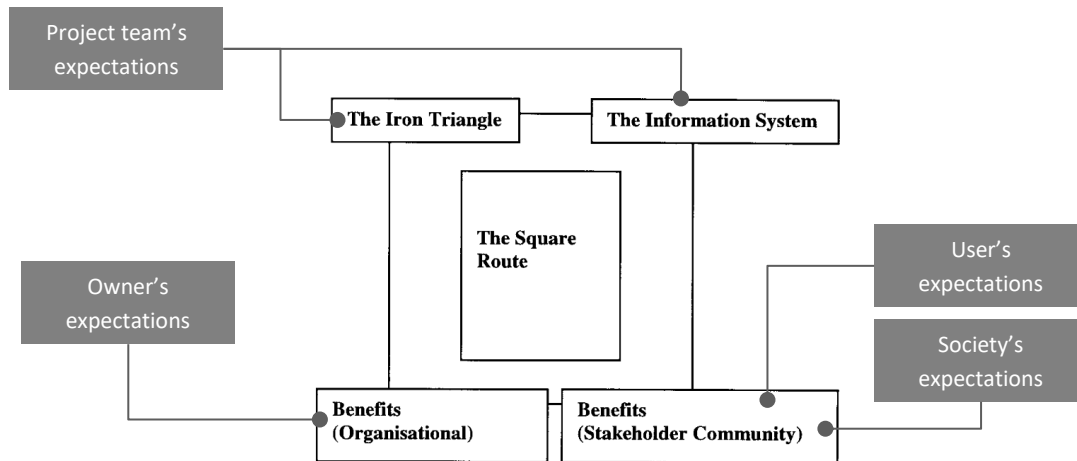


Figure 5-5: Link between The Square root framework, and stakeholder's expectations

[Source: Adapted from Atkinson (1999)]

(5) Multidimensional strategic Project Success:

Shenhar et al. (2001) developed a multidimensional strategic success framework (Figure 5-6) with four dimensions: meeting time, budget, and requirement goals (first dimension), benefits to the customer (second dimension), benefit to the performing organisation (third dimension), and preparing the future (fourth dimension). One relevant feature of this model is the incorporation of a time variable for the different dimensions of Project Success. The stakeholders' expectations can be linked to each dimension of Shenhar's model. Additionally, they can be framed in a certain period of time within the project lifecycle. The first dimension is in the interest of mostly the project team. The second dimension can be the expectations of the users. The third and fourth dimension should be in the main interest of the owner.

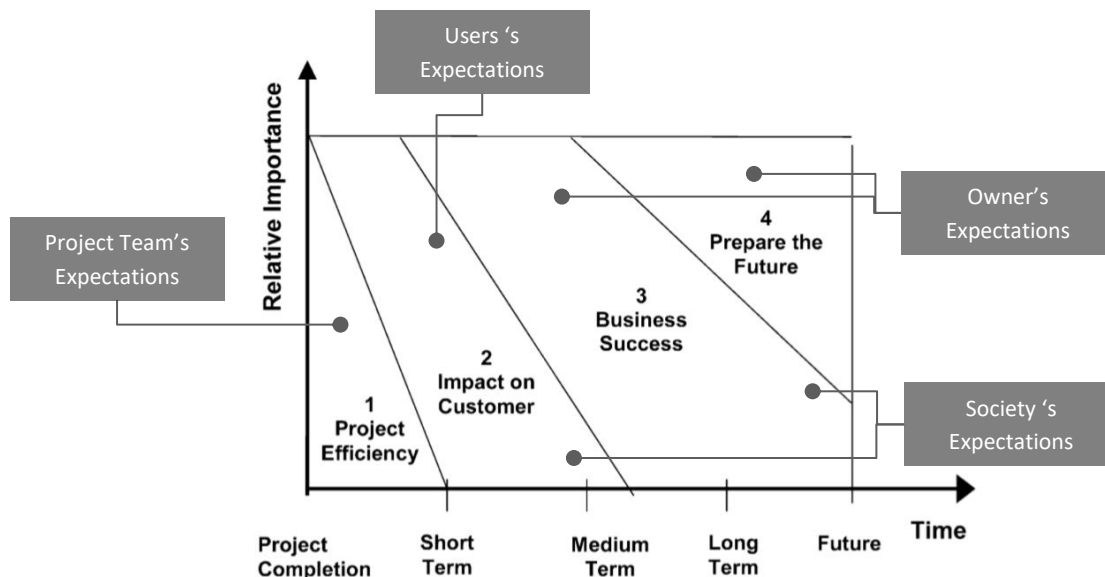


Figure 5-6: Link between multidimensional Strategic Project Success framework, and stakeholders' expectations

[Source: Adapted from Shenhar et al. (2001)]

(6) Performance framework for mega projects

Toor and Ogunlana (2010) recognise nine criteria to measure the performance of megaprojects (Figure 5-7). The expectations of the project team are related to time, budget, according to specification, safety, free from defect, efficiency, and minimum dispute and conflict. The expectations of the owner are related to safety, efficiency, and minimum dispute and conflict. Free from defect might be what the users can expect. One missing criterion is 'conform to stakeholders' expectations'. This cannot be assigned to any stakeholder, because it is not recognised by the analysis as a success criterion (discussed in section 5.3.3), but as a requirement for Project Success (in Figure 5-5 society and user are linked to 'conform to stakeholders' just as an example).

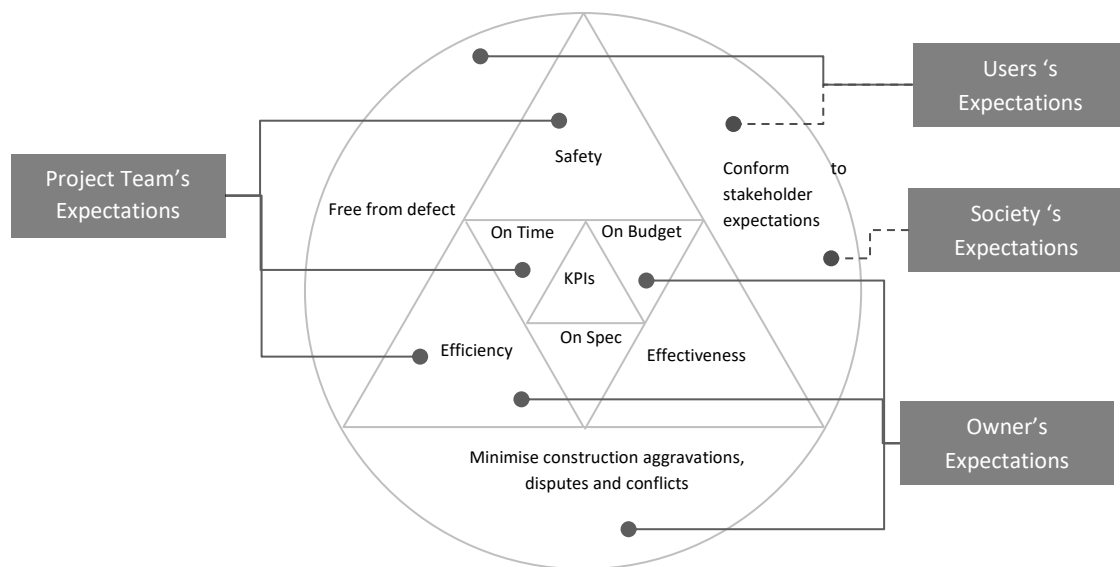


Figure 5-7: Stakeholders' expectations over the performance framework (example)

[Source: Adapted from Toor and Ogunlana (2010)]

(7) Project status model

Bronte-Stewart (2015) developed a project status model (Figure 5-8) that graphically shows the performance of a project using six criteria: timescale, cost and budget, scope of deliverables, benefits, quality, and risk. Similar to the emulation of the other frameworks in terms of stakeholders' expectations, each criterion could be linked to one or more stakeholders. Timescale, cost and budget, and scope of deliverables are the expectation of the project team. Benefits should be the expectations of the user and the owner. Quality is usually within the interest of the users. Finally, risk is something that concerns primarily the owner and the project team.

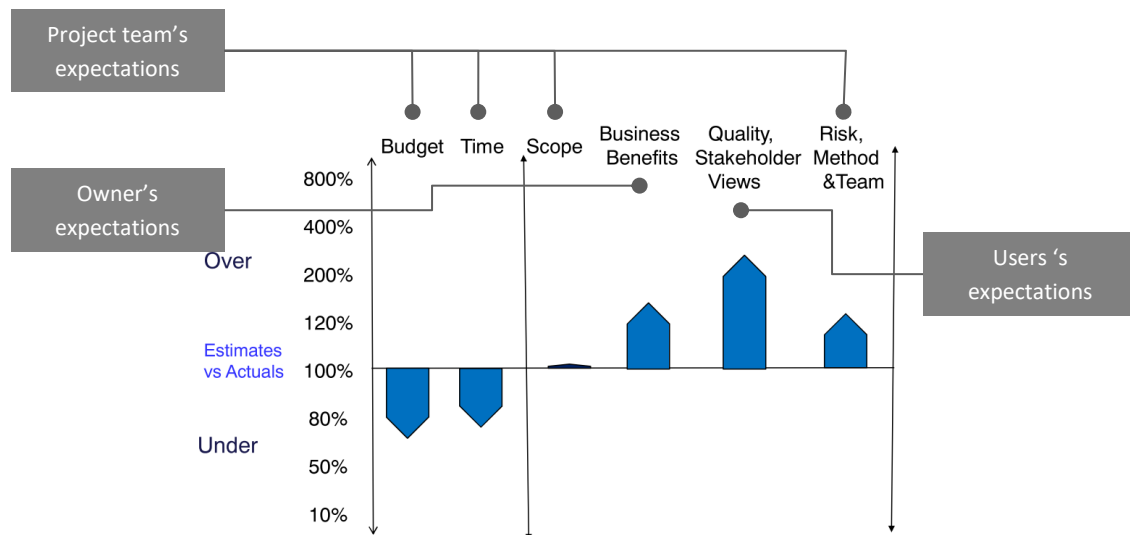


Figure 5-8: Link between Project Status Model and stakeholders' expectations

[Source: Adapted from Bronte-Stewart (2015)]

Overall, the analysis of the seven success frameworks suggests that the stakeholders' expectations are represented in all of them; i.e. the analysed frameworks could be specific cases of an overarching understanding of Project Success. Some authors such as Turner and Zolin (2012) and McLeod et al. (2012) use stakeholder expectations explicitly to understand Project Success, supporting, even more, the connection that it is argued in this sub-section. Moreover, the recent model of Davis (2018) studied the perception of Project Success across the multiple stakeholders to increase the shared understanding of factors and criteria, confirming the relevance of this new approach.

5.3.5. A successful project is different from a perfect project

Throughout this chapter, it has been stated that the definition of Project Success has reached a general agreement regarding the relevance of stakeholders to determine how successful is a project. The employed definition meeting stakeholder's expectations rigorously means that no stakeholder can complain or be unsatisfied with any component or perspective of a project in terms of the outcome or in terms of the process to deliver the outcome. However, it is possible to find projects that are considered successful, even though not all their expectations were met, even for a single stakeholder. Meeting all stakeholders' expectations is a very abstract and even utopian definition as in reality this is very unlikely to happen. This definition can be linked to a theoretical 'perfect project', which differs from a successful project. For example, the judgement of a user regarding a specific project can be that most of their expectations were met, some irrelevant expectations were not met, and the final judgement is that the project is successful; this is completely possible and logical in real-world projects. Moreover, the project manager (PM), as a stakeholder, has mostly its interests in time, budget and scope (Wateridge, 1998) and he can consider a project successful even if the three criteria were not met because at the end it is a judgement with many other considerations.

Based on the argument that a successful project is not the same as a perfect project, Figure 5-9 presents four project levels, based on how the stakeholder expectations are met. This conceptual characterisation synthesises the existing knowledge about Project Success and adds the new attribute that gives more precision and clarifies the situation of real projects when they are considered successful without meeting stakeholders' expectations.

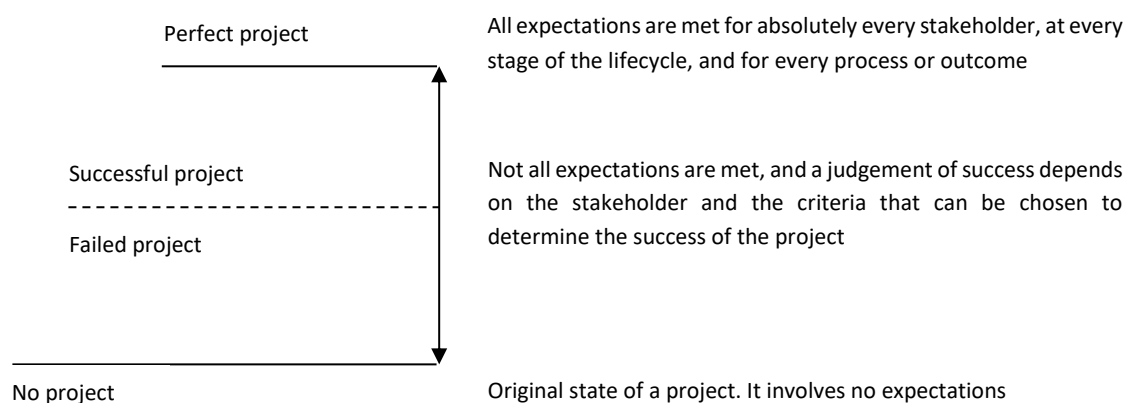


Figure 5-9: Characterisation of projects in terms of stakeholders' expectations

[Source: Author's]

5.3.6. Proposed refined Project Success definition

The analysis of the Project Success concept as the selected approach for developing a PPP ex-post impact evaluation approach was conducted in four subsections, namely: the existing connection between success criteria and stakeholder expectations (5.3.1), the subjectivity of success criteria (5.3.2), the misunderstanding of stakeholder satisfaction as a success criterion (5.3.3), stakeholder expectations embedded in existing success frameworks (5.3.4), and the difference between a successful and a perfect project (5.3.5). A refinement of the Project Success definition is proposed here: ***'meeting a substantial portion of stakeholders' expectations'***. Hence, a successful project will be a project in which not all, but a certain amount of the stakeholders' expectations are met. This definition differs from the current and general 'meeting stakeholders' expectations' but restates some correct attributes of it. The revised definition:

- Does not consider 'satisfaction of stakeholders' as a success criterion: Satisfaction is the achievement of success when stakeholders have met their expectations measured using success criteria. Hence, the satisfaction criterion is redundant.
- Makes any success criterion to be required by at least one stakeholder in order to be valid: Stakeholders are the trigger for the measurement of success, it is not possible to assess Project Success with criteria that come from outside the project; this will create an artificial answer of Project Success and probably the result will not be valid for the specific stakeholders of the assessed project.
- It is consistent with the existing success frameworks: all the reviewed success frameworks can be interpreted as structured ways of understanding stakeholder's expectations within a specific project.
- Specifies that not all possible expectations need to be met in order to achieve a successful project (section 5.3.5): there is a minimum of expectations that each stakeholder can accept for judging a project as successful. This means that a project does not necessarily need to meet all expectations to be judged as successful, but just the subset of expectations with levels of achievement for each success criterion. This is the main difference with the previous success definition, being more precise for describing what happens in real projects in which different stakeholders judge success even if not all their expectations were met.

Even though authors such as Turner and Zolin (2012), Davis (2018), and Shenhar et al. (2001) have discussed these issues implicitly (as presented accordingly throughout this section), the development of a refined Project Success definition that can address all of them together is a contribution of this

research. Further development needs to be done for addressing the question of how to measure stakeholder expectations; which is presented in the next section.

5.4. Stakeholders' expectation framework for Project Success

To support the refined Project Success definition, a framework for describing how stakeholder expectations can generate a Project Success judgement is presented in this section. Here, three main arguments are presented. First, '*meeting a substantial portion of stakeholders' expectations*' can be measured as the difference between the expectations and the perceptions of the stakeholders at a specific time. Second, the stakeholders' expectations are based on a set of success criteria that depend exclusively on the interest of each stakeholder. Finally, the stakeholder can change the significance of their expectations throughout the project lifecycle. This means that, for example, something one stakeholder finds very important now could be irrelevant for that same stakeholder in the future.

The three arguments of the framework are partially described in the literature. Xiong et al. (2015) employ the difference between stakeholders' expectations and perceptions as a measurement that can be conceptually linked to Project Success. Davis (2016) states that the same dimensions (criteria) of success are valued differently by each stakeholder. Finally, Shenhar et al. (2001) incorporate a 'time' variable in the analysis of Project Success.

The stakeholder expectation framework is a descriptive model of how the expectations of different stakeholders behave during the project lifecycle. The variables of the model are (1) success criteria, (2) relevance of success criteria for each stakeholder, (3) perceived performance in that success criteria, and (4) time. At a certain moment of time, every stakeholder has a Project Success judgement that is based on different success criteria, each one with a certain relevance and a perceived performance contrasting expectation and perceptions.

Figure 5-10 graphically represents an example of the behaviour of a single stakeholder's expectation over the project lifecycle. The area of each criterion represents the relevance at a specific period of time. The data for representing the profile was not taken from a real case and is just an example of how the relevance of each criterion can vary through the project lifecycle for one stakeholder.

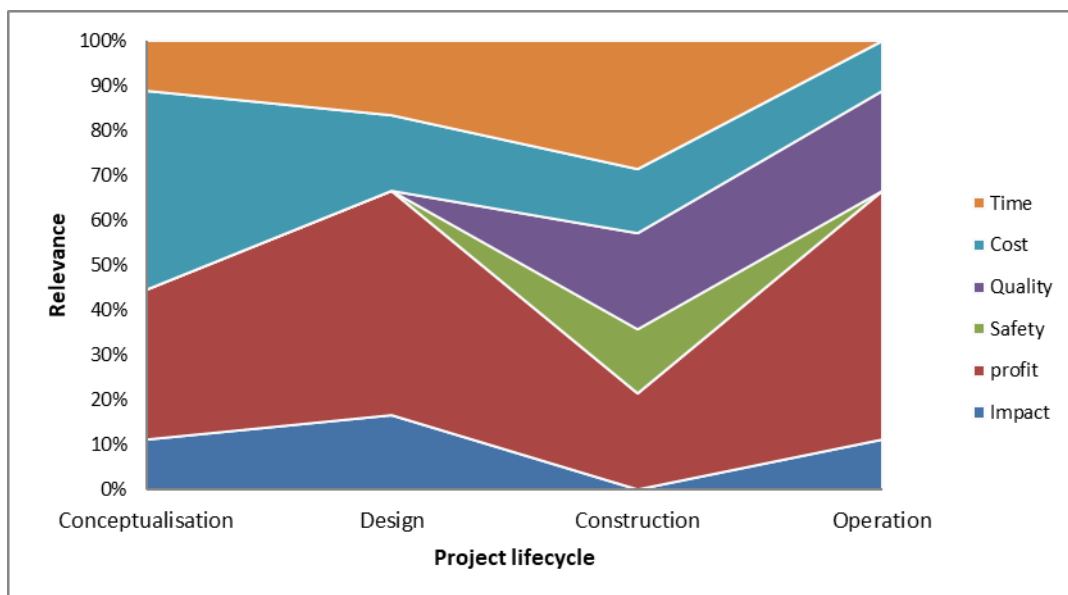


Figure 5-10: Example of one stakeholder's expectations profile in terms of criteria, project lifecycle, and relevance

[Source: Author's]

A mathematical equation describes the success judgement of each stakeholder during the project lifecycle. The perceived success of a stakeholder 's' is determined by a function that combines a 'j' criteria, during a 't' timeframe within the project lifecycle.

The expression is:

Equation 5-1: Project Success of stakeholder S

$$Project\ Success\ of\ stakeholder\ s = f\left(\int_0^t \sum_j criterion_{jt} weight_{jt}, \forall j, t\right)$$

[Source: Author's]

The combination of the different stakeholder's perceived success can determine the success of the project by employing the following expression:

Equation 5-2: Project Success expression

$$Project\ Success = f\left(\sum_s Project\ Success\ of\ Stakeholder_s\right)$$

[Source: Author's]

Function f represents any kind of operation that transforms the gathered information in terms of criteria, relevance and time into a possible measurement of Project Success. It is important to notice that the integer and the addition represent a form of aggregation and not exactly the mathematical formula. The next section proposes a conceptual function f that can be used to transform the gathered information into a Project Success judgement.

5.5. The Project Success judgement process for multiple perspectives

This section will develop the theory behind the judgement of project success. It will argue that a reasonable way to do it is to, first, understand the project as a system and then judge that system using 'stakeholder judgement principles', which are considered a common ground (a baseline) for the multiple stakeholders' views.

Three subsections develop the judgement process of Project Success when multiple perspectives exist. The first one (5.5.1) draws over the meaning of success in the Project Success concept. The second subsection (5.5.2) proposes the Alan Turing test approach to guide a Project Success judgement process. The third subsection (5.5.3) argues on the three principles that should be attributes of a successful project-system.

5.5.1. Meaning of success in Project Success concept

It is possible to agree that success is a perception. This means that currently there is no accepted way to measure success without considering someone's judgement. Even though this might be obvious, it is a starting point that must be addressed. The notion of a judgement is well recognised in the existing literature; Shenhar et al. (2001) state that the success of a project should be more accurately expressed as 'perceived success' because the ultimate validation of the concept is through a personal judgement, which can lead to a consensus or agreement of success. R. Müller and Turner (2007) describe success criteria as a set of satisfaction criteria which include end-user satisfaction, supplier satisfaction, team satisfaction, other stakeholder's satisfaction. This supports the argument that success is mainly a judgment, and that it depends on the perspective taken to judge it (Koops, Coman, Bosch-Rekveltdt, Hertogh, & Bakker, 2015).

The Oxford Dictionary defines success as a fact, an accomplishment, an achievement, an attainment (Waite, 2012); in other words, it is an attribute that can be achieved or not, with no other middle options. This makes the success concept to be considered a 'binary variable' meaning that the spectrum of variability has only two states: success and failure. In the literature, the discussion around Project Success considers this binary attribute in an implicit way (Atkinson, 1999; Baccarini, 1999; Pinto & Slevin, 1988; Shenhar et al., 2001; Williams, 2016). This is not surprising, as the word is used formally and informally with that spectrum of variability. This is consistent with the philosophical perspective of Wittgenstein, who states that the meaning of a word does not exist, but it does not matter because it is used anyway. This statement confirms that the nature of success is binary simply because it is employed in that way.

A scale of success with more than two states (success and failure) has not been found in the existing literature. Moreover, it is possible to state that even the word does not allow any kind of bigger spectrum. For example, '84% successful' is a useless assessment in any case that it can be imagined. A possible way to get an answer in those terms would be to identify criteria and allocate weights to get an overall score. If the overall score is within the success range (for example, 'above 50 points'), then the project is successful, if not, it is a failure. This kind of solution is considered incorrect as it creates just an artificial answer (de Wit, 1988) and it does not support the basic definition of success. For example, the Sydney Opera House in Australia would never be inside the success range, unless there is no weight for time, budget and scope, which is illogical. The Opera House was 10 years late and 1457% over budget, but it is rarely mentioned as not successful. It was declared a world heritage building by UNESCO, it has 8 million visitors each year, and it's undeniably the most well-known built landmark of Australia.

Recognising the difficulty to define success and the need for advancing the discourse in this topic; the search for new approaches to measure Project Success becomes relevant. A disruptive theory is needed to answer the question 'is this project successful?'

5.5.2. Application of Turing's test approach to the Project Success understanding

Understanding that success is a binary concept, the starting point is to state that the problem for measuring success is the *grey area* that exists between success and failure. It is difficult to establish the threshold when a project has just ended being a failure and started being successful, because the concept, by definition, is not defined with any kind of quantitative scale. Therefore, a way to simplify or reduce this *grey area* is required.

A similar problem can be recognised in the definition and measurement of intelligence. The Oxford dictionary defines intelligence as "the ability to acquire and apply knowledge and skills" (Stevenson, 2011). This means that the concept of intelligence also has a binary nature; this 'ability' may or may not exist. There are existing scales of intelligence (i.e. the intellectual coefficient scale), but usually, they cannot be applied outside the specific field in which they were created.

In the Computer Science (CS) field, the meaning of intelligence became essential with the rise of Artificial Intelligence; the ability of machines to think and solve problems. In 1950, Alan Turing (mathematician and philosopher) published the article 'Computing machinery and intelligence', which starts with the question 'can machines think?' this philosophical dilemma was disruptive, especially in 1950 where machines were just starting to be developed and the concept of Artificial Intelligence (AI) was not invented yet. To answer this question, Turing explained that 'thinking' and 'machine' were concepts that needed to be defined; and the search for the definition can be 'dangerous' and wrong for reaching an answer to the question. He stated that a new question was necessary to replace these

ambiguous words. The new question is: 'Are there any imaginary digital computer which would do well the imitation game?'

The imitation game involves one human having a text-based conversation with two entities. One of the entities is a human, and the other is a machine. The idea is that the first human needs to discover which one is the machine after 5 minutes of conversation. The theoretical implication of the game is that if a machine can imitate a human, then the machine must have the closest attribute to human intelligence. Consequently, machines should think. The Turing Test has been criticised widely, but it has also been highly influential in the philosophy of AI (Craven, 2014). In fact, since 1990 an annual competition of the Turing Test is held for chatterbot developers; there are no winners so far (Milton-Barker, 2016).

Insights of the Alan Turing's reasoning

Turing's approach for answering the question 'can machines think?' has four relevant insights that could inform the measurement of project success.

First, the question has a basic structure conformed by an 'entity' and an 'attribute' to be judged. The entity is the machine, and the attribute to be judged is thinking. Project success has the same structure; project is the entity and success is the attribute to be judged.

Second, the entity was redefined. He replaced machine' with 'imaginary digital computer', which consider also future improvements that technology can bring to 'machines'.

Third, the attribute (thinking) has a binary scale of variability and is simplified with a threshold question. He suggested that thinking was a binary human attribute; think or not think. Then, he transformed a question about the attribute (thinking) into a simpler question about a theoretical (but clearer) test, which can be called 'threshold question. he proposed an imitation game as a threshold question with no scales or scores; it is a yes/no game. The imitation game is the threshold and has the objective of reducing the grey area that exists between 'thinking' and 'not thinking' (Figure 5-11). With this game, it should be easier to judge if an imaginary digital computer is thinking or not avoiding a deep discussion (grey area). Project success is also understood as a concept with only two outcomes: success and failure, without middle levels, but there is no simpler theoretical test to answer the project success question.

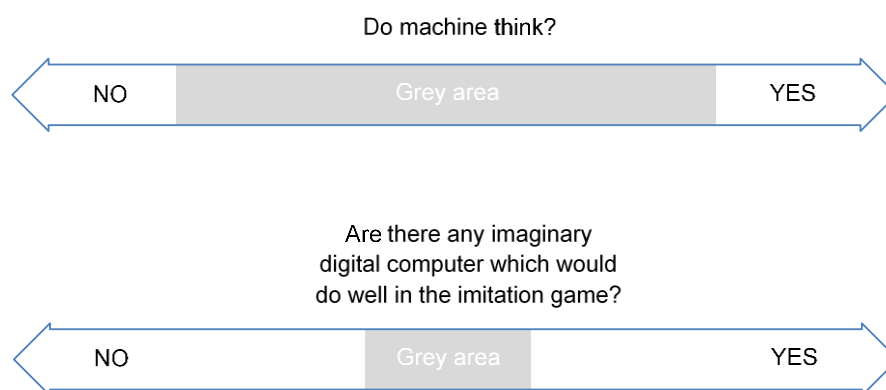


Figure 5-11: Representation of the Turing's approach for reducing the grey/discussion area in machine thinking dilemma

[Source: Author's]

Fourth, the nature of the threshold question (imitation game) is synthetical and not analytical. It means to comprise the whole concept and not parts of it. Synthetical comes from 'synthesis' meaning to concentrate and simplify a complex object or system, which is the opposite of 'analysis' (disaggregation). In graphic design, logos are called to be synthetic when they are able to transmit the concept behind brands in few and simple lines (Nike's logo representing a wing, YouTube's logo representing a screen, etc.). The imitation game is synthetical because it does not disaggregate thinking into, for example, mathematical thinking, conceptual thinking or any kind of 'sub-thinking'. The imitation game is a test that assumes it incorporates all the different aspects of thinking. Even though the critic to Turing's solution is the reliability of this assumption (to synthesise intelligence into communication behaviour), the concept of 'synthetic threshold question' still stands.

Application of Alan Turing's reasoning to Project Success

The four insights from Turing test can be applied to the question – is this project successful? – which are: (1) the question to answer has a structure with an entity (machine) and an attribute (thinking) to be judged, (2) the entity is redefined, (3) the attribute has a binary scale and is simplified with a threshold question, and (4) the threshold question is synthetical rather than analytical.

First, as stated previously, the question - is this project successful? - has the same structure as the question - Do machines think? Additionally, the judgement of project success needs to be done by someone, which makes no difference with the actual practice; there is an agreement that the way for measuring project success is through a judgement. Second, to redefine the entity 'project' to accurately represent the stakeholders' roles by which project success is now driven. Third, success needs to be simplified with a threshold question. A search for possible characteristics of successful projects is proposed for transforming success into a less ambiguous concept. Fourth, these characteristics of successful projects must be synthetical, meaning they could not represent a partial view of the project; i.e. 'on time' is not synthetical, as it provides only one component of the project leaving out other aspects.

Redefining the concept of 'project'

Most of the reviewed literature separates projects and stakeholders without conceptualising the two as a united system. In applying the Turing's approach, we need to redefine 'project' using a system view that understands project and stakeholders within the same project-system. This new concept can change the way we address the project success question by re-establishing a more system-based view.

The literature defines a project as an endeavour and the stakeholders as the people surrounding that endeavour. According to the PMBOK, a project is a 'temporary endeavour undertaken to create a unique product, service or result'. It has a beginning and an end time, and it has a defined scope and resources (Project Management Institute, 2013, p. 1) Even though it is not explicit in the literature, a project is conceptualised as different from the stakeholders, which are observing, contributing, and/or receiving the impacts of the project. The project manager, the owner, the client, the project team, the user, the community, etc. are all stakeholders that interact with the project; meaning the project refers only to the endeavour, not including the people surrounding that endeavour. Moreover, stakeholders are analysed by how they connect with the project (the endeavour) being classified with levels of power, influence, urgency, passiveness, activeness, dominance, demandingness, dependency, etc. (Johnson, Scholes, & Whittington, 2008, p. 158; McElroy & Mills, 2000; Mitchell, Agle, & Wood, 1997).

'Project' is defined in the PRINCE2 standard also as a 'temporary organisation that is created for the purpose of delivering one or more business products according to an agreed business case' (Hinde,

2012). Even though ‘temporary organisation’ includes the stakeholders and the endeavour, the PRINCE2 definition is restricted to the business context in which the organisation must follow a business case that was already agreed by a limited group of people. This definition is for a specific type of business project and is not applicable to all the other projects that do not comply with delivering a business product and following a business case.

We propose that ‘project’ should be understood with a system view to increase the ability to define project success and measure it more clearly by mimicking the Alan Turing’s approach. As projects are endeavours judged by multiple entities; a project can be understood as a closed-system in order to assess its success ‘as a system’. For the application of Turing’s approach, a project is here understood as “the set of different stakeholders implicitly connected to develop a temporary endeavour, which is undertaken to create a unique product, service or result”. This revised perspective of ‘project’ conceptually makes stakeholders part of the project and not just surrounding it (Figure 5-12). The connection between them is considered implicit because stakeholders (in practice) have different levels of involvement and some are only affected by the project without having much power or interest (Johnson et al., 2008). Nevertheless, it is a group of (some) stakeholders that triggers the endeavour, and it is the sum of all the stakeholders that allows the endeavour to reach its endpoint.

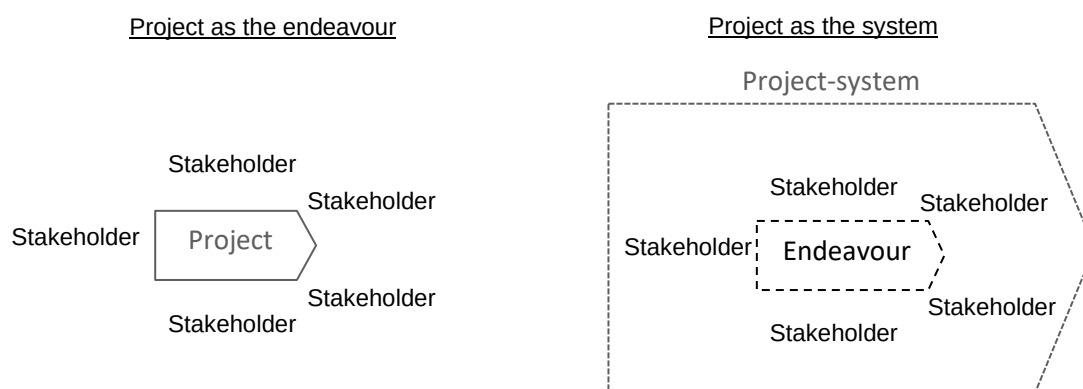


Figure 5-12: Project as an endeavour versus project as a system

[Source: Author’s]

Following the Turing methodology, and having redefined ‘project’, the original question - is this project successful? - can be rephrased as - is this project-system successful? Moreover, a better way to understand success in this case would be by rephrasing the question again as - is the project-system successfully working? Renger (2015) describes how system theory and system thinking can be applied to system evaluation. A system is an integrated composite of component with a need or objective. Here the project-system is understood as the result of having a project (the original definition) that inevitably needs to interact with the existing subsystems that surrounds it, generating a ‘project-system’. Therefore, this new system could work successfully or unsuccessfully, depending on what is considered successful working. This last point leads to the next conceptual development, as still the word ‘success’ needs to be transformed into a threshold question. Now, the search for threshold questions aims to simplify and clarify how a project-systems should work.

5.5.3. What successful projects do have in common

Turing’s approach provided a threshold question to reduce the grey area between thinking and not thinking. This section searches for threshold questions that can reduce the grey area between success and failure (Figure 5-13) using the new question - is this project-system successfully working? The threshold questions should be synthetical and not analytical. This means they cannot represent parts

of the project like budget, time, or quality; because it reduces success to the original solution of weighting, which ends in an artificial answer of success (de Wit, 1988).

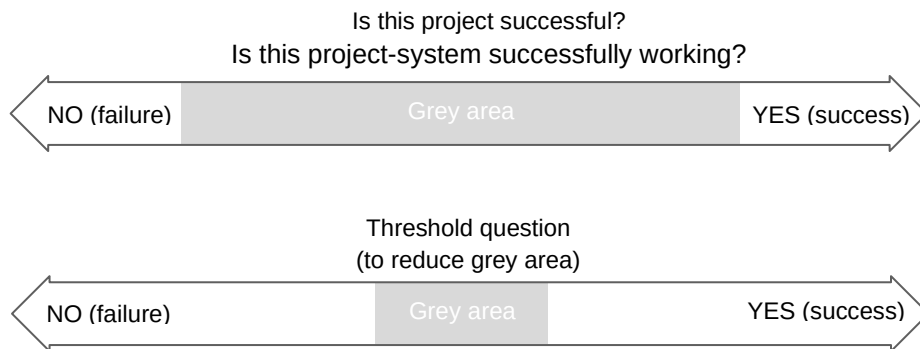


Figure 5-13: Representation of the project success dilemma mirroring Turing's approach

[Source: Author's]

Synthetical threshold questions to synthesise 'success' could be understood as common characteristics of successful projects-systems. The proposed definition of a project, which has the stakeholders in its core allows the exploration of the characteristics of successful projects differently from the existing approaches in the literature (e.g. iron triangle, success factors, performance measurement, etc.). These new characteristics focus on the successful functioning of a project-system and can be formulated as synthetical threshold questions.

The search for common characteristics of successfully working projects-systems was developed through literature review, critical thinking, and logic. The objective was to think of characteristics that could be applicable from the most basic and simple projects, to the most complex and large ones. The search for principles was based on 'what do the different stakeholder perspectives have in common?'; can these perspectives agree on something? As different stakeholders are using different criteria to judge success, what are the constraints of those judgments? What should those judgements need to comply with? What minimum rules should prevail? Philosophical approaches of Adam Smith, David Hume, Karl Marx, John Rawls, among others, were examined while reviewing the existing literature of project success. Each of the characteristics is explained as a principle for judging projects and they are discussed in detail.

As a starting point for further research in this area, this article explores three characteristics that are present or 'should be' present in successful project-systems. (1) Successful projects have a general willingness to be repeated; (2) successful projects should not have evident excessive negative outcomes impacting any of the stakeholders (minimum fairness); and (3) there should not be an evident better alternative. These three characteristics presented in this work are only the starting point in this strand of research. There might be other characteristics for exploration by future research.

First attribute: Stakeholders should be willing to hypothetically repeat successful projects

The first characteristic of a successful project is related to the judgement of the stakeholders. It is supported on the traditional satisfaction survey. Are the stakeholders satisfied with the project? This question (very common performance measurement) lacks precision and rational thinking because it is a judgement that does not involve subsequent decision-making. For example, the judgement about something good or bad can be considered vague, because the course of action is not set. A deeper analysis of the good/bad dilemma would be, is it good or bad for whom? good or bad compared with what? what does it mean that something is bad? that it will be changed? it will be changed for

something better? Will be improved? How better is the new thing? Is there any cost of adapting to the new thing? etc. The same analysis is done for success; the decision behind the success judgement is not clear.

The willingness to repeat the same action knowing their result is a trigger that captures a robust judgement about that action. It involves a decision-making (repeating), it keeps an accurate version of the situation that is being judged (exact action with exact results), and it forces the judgement to consider every aspect of the situation that is been judged (the black box includes all possible criteria, precise weighing process, and any possible prioritisation or exclusion).

The judgement about repetition is supported by two bodies of knowledge, self-interest, and social behaviour. Self-interest (Galles, 2015) pleads for the assumption that an individual will always seek for their best interest, then, their judgement will always contain a cost-benefit ratio that intrinsically solves the good/bad dilemma. The willingness repetition is a dilemma that is consistent with the self-interest principle; it clarifies that a stakeholder must consider its own interest and decides if it's a good or bad decision to repeat the same project. Within the self-interest, other criteria and individual values are intrinsically part of the decision process, such as fairness, solidarity, or legal compliance, which affect the decision about repeating the project.

The case that most of the stakeholders can be theoretically willing to repeat the project in the same general terms that was originally developed represents in a much clearer judgement of Project Success. Traditional criteria such as cost and time are related to 'the endeavour'; whereas this principle is related to 'the union of stakeholders'. A stakeholder needs to address their own set of criteria like time, quality, safety, participation, etc. and then make a judgement regarding the willingness for repetition. In other words, the principle is synthetical, comprising the whole project and not a part of it.

The willingness for repetition principle can be inferred from the existing literature. For example, R. Müller and Turner (2007) identified 'reoccurring business' as a success criterion, mainly for private projects. To meet this criterion, the client must assess all the technical criteria of their interest, such as time, quality and cost, and then judge the continuity or the breakup of the relationship with the service provider. This is consistent with common industry notion of 'if the client is happy, then the project or the business is successful' (from the perspective of the contractor). However, a clearer question to replace the 'happy client' would be - is the client willing to recommend the contractor for another job? Or, does the client want to continue doing business with the same contractor? These questions can be interpreted as the willingness for repetition in a particular stakeholder relationship (contractor-client). Consequently, do the workers want to continue in the same company? Or, have users selected the same service provider in the future? All specific situations that can be unified under a broader willingness for repetition.

The principle creates a threshold for judging a project as successful even though it did not fulfil all the expectations of the stakeholders. A theoretical threshold exists for projects that are not perfect, but still successful. Using the famous example of the Sydney Opera House, which was over budget and was delivered with a big delay; it would be rare to think that any stakeholder will be against building the Opera House again, even after knowing all the problems that the project will have in terms on time and budget.

Applying this characteristic to the Turing's Test approach, the threshold question to guide a success judgement would be 'Are the stakeholders willing to repeat the same project, knowing that they will receive the same outcomes, and considering that they have only the options that were defined at the start of the project (including the 'do nothing' option)? The willingness of the stakeholders to repeat the project is argued to be a clearer way to judge successful projects.

Second attribute: successful projects should not have evident excessive negative outcomes impacting any of the stakeholders (minimum fairness)

Having no excessive negative outcomes for any stakeholder is a requirement that every successful project should pursue or comply. With the proposed project definition that focuses on the stakeholders, the 'no excessive negative outcomes' is considered a minimum fairness that should exist in any system, network, project, or organisation. All stakeholders (that enter voluntarily or involuntarily into the project) should be impacted positively or negatively 'within a fair norm'. This fair norm is further discussed.

To describe and understand the meaning of negative outcomes 'within a fair norm' the philosophical perspective of John Rawls's 1971 'The Theory of Justice' is employed. He attempts to solve the problem of the distributive justice by presenting a theoretical situation in which the members of a community need to decide the rules of their society without knowing their future place in that society; what he called 'the veil of ignorance'. The set of rules created in what he refers to the 'original position' of the members, allows (according to him) the most fair consensus of society, assuming that no one can be influenced by their place in the future society because they do not know which is going to be (Rawls, 1999).

Rawls's theoretical experiment could be applied for determining what is 'within the norm' of fairness, by judging if the stakeholders are willing to accept any of the impacts that other stakeholders are receiving. A specific impact for a stakeholder is considered 'within the norm' if the other stakeholders can accept that impact for themselves; this is related to what is commonly known as 'skin on game'. Rawls's theory applied to the analysis of successful projects establishes a guideline to assess a successful project in term of the results for each stakeholder. Fairness is understood as the consensus that all the involved stakeholders (society in Rawls theory) could have agreed (with the veil of ignorance) for the outcomes of the project that affects each of them.

The implication of this principle helps to clarify the threshold in the judgement of Project Success for any project in which a stakeholder got an evident and relevant negative outcome. In the case of a mega infrastructure project, if the private company got bankrupt, then, no one would like to become the private company on the project assuming there is an opportunity to switch roles. This principle also helps to analyse an evident (out of the norm) negative result for the company's employees, or if the client cannot recover the investment, or if the community ends with an excessive environmental problem. However, the answer to this principle, if negative, should not be applied when there are reasonable arguments for the unfair result (e.g. bad financial management from the company that bankrupted) which are not a direct effect of participating in the project. As it was earlier put forward, these are principles and not rules or algorithms; they help to guide a discussion, rather than supporting a straight answer.

Evident excessive negative outcomes for at least one stakeholder is a principle to judge Project Success. The threshold question in the context of the Turing's test approach becomes: 'Are the stakeholders willing to exchange position with any other stakeholder of the project, considering that the outcomes are known for each of them? If the stakeholders have no problem in 'receiving' the impacts of any other stakeholder, it is possible to infer that the outcomes are 'within the norm' of minimum fairness. Similar to the repetition principle, the 'no excessive negative outcomes' principle helps to clarify the vague success concept and, consequently, it is considered a characteristic that successful projects should have.

Third attribute: If stakeholders are not willing to repeat a project and there is no minimum fairness; then, the project should be the best available options in order to be considered successful

A third characteristic that successful projects have in common is an exception (or a special case) of the two previous principles. 'Willingness for repetitions' and 'fairness between stakeholders' assumes that the project outcomes are positive and that the project was intended for the benefit of someone or for the improvement of a situation. However, many projects have the goal of reducing negative effects or creating the minimum negative results as the best possible outcome. Examples are the isolation of humans for epidemic and health issues, battle machinery and infrastructure, or environmental constraints due to natural disasters. In none of these examples the stakeholders are willing to repeat them, and the fairness between stakeholders probably does not exist.

To address these negative outcome cases, a third condition/principle is established. Successful projects should be an evident alternative option at the decision-making phase. This principle provides strong support for the judgement of successful projects in cases where there is no willingness for repetition and no fairness between stakeholders. It supports all projects in which there is no evident possibility to have a better option; many times, 'the less bad option is the best option'. To be the best available option is not a requirement for all projects, as 'not-best-option projects' can still be successful without the need to be compared with the original alternatives; a successful project does not mean perfect or optimal.

The application of the Turing test approach for this principle results in the threshold question: 'is the project one of the best available options considering the original alternatives developed and discussed at the conceptualisation stage?'

Framing the three principles

As stated before, the principles are guidelines and not rigid formulas, thus, they can guide a logical argumentation for judging a project-system. For example, the following argumentations can be used. 'Fairness between stakeholders' and 'willingness for repetition' principles could create a frame for the success judgement. The willingness for repetition principle could be the upper limit of a theoretical success scale, being a convincing argument for judging a project as successful. Conversely, the 'fairness between stakeholders' principle could be the lower limit in the success scale; an unfair project is a strong argument for judging it as a failure even if there is no reason that explains the unfairness. Between both principles, the judgement discussion is narrowed, allowing a clearer success judgement.

There are many possible scenarios in which the principles can interact for supporting a discussion and guiding a judgment of project success. Moreover, as mentioned before, there is no constrain for proposing in the future new principles that can act as threshold questions when describing a 'successfully working project-system'. For example, one way of applying them could be as follow:

- If within a project, there is a general willingness for repetition from all the involved stakeholders, this is a strong argument for judging the project-system as successfully working.
- If there is no general willingness for repetition, then, it is necessary to, at least, fulfil the requirement that is not possible to declare any evident excessive negative outcome to any stakeholder. Not meeting this principle is a strong argument towards judging a deficiently working project-system (failure).
- If no stakeholder wants to repeat the project and there is unmistakable evidence of excessive negative outcomes, then the question becomes 'was there a better alternative to do in the first place?' would have other project alternatives generated positive answers to the two previous threshold questions? The non-existence of a better alternative is a strong argument for a positive judgement of a project-system.

This framework is proposed under the assumption that stakeholders (as a group) can recognise the ‘reality’ in which the project is being developed. It is assumed that, for example, environmental problems and long-term risks should be part of at least one stakeholder’s interest. Conversely, if a project is recognised as successful by all stakeholders without considering (for example) long environmental impacts, then it would be a theoretical tragedy of the commons. This shows a potential limitation of the framework for analysing projects (or endeavours) that are beyond the traditional project management field, e.g. world wars, climate change, indigenous rights, etc.

Employing the Turing test approach, a new perspective for understanding Project Success has been created. To answer the question ‘Is this project successful?’ a combination of new questions are to be answered for reaching a general judgement. Considering that there are as many perspectives of the same project as the number of stakeholders involved, the new questions support that diversity by clarifying the concept of success. Three principles are proposed for judging Project Success and presented as questions: ‘Are the stakeholders willing to repeat the same project, knowing that they will receive the same outcomes, and considering that they have the same options that defined the project (like do nothing option)?’ ‘Is the project one of the best available options considering the original alternatives that were developed and discussed at the conceptualisation stage?’ Similar to Turing’s original approach, this set of questions is meant to guide a judgement about success, rather than making a binary discrete approach about success. It’s a step forward towards the understanding of the success concept and not a definitive answer to the original problem.

The theoretical development for characterising a successful project is the final step for developing the conceptual framework ‘Project Success Evaluation Pyramid Model’ that is presented in the next section.

5.6. Project Success Evaluation Pyramid Model (PSEPM)

This section presents the results of the theoretical development described from sections 5.1 to 5.5 which is a proposed conceptual framework called ‘Project Success Evaluation Pyramid Model’ (PSEPM). It is essentially a model that identifies a big spectrum of success criteria from each of the involved stakeholders and then employs a set of principles to assess the validity of all those criteria. For each stakeholder, a unique perspective is generated that contains the relevant information that supports the stakeholder’s judgement towards the project. To better understand the multiple stakeholder perspectives, ‘project’ was redefined as the system of stakeholders that are interacting within an endeavour. Consequently, the aim is to judge if a project-system is successfully working, rather than to judge if a project is successful. The outcome of the framework is the process for reaching a judgment rather than the measurement of Project Success.

The model (Figure 5-14) analyses project-system success using three components: (1) success criteria, (2) stakeholders success judgment, and (3) stakeholder judgment principles. Each component/level of the pyramid informs and supports the next level (from bottom to top) in order to achieve a theoretical judgment of Project Success. The model illustrates that a set of success criteria (SC) informs a single stakeholder success judgment (SSJ). Then, the combination of all the stakeholder success judgments is tested against a single ‘stakeholder judgment principle’ (SJP). Finally, all the stakeholder judgment principles inform a judgment of Project Success (PS). Each component is further explained.

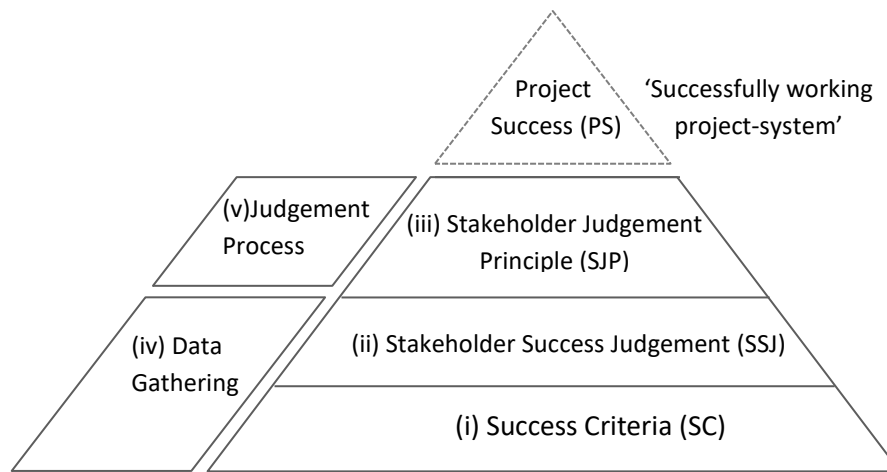


Figure 5-14: Project Success Evaluation Pyramid Model (PSEPM)

[Source: Author's]

i) Success criteria (SC)

The first level of the pyramid represents all the existing success criteria (SC) that can be employed to measure Project Success, including time, cost, quality, safety, profit, environmental sustainability, Value for Money, etc. A specific subset of this pool of SC is what stakeholders use to inform a success judgment. The model establishes that stakeholders must recognise their single perspective and be accountable for the way they judge the project. In other words, they must provide a set of criteria that will represent their perspective.

Existing literature suggests that success depends on each stakeholder (Baccarini, 1999; Bourne & Walker, 2005; Davis, 2013; Müller & Jugdev, 2006; Shenhar et al., 2001; Turner, 2009; Westerveld, 2003), which means that every stakeholder has a different success judgment. According to de Wit (1988), this is the main complexity for addressing Project Success as it represents a different thing to everybody. Section 5.4 discussed how the Project Success concept is linked to stakeholder expectations, and section 5.1 argued about the problem of having a single perspective evaluation approach when there are multiple stakeholders with their own perspectives.

ii) Stakeholder success judgement (SSJ)

The second level of the pyramid represents the judgement that every stakeholder must generate employing their unique criteria. They must declare if the project is meeting their expectations in every criterion. Based on Xiong et al. (2015), a gap in expectation can be expressed as the difference between what stakeholders expect and what they perceive in a specific success criterion. Every gap in expectation must be expressed in terms of quantitative and qualitative information to understand the gap and assess its relevance and validity. A 'j' gap in expectations is defined as follows:

Equation 5-3: A gap defined as the difference between a perception and an expectation

$$Gap_j = (Stakeholder\ perception_j - Stakeholder\ expectation_j)$$

[Source: Author's]

To ensure the validity of stakeholder expectations, one limitation is proposed: The addition of all the expectations should not make the project unfeasible. This means that a project should, at least theoretically, be capable to meet all the different expectations of the stakeholders. For example, if a hospital is expected to have the newest technology (a health authority perspective), but it is also

expected to be within a maximum budget (national budget authority perspective) that makes it the technology impossible to buy, then the project is not 'feasible'. Consequently, Project Success will never be reached even theoretically speaking. The overall purpose of the project (i.e. business case, contract, or approved scope) is the source for validating the expectations. This constraint is expressed in the following equation for 'n' stakeholder with 'j' expectations:

Equation 5-4: The sum of all stakeholder expectations should allow a feasible project

$$\sum_1^n \sum_1^j \text{stakeholder expectation}_{nj} \leq \text{feasible project}$$

[Source: Author's]

The judgement of the stakeholder about the success of the project is called a 'perspective', which is supported on the criteria that contains specific expectations, and traceable perceptions. This judgement is no more than what it has been conceptually stated in the literature; a stakeholder has its own judgement of success, and that judgement is based on the expectations over certain success criteria. Section 5.4 explained this stakeholder framework with a descriptive model of how expectations behave during the project lifecycle.

iii) Stakeholder judgment principles (SJP)

The new success question 'is this project-system successfully working?' needs to be answered as an overall statement informed by the different stakeholder views. For doing so, successfully working project-systems should be defined with certain characteristics in which the different stakeholder views could merge into. This framework proposed the concept 'stakeholder judgement principles' as overarching attributes for guiding the analysis of a successfully project-system. These principles are the key characteristics that project-systems should have in order to be considered successful from the perspective of multiple stakeholders. They have a 'high-level of abstraction' and they differ from the traditional success criteria such as cost, time, and quality; because they address the perspective of every stakeholder, rather than the technical measurements. This research has developed three judgement principles for guiding the evaluation of a project-system:

- Willingness to repeat the project: Are the stakeholders willing to repeat the same project and receive the same outcomes assuming that they can only consider the original project alternatives? This willingness to repeat the project supports the assumption that the stakeholder can accept the gaps in expectations when they are compared with the benefits they received. It is an active principle that involves a decision-making (to repeat or not repeat), different from a passive action such as level of satisfaction that has no real impact embedded.
- No excessive negative outcome: does any stakeholder received an evident negative outcome that could have been avoided? This requirement provides a minimum of 'fairness' in which an excessive and evident negative outcome cannot be accepted for considering a project as successful.
- No evident alternative option: In the case that the project does not want to be repeated and it has evident negative outcomes; do the stakeholders think that the project is still the best option from the original alternatives that were available at the conceptualisation of the project (including the 'do nothing' option)? Not having a better alternative is a strong argument for judging a project as successful. It helps to modify (calibrate) the expectations with the real available options and the estimated results.

Three limitations need to be considered for the principles: (1) they are guidelines to support a judgment and not formulas, so they do not fit perfectly in every case; (2) they are generic and should be adapted for each project and stakeholder; and (3) it is also possible to add other principles (with a

similar level of abstraction) for specific cases. For the PSEPM, the concept of 'stakeholder judgement principle' (SJP) is the relevant component, rather the specific principles developed by this study. The underlying theory for each of the proposed principles was developed in section 5.5.

iv) Data gathering and (v) judgment process

The goal of an infrastructure PPP project is to provide a public service using private funding. Consequently, program evaluation from public policy can be employed to assess a PPP project.

As it was described in the literature review chapter, program evaluation, as a discipline, establishes that 'evaluating' involves two processes: a measuring process (gathering data), and a decision-making process (judgement) (Langbein & Felbinger, 2006). The PSEPM is consistent with the program evaluation theory including the two steps that operationalise the pyramid. The first step represents all the information that needs to be gathered from each of the multiple perspectives; which involves the two first levels of the pyramid. The second step represents the judgement process that analyses the information and generates a decision; this refers to the third level of the pyramid with the judgement principles.

vi) PSEPM outcome

The outcome for determining the impact of the PPP on stakeholders is a judgment employing the stakeholder judgment principles (SJP). The judgement is based on the set of gaps in expectations that were gathered and assessed from the multiple stakeholders. This outcome specifies the impacts on each stakeholder; and the connection to the overall view of the project.

The contribution of the conceptual framework is to solve the multiplicity of views that exist within a Project Success judgement. The described framework (1) understood a project as a system, (2) structured the way the stakeholders provide their judgements, and (3) proposed general principles for judging successfully working project-systems. This theoretical understanding of Project Success is considered an independent outcome of this study, contributing to the theoretical understanding of Project Success. The conceptual framework will be the theoretical support to develop a PPP ex-post impact evaluation method, which is the purpose of this research, but this chapter could be applied to any other endeavour that needs to deal with successful projects e.g. public policy, innovation projects, etc. For PPPs, the Project Success Evaluation Pyramid Model is useful mainly because it recognises the interaction of the stakeholders in a complex project.

5.7. Summary

This chapter describes the conceptual development 'Project Success Evaluation Pyramid Model' (PSEPM), which is the first deliverable in the design process and it answers to the first research objective: *Develop a conceptual framework that informs a PPP evaluation.*

The theoretical development of this chapter has five components.

1. An argument is presented that for assessing PPP projects multiple-perspective evaluation approaches are preferred as these kinds of projects require the multiplicity of views that come from all their stakeholders.
2. From the literature review, the Project Success concept is selected as the evaluation approach that can support the conceptual framework, due to its explicit involvement of all stakeholders in the judgment process.
3. The Project Success concept is refined, based on five issues that are considered conceptual ambiguities and inconsistencies: (a) success criteria are essentially stakeholder expectations; (b) success criteria are subjective; (c) stakeholder satisfaction is not a success criterion; (d) existing success frameworks can be represented as sets of stakeholders' expectations; and (e)

a successful project is different from a perfect project. A new project definition is proposed as meeting a substantial portion of stakeholder's expectations.

4. A stakeholder expectation framework is proposed employing explicitly the gap between expectations and perceptions for every success criterion.
5. The philosophical approach of the Alan Turing's test is employed to understand the concept of Project Success as a judgement, guided by three principles: (1) willingness to repeat the project, (2) no excessive negative outcomes, and (3) no evident better alternative.

The result of the theoretical development is the conceptual framework 'Project Success Evaluation Pyramid Model', which is essentially a procedure that identifies a big spectrum of success criteria from each of the involved stakeholders and then employs a set of principles to assess the validity of all those criteria. It has four components: (1) the success criteria, (2) the stakeholder judgements, (3) the stakeholder judgement principles, and (4) the theoretical Project Success. The PSEPM is the basis for developing the PPP ex-post impact evaluation method, which is the second deliverable of the research.

The next chapter will present the development of the prototype evaluation method which is based on (1) the expert's feedback about the context of PPP ex-post evaluation, (2) expert's feedback regarding the proposed conceptual framework (PSEPM), and (3) the prototyping exercises of the stakeholder workshops.

CHAPTER 6

PRESENTING THE CONCEPTUAL FRAMEWORK TO EXPERTS

6.0. Introduction

This chapter presents the feedback process of the conceptual framework, and the development of the prototype evaluation method. The conceptual framework 'Project Success Evaluation Pyramid Model' was presented to 13 PPP experts and practitioners in Chile and Australia. They provided feedback about the general evaluation context of PPPs and about the developed conceptual framework. This chapter also includes the prototyping exercises for the development of the 'prototype evaluation method'.

The chapter is structured as follow. Section 6.1 describes the process for developing the prototype evaluation method. Section 6.2 groups the feedback from the experts regarding the evaluation context of PPPs. Section 6.3 describes the experts' feedback about the conceptual framework. Finally, section 6.4 presents the observations of the prototyping exercises.

Part of this chapter has been published in the following article: Oliveros-Romero and Aibinu (2019).

6.1. Development of prototype evaluation method

This section presents the development of the prototype evaluation method. It explains how the feedback from the experts was collected and then analysed, how were the prototyping exercises conducted, and how these two processes led to the development of the prototype evaluation method.

Interviews with PPP experts and practitioners were conducted, and they had two objectives: (1) explore and discuss the challenges for designing a PPP ex-post impact evaluation, and (2) gather feedback from the presentation of the PSEPM. The data collected from interviews was transcribed verbatim (word for word) and a thematic analysis was conducted employing MS Excel. The analysis generated 170 pieces of information coded with more than 40 key concepts. Each piece of information is an idea presented during the interviews in the form of a sentence or paragraph from the interview transcription. The pieces of information are grouped into themes. For design purposes, each theme is considered a refinement insight that contributes to the development of the evaluation method. Three actions have been created for categorising the refinement insights: acknowledgement, declaration, and inclusion.

- 'Acknowledgement' means that the theme do not generate a specific change in the evaluation method and they are not included as part of the design. However, they could support other parts of the research, such as the relevance of the topic, the significance of the findings, the further research development and/or the method applicability.
- 'Declaration' means the idea in the theme needs to be presented as part of the method's documentation. However, it does not inform a modification or an addition to the design.
- 'Inclusion' refers to any idea or development that needs to be included in the evaluation method procedure.

71 pieces of information were categorised as evaluation context (EC) and grouped into 9 themes. Evaluation context refers to the first objective of the interviews: explore and discuss the challenges for designing a PPP ex-post impact evaluation. PPP experts and practitioners stated that conducting an ex-post evaluation on PPP projects was not common and the topic was usually not present in the projects they have worked on. Specific ex-post evaluation methods could not be recalled, except for (1) financial approaches, (2) Gateway Review method of the Victorian Government, Australia, and (3) short and long-term ex-post evaluation of the Social Development Ministry in Chile. However, none of the interviewees was confident about the effectiveness of those methods for ex-post evaluating

thoroughly the impact of PPP projects. Nevertheless, they all agreed on the relevance of conducting ex-post evaluations to PPP projects and that it should be pursued in the future. Interviewee F stated, *“There are a few evaluation tools out there that I have seen for helping lessons learnt so I cannot think of anything out there”*. Table 6-1 presents the summary of the 9 evaluation context (EC) themes.

Table 6-1: Design inputs from evaluation context themes

Evaluation Context (EC)			
ID	Issue	Description	Action
EC1	Contribution and purpose of the evaluation	General agreement over the relevance of PPP ex-post evaluation. Four main purposes: transparency and accountability, PPP legitimisation, industry learning, and government agency learning	Include
EC2	Value for Money, urgency and importance	Value for Money is being questioned. Urgency and importance are presented as alternative decision-maker drivers	Acknowledge
EC3	the role of ex-ante evaluation in ex-post evaluation	The comparison before-after legitimises the evaluation process. Therefore, the ex-ante evaluation should be robust and considered in the PPP ex-post evaluation	Acknowledge
EC4	Separation between procurement method, project, and public policy	Separating the effect of the public policy, the project, and the procurement method should help to assess correctly the impact of the PPP	Include
EC5	The ideological component of PPPs	The political bias can alter PPP evaluations, similarly as it influences the acceptance and rejection of PPP projects	Acknowledge
EC6	Stakeholder expectations	The change of stakeholder's expectations thought time should be considered in a PPP ex-post evaluation	Acknowledge
EC7	Accountability	Final accountability should remain in the public sector, but the long lifecycle duration makes accountability difficult to track	Acknowledge
EC8	Type of PPP from an exposure perspective	Social infrastructure PPP project are more exposed to the general public, that has less engagement and less technical knowledge to assess the project. This can lead to judgements that over rely on perceptive single experiences	Acknowledge
EC9	The misconception of PPPs as a donation	PPPs are financially a loan that needs to be paid by the public sector. This misunderstanding can lead to assess PPPs wrongly	Acknowledge

[Source: Author's]

The remaining 99 pieces of information were grouped into 6 themes about method feedback (MF), in reference to the second objective of the interviews: gather feedback from the presentation of the PSEPM. Experts recognised the conceptual framework's contribution and its potential applicability. However, some issues were highlighted for its practical application such as the inclusion of ex-ante evaluation, the stakeholder bias treatment, and the acknowledgement of the stakeholders' purpose towards the project. Table 6-2 presents the summary of the 6 themes.

Table 6-2: Design inputs from method feedback themes

Method Feedback (MF)			
ID	Issue	Description	Action
MF1	Method contribution	Experts have mostly a positive impression of the core evaluation idea, but the three principles to address a Project Success judgement generated scepticism	Acknowledge
MF2	Method applicability	There are few incentives to apply PPP ex-post evaluations from the public and private sector. Independent organisation should be interested in using this kind of method, and to include it in a regulation was also suggested	Acknowledge
MF3	Link with ex-ante evaluation	Ex-ante evaluation should be robust and well linked to the PPP ex-post evaluation to avoid any bias related to the evaluation time and context (media coverage, political elections, etc.)	Include
MF4	Treatment of stakeholder expectations	The evaluation method should consider that stakeholder expectations (1) change through time, (2) have different timeframes, and (3) can be official and unofficial (hidden)	Include

MF5	The bias of the evaluator	As the method involves judgements and mediation of the evaluator, the bias risk exists and should be minimised and declared	Declare
MF6	Stakeholder identification and purpose	The stakeholder identification process should consider (1) separating the different public sector entities, (2) including stakeholders' strategic and long-term objectives, and (3) stakeholders' motives to support the criteria validation	Include

[Source: Author's]

This chapter also includes the refinement insights (lessons learned) taken from the prototyping exercises (PE) of the stakeholder workshop, which is the main component for gathering the data for the PPP ex-post impact evaluation method. Lessons learned included the workshop structure, amount of information to provide to the stakeholder, workshop length estimation, and questions testings. This section has a different data analysis method, as it employs a Design Science Functional (black box) testing, in which the researcher search for problems in the artefact while is being used in a real situation (Hevner et al., 2004, p. 86). Table 6-3 summaries the observations.

Table 6-3: Design inputs from prototyping exercise observations

Prototyping Exercises (PE)			
ID	Issue	Description	Action
PE1	Bias between theoretical stakeholder and its representative	As stakeholders are groups of people, they have different views. Thus, a stakeholder representative can be biased to show that diversity. This should be acknowledged and mitigates	Include
PE2	The workshop requires to be guided rather than structured	The evaluator needs to be a workshop facilitator and prepare well for capturing, discussing and synthesising the representative's view	Include
PE3	Confidentiality issues	Confidential information that cannot be declared can be registered in less precise terms, e.g. profits were ok, or close to the expectations, etc.	Include
PE4	Time constraint	The work should fit in a 1-hour session, but the exercises took 2 hours. Modifications to simplify the workshops are required	Include
PE5	Capturing complex stakeholder	The perspective of users or contractors can be easy and straight forward. However, a University authority is a complex stakeholder with no easy objectives and purposes to declare. Thorough workshop preparation is recommended	Include
PE6	Unintended facilitator bias	The provision of pre-setting and recommendations to make the workshop more time-efficient generates an unintended bias of the facilitator by influencing excessively the representative's judgement	Declare
PE7	Default pre-set motives per stakeholder group	The motives of the stakeholders could be presented with pre-sets, keeping a balance between efficiency and low facilitator bias risk	Include
PE8	Stakeholder workshop form development	The stakeholder workshop form was a developed result of the prototyping exercises and presented in Appendix B	Include

[Source: Author's]

The prototype evaluation method is presented in Figure 6-1 as a 9-step procedure coloured in blue. It also includes in red the tools/meetings/technics that should be used in every step. This version is not fully documented because it is not considered a milestone, nor an outcome. The design is an ongoing process and many components of the evaluation were developed during the prototyping exercises and through the testing stages. Only the final version of the evaluation method is fully developed in Chapter 8. Nevertheless, the details of this prototype evaluation method are presented in appendix D.

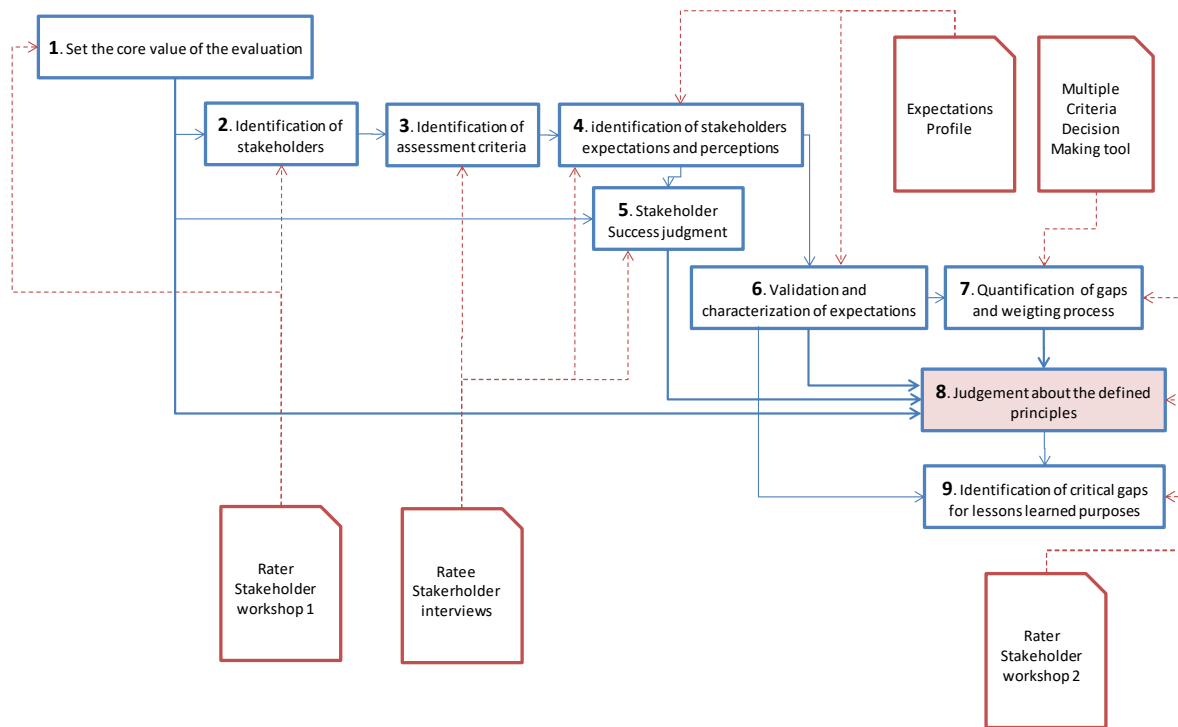


Figure 6-1: Prototype evaluation method (details are described in Appendix D)

[Source: Author's]

Figure 6-2 summaries the design process from the conceptual framework PSEPM to the prototype evaluation method. The themes captured from the experts and the observations from the prototyping exercises are colour coded to show how they relate and contribute to the prototype evaluation method and its 9 steps. In practice, this process was not entirely sequential, because some parts of the proposed evaluation were briefly presented to the experts, and then, it was part of their feedback. Nevertheless, the sequential description resembles the core idea of the research process.

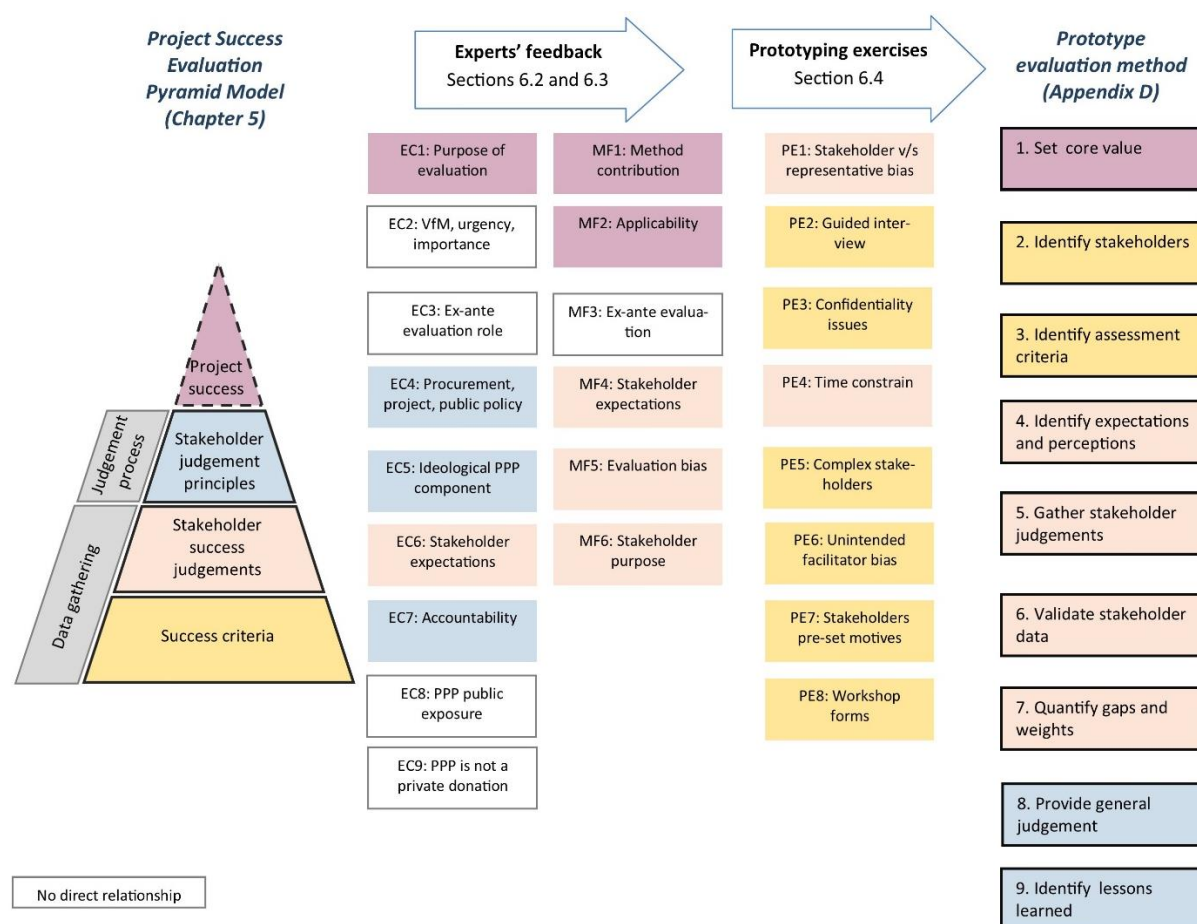


Figure 6-2: Development of the prototype evaluation method

[Source: Author's]

The next three sections will describe each of the refinement insights that were summarised in Tables 6-1, 6-2, and 6-3. The feedback regarding the context of the PPP evaluation, the feedback about the proposed conceptual framework, and the observations from the prototyping exercises.

6.2. Feedback about the context of PPP evaluation

The 170 pieces of information regarding the context of PPP ex-post evaluations were categorised into nine themes and coded from EC1 to EC9. Each theme discusses an issue that affects the development of PPP evaluations.

EC1: PPP ex-post evaluations contribute to PPP development, aiming for transparency and accountability, PPP legitimisation, industry learning, and government agency learning

Experts agree that a PPP ex-post evaluation can contribute to the development of better PPPs in the future; it is an effort that the public and private sector need to make. They recognised that there might not be sufficient willingness for implementing this kind of evaluation. One of the respondents stated that there are disincentives for conducting these kinds of evaluation, as neither the public sector nor the private sector wants to be monitored and accounted for something that has already been built (Interviewee B as described in the methodology chapter, Table 4-4).

In general terms, knowledge about PPP ex-post evaluations is rare. The existing approaches are related to ex-ante impact assessments, satisfaction surveys, financial audits, and some pure lessons learned

reports. Specific evaluation methods for assessing PPPs after completion are recognised as something that is not known or at least not sufficiently publicised.

According to the respondents, defining the evaluation's purpose is essential; four general aims were possible to conclude from the different interviewees' perspectives: transparency and accountability, PPP legitimisation, industry learning, and government agency learning (interviewees A, B, C, D).

For transparency and accountability purposes, the PPP ex-post evaluation needs to be presented in plain language allowing the general audience to understand and validate the evaluation. All the technical financial information could be presented in an appendix section of the evaluation report. Additionally, it should declare who is accountable for specific impacts that occurred during the project lifecycle, and make a clear distinction between the impact created by the PPP strategy and the impacts of the project (further explanation of this distinction in theme EC4).

The PPP ex-post evaluation also has the purpose of legitimising a PPP project or portfolio. However, this purpose is difficult to achieve due to the ideological component that supports PPPs. PPP performance focuses mainly on time/cost comparison as it is in Traditional Procurement projects. A more detailed approach of the real short and long-term effects should contribute to a better discussion over the general use of PPPs (interviewee C).

The ex-post evaluation creates lessons learned in the Industry. The recognition of the impacts that affect every stakeholder helps private companies assessing the results of previous projects; improving the precision of their proposals for new PPPs in terms of output quality and procedural strategy. If multiple perspectives are employed, the government performance in PPP management could also be assessed, helping private companies to make a better risk assessment in terms of macroeconomic issues, long-term financing options, etc.

Government agencies can also learn from a PPP ex-post evaluation. They can replicate successful components of PPPs and modify unsuccessful ones. Issues such as negotiation processes, monitoring strategies, revenue systems, quality assurance, etc. can be declared as a positive or negative outcome with a clear course of action.

The contribution and purpose of the PPP ex-post evaluation is included, as it supports the rationale for the study.

EC2: Value for Money, urgency, and importance are three relevant concepts for PPP decision-making

During the interviews, three concepts appeared as relevant drivers for PPP decision-making: Value for Money, urgency, and importance.

According to some experts, Value for Money (VfM) (described in section 3.2.2) could be biased towards the acceptance of PPPs when analysing the public-sector comparator (PSC). Interviewee G suggested that it is rare (or impossible) to find projects that have been rejected because the VfM of the PSC was better than for the PPP. The Value for Money concept is essentially an efficiency measurement to establish the amount of value that is possible to get with a certain amount of money. It differs from a cost-benefit analysis because value is a more holistic perspective of benefit, and money refers to a bigger approach that includes costs, as well as risks and opportunity costs, among other concepts. It has been stated that VfM lacks precision due to the inclusion of concepts that are difficult to quantify (intangible benefits). However, it is a concept that is associated with the evaluation of PPPs and the evaluation of international agencies projects (World Bank, regional development banks, etc.). The discussion confirmed that the VfM measurement is biased to the acceptance of PPPs when analysing the public-sector comparator (PSC).

Urgency is the principle for selecting a PPP approach because the need for service facilities cannot be postponed. Interviewee A stated that urgency in practice is not driven by technical measurements but political terms; VfM supports a decision, rather than being the main factor.

‘Importance’ is the counterpart of ‘urgency’. Political decisions are the ones that drive the development and the conceptualisation of PPPs depending on how relevant is a facility. Importance advocates for the highest long-term benefits with low relevance to the time of delivery.

This theme is acknowledged as it affects the application of the evaluation method, rather than its design.

EC3: The ex-ante evaluation must have an important role in the subsequent ex-post evaluation

Interviewees support that the ex-ante evaluation is a key factor for conducting a PPP ex-post evaluation. With a precise ex-ante evaluation scope, the PPP ex-post evaluation can contain the comparison of what was expected and what was achieved by the project. This comparison before-after legitimises the evaluation process. If there is no clear ex-ante evaluation, the ex-post evaluation can be biased by the drivers that are present at the moment of the evaluation such as current affairs, political scenario, media coverage, etc. (interviewee B). For PPP projects, Key Performance Indicators (KPI) are employed to control and monitor the facility operation, tracking the service quality; this, from the perspective of the public authority to assess the private consortium performance.

Even though the ex-ante evaluation is an important input for the ex-post evaluation, the researcher argued (during the conceptual framework presentation) that it cannot be the only source of information. Additionally, all the unforeseen events that are considered impacts to specific stakeholders need to be addressed. PPPs can last for more than 20 years, making the lifecycle of the project as relevant as the initial conceptualisation and purpose of it. As a response, interviewee B states that, according to her experience, ex-ante evaluation (for any project) rarely includes the user as part of the process, but when thinking of an ex-post impact evaluation, then, user’s satisfaction is commonly mentioned as a relevant input; this is an inconsistency.

This theme is acknowledged and considered for the refinement of the evaluation method

EC4: The evaluation must distinguish between evaluating the procurement method, the project, and the public policy

The PPP as a procurement method needs to be differentiated from (1) the project that is being procured, and (2) the public policy that supports the public effort. There is a misconception that a PPP is a goal by itself, whereas it needs to be considered only as a tool to pursue a public policy goal. The evaluation must consider a separation between the impact that is a direct effect of the PPP strategy, and the impact that is common to any kind of procurement method. According to Interviewee A, if the separation is not made, the PPP could be negatively assessed, because it is not possible for the procurement method to stand over a public issue that needs to be solved with public policy. For example, the urban segregation produced by an inner-city toll road is an impact of the project, with no relation to the procurement method. A toll road PPP project has the aim of moving cars for point A to B, whereas a general transport policy has the aim of harmonising between speed limit, congestion, security, connectivity, new infrastructure, and user experience. With a Traditional Procurement or with a PPP, the urban impact to the surrounding community will be the same.

In practice, the process of classifying the impacts is not easy to do and involves expert judgement, because there is a lot of ‘noise’ that affect the analysis. For the same case of the toll road, the impact could be analysed as a result of the proposed and accepted design of a Design Built Operate Transfer (DBOT) strategy, which could differ from the design in a Traditional Procurement strategy.

This differentiation is included in the evaluation method.

EC5: PPPs have ideological drivers

Something that is not well described in the existing literature is the ideological perspective towards PPPs. Political ideology affects the procurement methods and it has resulted in the acceptance and rejection of PPP projects. The political bias affects not only the approval of PPP projects but also the subsequent evaluation of them. PPP ex-post evaluations need to consider political bias when legitimising PPP projects. In general terms, the inclusion of the private sector for providing public services is considered contrary to the left-side political principles (maybe excluding centre-left). Conversely, right-side political principles advocate for a minimum involvement of the government in issues that can be solved by private entities. This ideological bias can have an impact on a PPP evaluation and should be carefully analysed, as one interviewee stated.

Another thing worth saying is that in my experience, normally everything that has to do with public investment is very tinged with an ideological component and therefore you have to be very careful when choosing stakeholders and evaluators and the data that is used to feed the model; It is the most complicated part of the evaluation process (interviewee B).

This theme is acknowledged because it supports the approach taken for developing the evaluation method.

EC6: Stakeholders' expectations can change through time and might be hidden

What is important about the PPP for the public authority, the private consortium, the user, or the community can change over time, and this might affect the ex-post evaluation. This is conceptualised as the change in the expectations of the stakeholders. This change is relevant because in practice, assessing the project by their original objectives is not the only way to legitimise the public investment. Projects that have been operating for 20 years can have a new surrounding environment and conditions that affect the judgement of the people involved with the project. For example, a change in the macro-economic conditions (affecting the consortium), or the availability of recent technology (affecting users' perception of quality). The ex-post evaluation should address this change of expectations and trace, from the conceptualisation to the completion of the project, how the expectations have varied, and possibly what unforeseen impacts occurred during the project implementation.

Another difficulty for evaluating PPPs is the hidden motives of the different stakeholders that are involved in the PPP; which are covered by 'official' interests. The profit driver of the private company is in many cases hidden from the general public. This lack of transparency is consistent with the idea that the process for providing a public service should be known by the public audience; by the taxpayer. The hidden interests also affect government agencies that have different specific objectives towards the same PPP project. The objectives can be related to urgency, to relevance, to budget control, to political revenue, etc. For example, the department of finance could reject a PPP proposal because of the risk exposure, whereas the real hidden interest is to maintain budget control. One interviewee said.

In the case of the public sector, there are stakeholders in the same state (role) that have different interests. With three administrations (governments) the 'Chacao' bridge passed from PPP, to then cancelled, and then procured as a public project (interviewee A).

Evaluating the impact of PPP projects over multiple stakeholders involves gathering stakeholders' expectations in many criteria. Then, an ex-post evaluation should consider that a stakeholder can

change their expectations over time (throughout the project lifecycle) and that they can have hidden expectations that will not declare in an evaluation.

This theme is acknowledged and considered for designing the evaluation method.

EC7: Accountability of PPPs is difficult to track due to their long timeframe

The accountability of PPP projects is difficult to track, because stakeholders change through the project lifecycle, changing also their view towards the project. However, experts suggest that the main stakeholder responsible for a specific problem will always be the government agency, more than the private consortium. Depending on how critical is the issue that triggers the search for accountability, the government in place at the moment of raising the issue is the most probable accountable. This accountability is also influenced by political bias; It is possible to blame previous authorities, especially if they were from a different political party. *“The accountability ultimately relies on the procurement agency that started the process in the first place, so it’s the department of health that procured the new hospital (for example). They have to say that they are responsible for the project” (interviewee F).*

The issue is acknowledged because it influences the applicability of the tool.

EC8: PPP evaluation might be affected by the project’s exposure to the public

Interviewee I emphasised that for evaluating PPPs, a differentiation needs to be noticed between the types of infrastructure projects: economic infrastructure (ports, energy systems, power plants, etc.) and social infrastructure (hospital, schools, prisons, etc.). Their evaluation is influenced by the project’s exposure to the public. In social infrastructure project, the final user is less engaged with the project (conceptualisation, design, service provision), than users in economic infrastructure projects. In a port, the end-users (clients) are companies that create strong commercial relations with the port management and their knowledge level in terms of legal commercial rights, risks, contract analysis, etc. is very high and complete. Conversely, regular individual users that are clients of a hospital or a school have less preparation and less engagement with the service that they are using; the relationship with the service is not as close as in economic projects, and the large number of clients makes the relationship even more distant (interviewee I).

This different engagement levels of the stakeholders can affect the PPP evaluation. A user with low expertise makes the evaluation to rely on an overall perception rather than more technical and verified issues. The service quality, the delivery time, or the service price are perceptions that users with low expertise cannot support with reliable evidence. When evaluating a PPP project, the relevance of each stakeholder needs to be assessed to effectively filter, analyse and weight the different perceptions, contrasting them with objective support.

This issue is acknowledged; it will be considered for the testing process but not for the method design.

EC9: PPPs are sometimes misunderstood as private donations

PPP projects, same as public investments are intrinsically evaluated by the society in general. This can be seen (in Chile and Australia) during election campaigns that include PPP topics as a valuable insight for voters. At a higher level, PPP projects are partially evaluated with political parties that are in favour or against this kind of projects. Interviewee F stated that PPPs can be misinterpreted as a private donation, which in any case is not true. PPPs can act as loans for the government (urgency), as a more efficient way to deliver a public service, or as a better way to employ private expertise and risk allocation; but in all cases, the government must pay for the provided service through direct transfer, or through user service fees. In terms of service fee, this is not a PPP exclusive financial method as

traditionally procured projects can also be financed by user service fees, while being a completely public designed, built, and owned project.

This theme is acknowledged, and it supports the study discussion section.

6.3. Feedback about the proposed conceptual framework

The conceptual framework PSEPM was presented to the experts and they provided feedback about their different components, commenting on the challenges and possible solutions to the problems they identified. The Method Feedback (MF) is grouped by themes, which are supported by the feedback of one or more experts and coded from MF1 to MF6.

MF1: The proposed conceptual framework was considered useful but with scepticism

Experts agree on the contribution that a PPP ex-post impact evaluation method could do. However, they differ regarding the contribution of the proposed model PSEPM.

Three comments about the contribution of the method were gathered. First, the idea of addressing multiple perspectives was considered a positive outcome of the method. Second, the traceability of the perspectives, meaning a clear criteria declaration and the reasons to support the perspectives, was considered useful. Third, the idea of reaching an agreement for Project Success using the principles described in chapter 5 (willingness for repetition, no excessive negative outcome, and no evident better alternative) was not well understood by some experts and it was questioned. However, the scepticism is considered reasonable, as the framework is completely new, and no real case was presented. Maybe a better explanation with a larger presentation could have helped. Nevertheless, the researcher at this point considered the feedback valuable, but without incurring any major modification to the original conceptual framework.

This issue is acknowledged, as it will be considered for the testing process, but not for the method design.

MF2: The evaluation method could be applied in real-world, but the incentives are not evident

Experts generally agreed that the idea of the proposed evaluation method should be and can be applied to real-world projects. There is a positive first impression of the method, regardless of the many concerns that need to be solved prior to its application. The incentive to apply the method is not evident as it is difficult to make an evaluation of the public sector. One expert suggested that the government is like 'a child' that does not want to be evaluated; which is supported by Langbein and Felbinger (2006).

Private entities, research centres, development banks and international organisations should be the first and more interested in applying this kind of evaluation.

Experts also suggested that an ex-post evaluation should be part of the regulation in order to be effective. A voluntary process will not have the expected results in term of accountability and transparency, especially if the assessed project presents a major problem.

This issue is acknowledged and will support the discussion section.

MF3: The proposed evaluation method must show a clear link with an ex-ante evaluation

The proposed evaluation method does not have a clear link with an ex-ante evaluation. This means that the original purpose of the project and, consequently, the way a PPP project is going to be evaluated could be influenced by the context in which the PPP ex-post impact evaluation takes place.

For pursuing objectivity, the ex-ante evaluation should be an important component of the designed ex-post evaluation.

This suggestion is acknowledged and considered when designing the evaluation method.

MF4: Stakeholders' expectations should be well traced and analysed by the evaluation method

Interviewee C said that is difficult to trace the original stakeholder expectations because they are not static as they change throughout the project lifecycle. An ex-ante evaluation should fulfil this requirement, but it has not been done thoroughly, and unforeseen events in the project lifecycle can generate legitimate expectations that are not planned or declared in any contract or ex-ante evaluation. It is unfair to consider expectations that were not available at the time of designing and conceptualising the project. Somehow this should be considered in the method.

The timeframe of the expectations was also discussed. For example, the contractor's expectations end when they deliver the infrastructure (end of construction stage). Conversely, the public authority, the users, and the community have expectations for all the project lifecycle including the end of the operational stage. This difference in the timeframe of the stakeholders' expectations is important to consider.

Another issue to address is the official versus unofficial expectations of the stakeholders. Political decisions, especially in the public sector hide some interests and goals that cannot be officially stated. This might lead to a confusing, or even wrong analysis when evaluating a project. A wrong judgement with incorrect (misleading) information could affect the objectives of the evaluation in terms of accountability, legitimisation and learning.

This issue is included in the final version of the evaluation method (Chapter 8).

MF5: The proposed method could have a bias of the evaluator conducting the evaluation

Interviewee E highlighted the possible bias of the evaluator and its effect on the evaluation outcome. The method has a bias risk because it involves a judgement over the different perspective's validity based on information and logic. The evaluators should solve possible conflict and inconsistencies between the different perspectives and provide a judgement probably disregarding (or lowering the relevance of) some arguments. Thus, the outcome of the evaluation could be highly influenced by the evaluator's judgement. This bias should be declared and reduced employing as much objective information as possible.

According to the researcher and as a counterargument, the bias of the evaluator is better addressed in the proposed method than in other forms of evaluation, in which the evaluator can influence even more the evaluation outcomes. In the case of the proposed method, the evaluator has less 'power' than in other evaluations because the role is narrowed to address (mediate) all the possible perspective, without considering his/her own perspective. This forces the evaluator to have a holistic view and declare the existence of the different perspectives. In other cases, the perspectives are selected by the evaluator, creating even more bias risk. This was partially analysed in section 5.1 when comparing single perspectives and multiple perspectives evaluation approaches.

The risk of having an evaluator bias is declared as part of the final evaluation method documentation.

MF6: The stakeholder identification process must be well developed in the proposed method

Interviewees B, C and E mentioned that the method must have a clear stakeholder identification process because a wrong identification will lead to wrong conclusions. For example, within the public sector, there are many agencies that should be considered different stakeholders, as they have clear

and independent aims and goals. Moreover, these agencies can have conflicts between them for the same project, as different views and goals get confronted. That is the case of the department of finance and the department of infrastructure. They both represent the interest of the public sector, but sometimes they could have opposite perspectives towards the same results.

An uncomplete stakeholder identification could also influence the recognition of long-term interests that are present in every project. Interviewee B discussed that the method should be able to address the long-term objective that at the first impression could not be addressed by the stakeholders. Long-term aims such as geopolitical advantage, urban development, cultural asset, etc. need to be addressed by at least one stakeholder. This means that the evaluation approach should identify the stakeholder that is evaluating the project using the long-term goals. For example, employing the criterion 'contribution to the overarching strategy, public policy, or national agenda'.

Regarding the amount of information that the evaluation can capture, interviewee B stated, 'the matrix that captures every stakeholder's perspective will grow and it will become big and very complex, but it is necessary to do it; it is a valuable insight'. However, the time consumption associated to this process needs to be considered. A more accurate and disaggregated stakeholder map means a more complex and time-consuming evaluation process. This needs to be accounted, as in practice, efficiency is an important parameter of an evaluation method.

The purpose of every stakeholder, in general terms and over the project, is relevant to identify. This information allows the evaluation to present a logical reasoning of why a stakeholder has a set of criteria for evaluating the project. The idea is to generate a robust argument for every stakeholder. Interviewee B explained that (for example) it is very different to have a syndicate (work union) evaluating the project, or the executives; they have completely different purposes for the project, and the declaration of those purposes help in the validation of the evaluation criteria and the subsequent criteria assessment from their own perspective.

The stakeholder identification process that complies with the suggestions and comments is included in the evaluation method.

6.4. Prototyping exercises

The stakeholder workshop is one component of the evaluation method, undertaken for capturing one stakeholder perspective. It is a formal meeting with a stakeholder's representative for declaring how that stakeholder judges the assessed project. It includes: description of the stakeholder, purpose of the stakeholder towards the project, key concepts that motivate the stakeholder function, success criteria for the project, valuation of the criteria for the project, and answer to the three judgement principles discussed in section 5.5 (are you willing to repeat the project? Is there any excessive negative impact to any stakeholder? Is there any evident better project alternative?)

To design the stakeholder workshop structure, two prototype exercises were. To test the interaction between the evaluation facilitator and the stakeholder's representative, a practitioner in the construction industry and a university authority were selected. The duration of each exercise was 2 hours and included discussion and modifications of the stakeholder workshop structure. The data analysis in this case is different from the expert interviews. Here, a functional (black box) testing (Hevner et al., 2004) was employed as a data analysis method, generating eight refinement insights.

The refinement insights are observations of the researcher that will be considered for refining the evaluation method. They are coded as prototype exercise (PE) from 1 to 8.

PE1: There could a bias between the stakeholder representative and the theoretical stakeholder

Project stakeholder can be defined as one person or group of people that have an interest in a specific project. If they are a group of people, they do not necessarily agree on the perspective they have towards the project; users, community, shareholders, etc. can have different opinions. Therefore, there is a bias between the stakeholder and the stakeholder's representative appointed for conducting the stakeholder workshop. The representative will never be able to represent exactly the interest of the theoretical stakeholder. Even though the representative is the closest approach to capture a perspective, it needs to be acknowledged as a potential bias risk.

This issue is included in the evaluation method as part of the facilitator analysis.

PE2: The stakeholder workshop requires to be guided within an overall structure

The evaluator conducting the workshop needs to be a facilitator because the information that needs to be gathered is complex, difficult to follow, and difficult to synthesise. The evaluator needs to explain each part of the process and guide the representative in areas that are not easy to understand (e.g. motives and judgement principles). Even though it is possible to follow a structure, the evaluator has to explain, discuss, synthesise, and also triangulate the information received by the stakeholder. Consequently, the method should consider extensive material for the facilitator to study and prepare prior to conducting a stakeholder workshop.

The observation has been included in the evaluation method.

PE3: Confidential information cannot be presented in a workshop

The stakeholder workshop captures the argumentation that supports the judgement of a stakeholder regarding a project. During the prototyping exercise, a confidentiality issue arose when the stakeholder representative selected 'profit' as a success criterion. The financial result of a project was not revealed by the stakeholder representative and it was not possible to declare it in a public evaluation. A relative scale approach was proposed to answer the profit criterion valuation (percentage of achievement in terms of profit). The confidentiality problem was not solved because the negotiation process of the private company could get exposed.

In a project, clients and contractors employ what is called in negotiation research as 'anchors'; low balls or high balls referring to the starting points for negotiating. These anchors are common in contract management. Declaring profit in an ex-post evaluation could reveal information that was hidden during negotiations. For example, if the contractor declares that they gained more profit than what was expected (e.g. 110%), it might contradict previous negotiations. Another example; an environmental lobbyist will not be interested in revealing higher result than expected because it can affect future campaigns.

This observation is included in the documentation of the evaluation method, arguing that even a non-specific declaration of a stakeholder will contribute to the evaluation.

PE4: The stakeholder workshop should be under 1 hour

The stakeholder workshop should take approximately one hour for practical reasons; it is easier to ask representatives for a 1-hour meeting, rather than a 2-hour workshop. In the prototyping exercises, 20 minutes were spent in explanation and instruction. The exercises took more than 2 hours each. The stakeholder workshop is redefined for fitting into a one-hour session. More 'ready to tick' information and templates need to be designed.

This issue is included in the evaluation method documentation.

PE5: Better preparation is needed to capture complex stakeholders

When generating a stakeholder perspective, the complexity of the stakeholder can make the workshop very difficult. Traditional stakeholders such as the contractor, or the end-user can be easy to capture. However, in a University structure, University authorities play a role that is difficult to declare. Their motives for existence are not easy to synthesise, and consequently, the success criteria that can be employed to evaluate a project from their perspective is difficult and time-consuming.

Thorough preparation and background check prior to the workshop with each stakeholder should be recommended in the evaluation method. Additionally, an introductory email with general information including the evaluation's aims could also help to better gather a complex stakeholder perspective. Forms and templates could also help; for example, a list of motives and success criteria for the representatives to agree, to change, or to completely disagree.

This issue is included in the evaluation method.

PE6: Excessive facilitator guidance in the workshop can generate an unintended facilitator bias

The refinement insights PE2, PE4, and PE5 (in this section) refer to how the facilitator should interact with the representative during the workshop. An additional observation is the bias of the facilitator while guiding the workshop. To make the workshop effective and 1-hour long, some information and suggestions could be provided by the facilitator including templates and recommendations for the stakeholder perspective. This process generates an unintended bias of the facilitator because the suggestions can modify the representative's real judgement. For example, recommending criteria could influence the full identification process that the representative should do by its own. Even though the facilitator could encourage the representative to modify and critic all the pre-setting, the risk of having a non-real perspective influenced by the facilitator cannot be eliminated.

This issue is declared as part of the evaluation documentation.

PE7: To have only pre-set motives could mitigate the unintended facilitator bias

Aligned with the previous issues (PE2, PE4, PE5 and PE6), the balance between being efficient with pre-sets and having low facilitator bias risk could be achieved by systematising the stakeholder motives. Stakeholders could be classified with generic motives, providing a less influential template to the workshops. Motives are the general drivers of the stakeholder, i.e. the reason why they are a stakeholder, for example, 'pursue justice', 'make a profit', 'achieve life quality', or 'represent workers'. Using pre-sets for the motives could help reduce the time for explaining and discussing them while maintaining a low facilitator bias risk. These pre-set forms might not be applicable to the success criteria, as they are more specific for the stakeholder and for the project that is being assessed.

This issue is included in the evaluation method together with all the refinements about the technics for conducting a stakeholder workshop.

PE8: Stakeholder workshop templates and forms were developed

During this prototyping exercises, the first draft form for conducting a stakeholder workshop was developed. The exercises started with a list of information to gather, and after them, an A3 size sheet was designed including space to fill every component of the stakeholder workshop. The final stakeholder workshop form is included in the final evaluation method (Chapter 8) and presented in Appendix B.

This development is included as part of the evaluation method documentation.

The three design inputs presented in this chapter (evaluation context, method feedback, and prototyping exercises) were employed for designing the prototype evaluation method. The next chapter will describe how this prototype was tested, further developed, and refined using three PPP projects and test cases.

6.5. Summary

This chapter presents the result of the first stage of the research (stage A), in which the conceptual framework 'Project Success Evaluation Pyramid Model' (Chapter 5) was presented to 13 PPP experts and practitioners in Chile and Australia. In total 15 themes were generated after applying a thematic content analysis to the interviews. Nine themes synthesised the feedback from the experts regarding the context of PPP evaluations; including the difficulties associated with PPP evaluation, the incentives and limitations, and the considerations that need to be addressed when planning a PPP ex-post evaluation. Additionally, six themes addressed the feedback from the experts regarding the proposed conceptual framework. This chapter also included eight refinement insights (lessons learned) that were a result of the prototyping exercises of the stakeholder workshop; the main component of the evaluation method. The three components (evaluation context, method feedback, and prototyping exercises) were the design inputs for the prototype evaluation method.

Regarding the evaluation context, PPP experts and practitioners stated that conducting an ex-post evaluation on PPP projects was not common and the topic was usually not present in the projects they have worked on. Nevertheless, they all agreed on the relevance of conducting ex-post evaluations to PPP projects and that it should be pursued in the future. In terms of method feedback, the experts recognised the conceptual framework's contribution and its potential applicability. However, some issues were highlighted for its practical application such as the inclusion of ex-ante evaluation, the stakeholder bias treatment, and the acknowledgement of the stakeholders' purpose towards the project. The lessons learned taken from the prototyping exercises were related to the workshop structure and templates, amount of information to provide to the stakeholder, workshop length estimation, and questions testings.

This chapter described the feedback about the conceptual framework PSEPM and the development of the prototype evaluation method. The next chapter will present the testing and refinement process for developing the final evaluation method.

CHAPTER 7

TESTING THE EVALUATION METHOD ON THREE PPP PROJECTS

7.0. Introduction

This chapter describes the development of the final evaluation method using three PPP projects as test cases. This is the testing, refinement, and improvement of the prototype evaluation method presented in Chapter 6 to achieve the second research objective.

The chapter is structured as follows. Section 7.1 explains the development of the final evaluation method employing the three test cases. Section 7.2 describes the results of the first test case, a social housing PPP project in Chile. Section 7.3 shows the results of the second test case; a toll road PPP project in Chile. Section 7.4 presents the final test case; a hospital PPP project in Australia. Finally, the analysis of the results for the three test cases is discussed in section 7.5.

Part of this chapter has been published in the following articles: section 7.2 Oliveros R. and Gonzalez-Mathiesen (2017); section 7.3 Oliveros R. and Aibinu (2018), and Oliveros R. and Aibinu (2019).

7.1. Development of the final evaluation method

The prototype evaluation method presented in Chapter 6 was applied to three PPP projects for refinement and improvement, and the result of each project evaluation is presented in this chapter. This application is not about the testing of the project-specific issues but testing of the ex-post impact evaluation method (the tool).

Figure 7-1 shows how the insights used for the refinement of the method were obtained from the three test cases and how they were applied towards the development of the final evaluation method. Each test case had a data collection process and a data analysis process. The data analysis process was done mostly in parallel (as shown in the diagram), allowing for some insights to be applied retrospectively into the test cases every time that was possible. For example, the structure of the evaluation conclusion was developed as a refinement insight from test case 3 but applied to all the test cases (green arrows in the diagram). Sections 7.2, 7.3, and 7.4 describe the test cases as a result of this interinfluence.

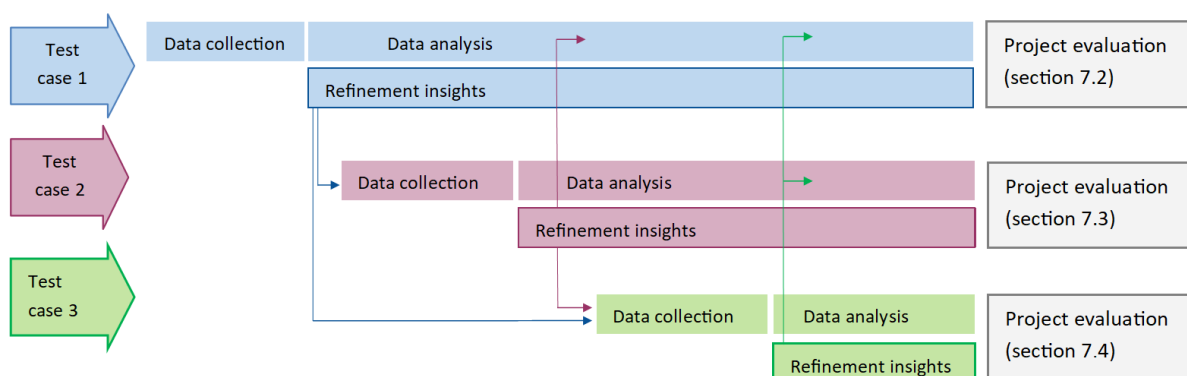


Figure 7-1: Representation of how the insights from the test cases inform each other

The testing and analysis of each test case is conducted and presented as a project evaluation with five components:

1. Test case overview/context: The context of the project is presented as a chronological description, with a special focus on the PPP characteristics. This description helps to understand the subsequent evaluation.

2. Stakeholder identification: The process of identifying and selecting the project stakeholders is described.
3. Stakeholder workshop analysis: The main data collection process of the evaluation comes from the workshops held with each of the identified project stakeholders, in which their perspectives towards the project were gathered. The data analysis process is different for every test case and explained accordingly.
4. Evaluation conclusion: A final judgement over the project is drawn across three main topic areas: (1) the project-system, (2) the procurement method, and (3) the evaluation objectives (accountability, PPP legitimacy, Industry learning, and public-sector learning).
5. Refinement insights: They are the researcher's observations from applying the prototype evaluation method to the test cases. They are the inputs for developing the final evaluation method.

The refinement insights are coded as 'SH' (in the social housing test case), 'TP' (in the transport project test case), and 'HP' (in the health project test case). Each refinement is also categorised as an acknowledgement, inclusion, or modification; similar to Chapter 6. These are conceived as follows:

- 'Acknowledgement' means that the insight/observation does not generate a need for a specific change in the evaluation method, and they are not included as part of the design. However, they support other parts of the research, such as the relevance of the topic, the significance of the findings, the further research development and/or the applicability of the evaluation method.
- 'Inclusion' refers to any insight or observation that needs to be included in the evaluation method procedure; thus, they are insights that are present in the final evaluation method presented in Chapter 8.
- 'Modification' means that a specific component of the evaluation method is modified by the gathered insight or observation. Thus, they should be present in the final evaluation method in Chapter 8.

Figure 7-2 presents the developed final evaluation method. This version incorporates all the refinement insights from the three test cases and is fully described in Chapter 8.

PPP ex-post impact evaluation procedure

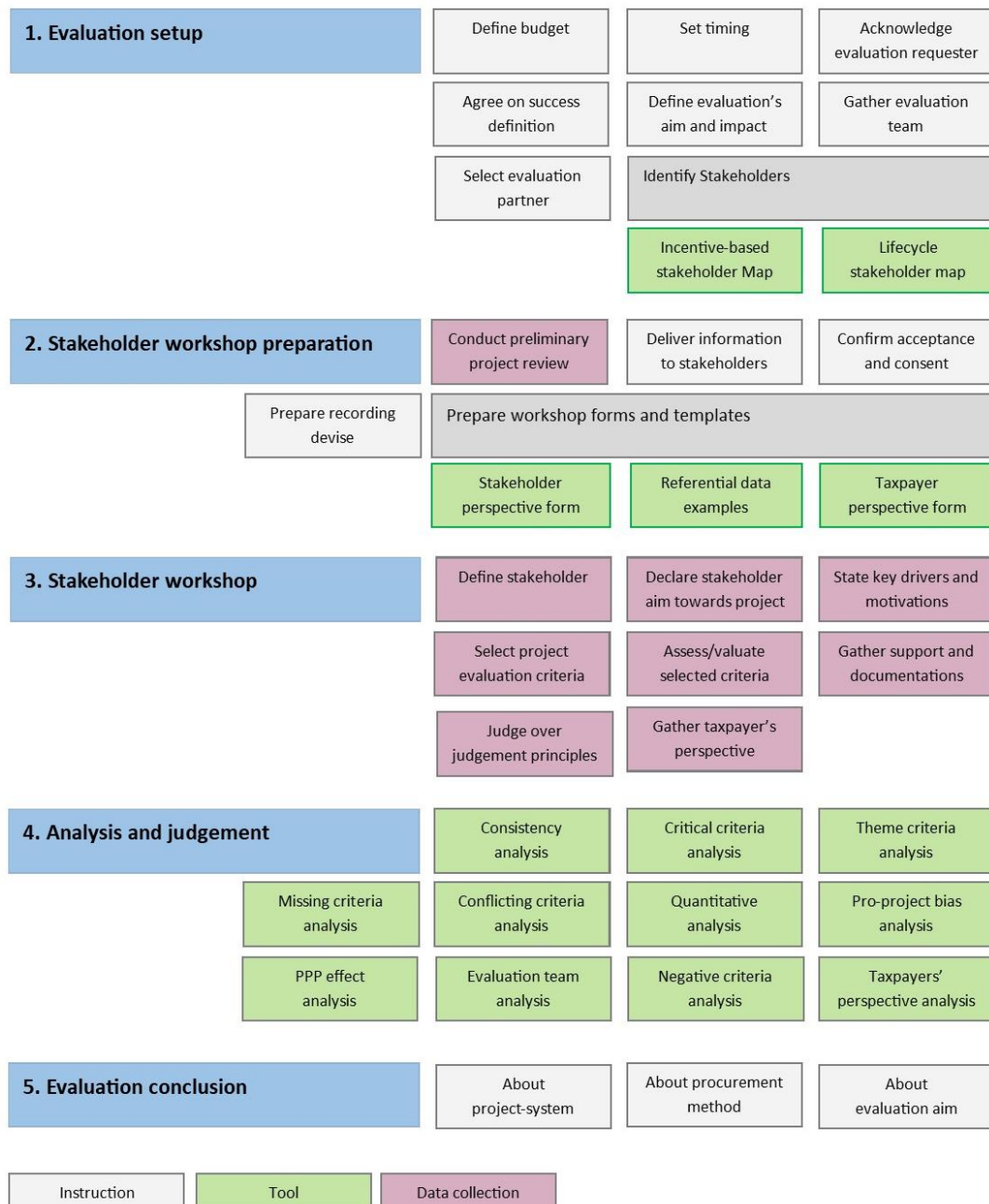


Figure 7-2: Final evaluation method (fully described in Chapter 8)

[Source: Author's]

Figure 7-3 summarises how the final evaluation method was developed as a result of the test cases, starting from the 9-step prototype evaluation method (chapter 6). The refinement insights of each test case are included to show how the insights are connected to the 5-step final evaluation method.

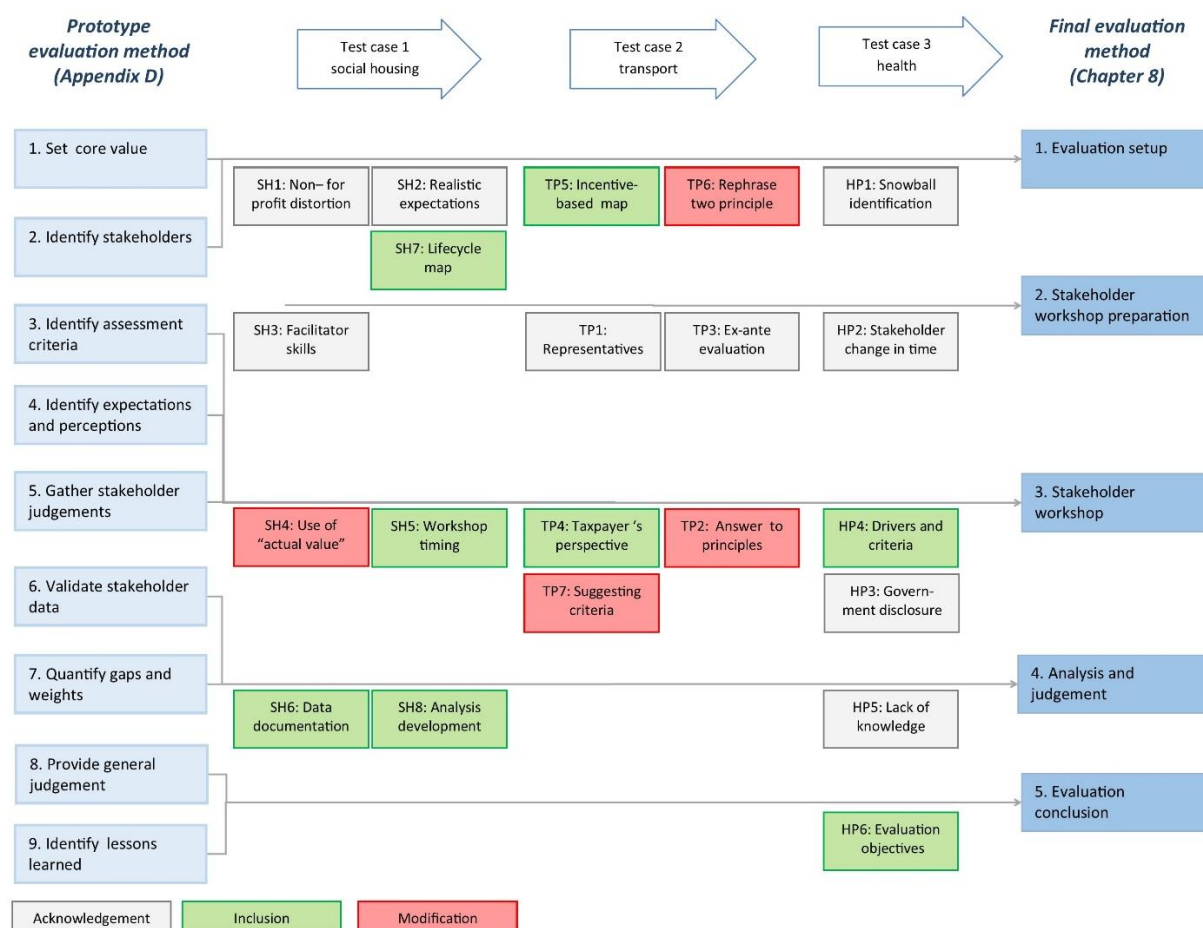


Figure 7-3: Test cases' contribution to the final evaluation method development

[Source: Author's]

7.2. Test case 1 – Social housing project

The result of applying the prototype evaluation method to a social housing PPP project is presented in this section. The analysis of the project evaluation results has eight refinement insights for improving and developing the final evaluation method.

The structure of the test case 1 project evaluation is as follow. First, the context of the project is presented (overview). Second, the stakeholder identification process is described. Third, the analysis of the stakeholder workshops is presented using five types of analyses namely, critical criteria analysis, negative impact analysis, missing perspectives (evaluator's perspective), PPP effect analysis, and stakeholder judgment analysis. Fourth, the conclusion of the evaluation is presented. Then, the analysis of the project evaluation results describe the eight refinement insights (lessons learned).

7.2.1. Test case 1 overview

The context of Chilean social housing policy

Social housing is the shelter provided generally by the government to supply a need for vulnerable families that cannot afford a private house. There are different methods to deliver the housing service, including own construction, legal right over the land, a direct subsidy to entities that build and operate a social housing rental service, or a subsidy that goes directly into the beneficiaries' houses. Public housing usually refers to publicly owned houses that are rented to vulnerable families at an under-

market price. Conversely, social housing refers to government-funded infrastructure that is finally owned by the beneficiary.

In Chile, the social housing policy only includes houses that, in the end, are owned by the beneficiary; there is no rent option. The different levels of poverty and vulnerability have different policy strategies, which differ in terms of savings requirement, house final cost, debt level, subsidy amount, etc. Currently, for the lowest poverty level (most vulnerable group) the social housing policy is regulated by the *Fondo Solidario de la Vivienda* (housing solidarity fund) DS 49/2012. The regulation establishes that the total house cost is USD20,000 approximately and is funded mostly by the government, with a beneficiary saving requirement of USD500 approx. with no debt attached. Even though they have certain temporal restrictions for the house occupancy, the family (the family leader) gets the ownership of the house, which is considered the *Chilean dream*.

In 2006, the housing policy changed with a new procurement method. The procurement method is the organisational system that assigns specific responsibilities and authorities to people and organisations and defines the relationships of the various elements in the construction process (Love, Skitmore, & Earl, 1998). Prior to 2006, a Traditional Procurement approach was employed, which means that projects were publicly funded and managed, and privately built. With the policy change, a Public-Private Partnership approach was set, and projects are now publicly funded, but privately managed and built. It is relevant to make the differentiation between the public policy and the procurement method; the latter is only a component of the policy. For example, the beneficiary eligibility is part of the policy and the procurement method (PPP) has no influence in this aspect.

The current social housing policy in Chile that started in 2006 allows private entities to generate social housing project from their conceptualisation until their delivery. This procurement method is based on a demand size subsidy creating a voucher system in which beneficiaries can acquire a social house from the existing projects or from future planned housing projects. The public sector certifies private entities (for-profit and not-for-profit) to become social housing management entities, i.e. social housing developers (*Entidad de Gestion Inmobiliaria Social, EGIS*). They are responsible for managing the different stages of the housing process, from the prospecting of housing needs, demand organisation, savings planning, feasibility study and provision of land, design of technical urbanisation projects, the design of the set and architecture of the dwellings, the design and implementation of a community self-diagnosis, the design and application of a social habilitation plan, the study of construction proposals and the contracting of works, the design and application of a methodology of technical assistance for the process after obtaining the subsidies, and finally the technical inspection of the works and the processing of the final receptions concluding with the corresponding writings and inscriptions (Toro Blanco, 2006). This model generates a market of beneficiaries that are potential subsidy receivers, and social housing developers that search for beneficiaries to build housing projects.

This kind of PPP is here defined as Manage, Design, and Build (MDB) because the social housing public policy is a demand-side subsidy that transfers to private companies the responsibility and accountability for the entire process of developing a social housing project, including the demand management (search and engagement of eligible beneficiaries). However, there are certain aspects of the Chilean social housing systems that are different from a common PPP. The risk allocation is different from the traditional PPPs, as the government allocates less financial risk to the private company by financing throughout the process. The EGIS must be a different entity from the builder (a consortium is not allowed), and they must provide proof of a transparent tendering process before the public sector can effectively pay for the construction directly to the builder. This kind of model may not be considered a PPP by the general literature which considers no middle step between formal privatisation and a concession that includes an operational stage (Moskalyk, 2011). However, for the purpose of this study, it is considered that the Chilean social housing PPP approach is certainly a

partnership, as the engagement of the private sector (which is the variable for classifying a PPP) is broader than in a Traditional Procurement approach, and definitely, the system is not privatisation.

The prototype evaluation method was tested in a completed social housing project, delivered by the NGO TECHO.

The context of the selected social housing project

In the city of Penco, Chile, the slum called 'La Greda' was populated by 60 families around the year 2004 in an area that was government-owned. In 2005 the NGO TECHO started working in the slum with volunteers in social activities to promote work, education and health issues, among others. Employing the social housing policy, TECHO was certified as EGIS in order to manage a social housing project. TECHO's first project in the region was for 'La Greda' slum community.

A site in the city of Penco was found for the development of the project, however, the piece of land needed extra infrastructure to manage the big slope.

The first design presented to the community was a four-storey building without a backyard. The community was completely against this design, and after searching for new house designs, the project from the architectural office ELEMENTAL (from the Pritzker Prize-winner Alejandro Aravena) which was already built in the north of Chile, was presented as an option. Even though the community wanted individual houses, the constraints of budget and the extra cost of sloping site preparation allowed this option, that in general terms, was accepted by the community. The final capacity of the project was established in 98 dwellings, with a three floors-continued façade design.

The slum community consisted of 60 families that could apply to the subsidies, so there were +30 units that needed to be occupied by other residents out of the original community. The application process to the new 'social housing project' was managed by the community leading team, prioritising the relationship with the neighbours (family nexus) and the geographic origin (same city).

The surrounding community (close neighbourhoods) were against the new social housing project, mainly because they were afraid of the quality of the people. They had rumours about their bad behaviour, criminality issues, etc. After several conversations, and meetings with the slum community, the problem was not solved, but the project was able to continue with the construction process.

The project finished in 2010, the families moved to the new neighbourhood, and the original slum site was closed with fences, to avoid other people to populate the site.

After the arrival and up to date, the community still works as a group. They have an admin team and have community projects such as recreational activities, infrastructure for the neighbourhood and services for the extended area, such as a childcare service.

The context of testing the evaluation method

The social housing project reached practical completion in 2010. Therefore, the evaluation of this project employing the developed ex-post impact evaluation method is conducted 8 years after the community started living there. The aim is to evaluate the impacts that this project has generated on their stakeholders using the proposed evaluation method. Some evaluation formalities were not included in the test case because the objective at this stage is to test and design the main components of the evaluation. Thus, steps such as the declaration of the evaluation requester, or the analysis of the best time-period for evaluation were not applied.

7.2.2. Test case 1 stakeholder identification

This identification of the stakeholders was first described in the prototype evaluation method, but then fully developed during test cases 1 and 2. Two tools were designed for this purpose: lifecycle stakeholder diagram (coded as SH7), and the incentive-based stakeholder map (coded as TP5), which are applied to all test cases.

Figure 7-4 shows the lifecycle stakeholder diagram applied to test case 1. The researcher judged in which part of the project lifecycle the stakeholders were relevant, and selected a stakeholder representative for each of them. Some stakeholders were not possible to reach, and they are shown in Figure 7-4 as ‘not included’. This is not a problem for the design process; as the main objective is to test and refine the evaluation method, which can be achieved without the involvement of all of the identified stakeholders. Nevertheless, this issue must be recognised when using the test case conclusions regarding the project (contextual results). This work is not about the testing of the project-specific issues but testing of the ex-post impact evaluation method (the tool).

Stakeholder Group	Conceptualisation	Design	Construction	Operation	Dispose	Stakeholder Representatives
Private sector		EGIS - NGO perspective EGIS - technical department	Contractor			NGO office director Former NGO technical department director Not included
Public sector		Housing and Urbanism Ministry (technical entity) Housing and Urbanism Ministry (political entity) Council				Ministry Project controller Not included Former city mayor
Users/beneficiaries		Community leaders (slum) Beneficiaries (families from the slum)				Former slum and beneficiary group leader Not included
Community			Surrounding neighbours Urban experts			Neighbourhood committee leaders Not included

Figure 7-4: Project lifecycle stakeholder diagram (Social housing test case)

[Source: Author's]

Figure 7-5 shows the second tool developed for the stakeholder identification process from the evaluation method: the incentive-based stakeholder map. It is a diagram for identifying stakeholders according to their main roles and incentives toward the project.

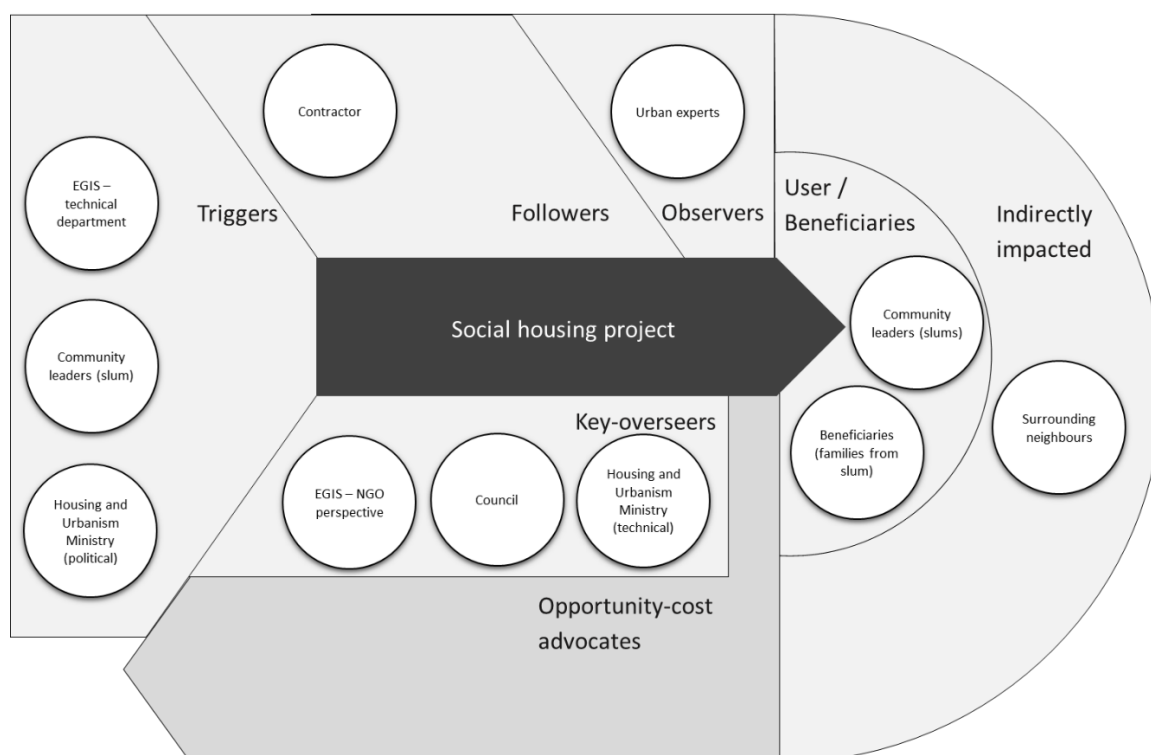


Figure 7-5: Incentive-based stakeholder map with the identified stakeholders (Social housing test case)

[Source: Author's]

The next step in the evaluation method was to conduct workshops with each of the stakeholder representatives. Their stakeholder perspectives were gathered using a stakeholder workshop form developed as a refinement insight from the prototyping exercises (PE8 in Chapter 6).

7.2.3. Test case 1 stakeholder workshop analysis

Each workshop was a one-hour meeting with a stakeholder representative in which they had to describe/declare a 'stakeholder perspective' towards the project. It comprises six components (for this test case):

1. Stakeholder's definition (who is the stakeholder?),
2. Stakeholder's aim towards the project (why is it relevant to the project?)
3. Key motivations and drivers of the stakeholder (what drives stakeholder's actions?)
4. Evaluation criteria (which criteria the stakeholder uses to assess the project?)
5. Assessment of the evaluation criteria (assess/evaluate the selected criteria)
6. Judgment of stakeholder principles (principles defined in section 5.5: willingness for repetition, no excessive negative outcome, no evident better alternative)

For each stakeholder that was included in the evaluation, a stakeholder perspective was gathered. The data is now analysed employing five types of analysis: (1) critical criteria analysis, (2) negative impact analysis, (3) missing perspectives, (4) PPP effect analysis, and (5) stakeholder judgment analysis. These analyses were developed as a refinement insight (coded as SH8) and applied to the three test cases.

Critical criteria analysis

Figure 7-6 shows the criteria selection and assessment for each of the stakeholders. The figure makes a separation between the critical and non-critical criteria. Critical criteria refer to the criteria that (according to each stakeholder) could change the overall judgment of the stakeholder towards the project. The sum of six stakeholder perspective shows that 13 criteria were categorised as critical and assessed as positive (green colour) by their stakeholders.

	Critical criteria		Non-critical criteria							
NGO (head director)	Technical house quality	City integration	Surrounding physical habitat	Community habitat	Reduce or close slum					
	Self-organised community	Link between volunteers and families								
NGO (technical department)	Localization heterogeneity		Project closeness to services	Community organisation	House size standard	House growth capacity	Level of densification	Integration of surrounding community	Capacity building of technical area	Time
Housing and Urbanism Ministry (technical entity)	Within the regulation		House use (not for renting)	House functionality	Gap between used regulation and new one	Quality of the EGIS	Quality of the Contractor	Integration (location and densification)		
Council	Neighbourhood generation		Conflict between neighbours	Efficiency in land use	Capacity to deliver social housing	Efficiency in public resources	Ownership sense form the city	Guidance of private entity	Community leader development	Urban design
Community leaders	Life change (calmness)	Co-property	Each family satisfaction	House size and materiality	Time					
	Community organisation									
Surrounding neighbours	Quality of people (in the new project)	Privacy with the existing neighbourhood	Social housing land use	Safety	Organisation of new neighbourhood	Infrastructure quality	Level of integration	Disturbs during construction		
	Drug dealing									

Figure 7-6: Evaluation criteria for project stakeholders separated by critical and non-critical (Social housing test case)

[Source: Author's]

Based on the conceptual framework 'Project Success Evaluation Pyramid, this analysis suggests that the core expectations of the stakeholders are met. This result is a strong argument in favour of a positive final judgement about the question, in this project-system successfully working? In practice, this is a holistic view of the project that shows the core impacts of every stakeholder as positive.

Negative impact analysis

The conceptual framework in which the prototype evaluation method is based has an 'optimistic' starting point. Similar to the presumption of innocence, in which a person is declared innocent until the contrary is proven; this evaluation method presumes that the project-system is successfully working until the opposite is suggested. Therefore, negative and neutral impacts are analysed individually, estimating its validity and relevance. The criteria in which neutral and negative impacts were recognised are (1) safety, (2) time, and (3) disturbs during construction.

- **Safety:** According to the surrounding neighbourhood, safety was considered a neutral impact, as they felt less safe than before. However, the data to back up that claim is difficult to obtain. According to the same stakeholder, in the beginning, the surrounding neighbourhood did not trust the occupants of the new social housing project, but currently, they have a much better relationship and is possible to trust the new incoming community. Certain crime episodes, that presumably did not occur before, are suggested to be because the new project design makes the adjacent neighbourhood to be more exposed to the public transit, providing easier access. Even though this might be a legitimate negative impact of the project, it is considered

by the evaluators to be a normal externality of urbanisation, an intrinsic outcome of any urban project.

- Time: According to the roadmap that the social housing developer (NGO) presents to the community, a social housing project should take 3 to 4 years, from the conceptualisation to the final inauguration of the new neighbourhood. However, this project took almost 8 years to finish. Multiple factors could be described for this delay. First, the expertise of the NGO was not fully developed, as this was their first managed project in the region. Second, the contractor had financial problems in the middle of the construction period and a new contractor had to be engaged. Third, community groups who were opposed to the project created another delay in the process. Fourth, the shape of the land required extra funding for slope adjustment, which is something that other projects do not necessarily need. Fifth, the 2010 earthquake also affected the project delivery deadline. Finally, problems with the sewage company influenced the project schedule. At this point, it is difficult to find accountability for the delay, especially because the relationship between the community and the NGO was based on trust. Nevertheless, the NGO should have the final responsibility for estimating a timeframe of 3 to 4 years, whereas, it was 8 years.
- Disturbs during construction: The disturbs during construction was considered a negative impact, however, even the stakeholder representatives of the adjacent neighbourhoods said it is an unavoidable impact and does not influence the perception of infrastructure projects.

Missing criteria and perspectives in the stakeholder evaluation

The purpose of this analysis is to identify any stakeholder or criterion that was not included by the stakeholders that are considered part of the project-system. This identification is done by the evaluators (researcher in this case) and presented as an additional perspective. This analysis mitigates two risks of the method that were addressed by the experts in the first stage of the research (Chapter 6). First, the risk of not recognising a valid stakeholder that should be part of the evaluation; or in this case, not been able to reach a stakeholder. Second, the risk of conducting the workshop with a stakeholder representative that do not address all the criteria that are relevant to that stakeholder. To mitigate those risks, this analysis explicitly suggests missing criteria or stakeholder perspectives that were not gathered during the stakeholder workshops, and according to the evaluators (researcher in this case), they are important to consider. Two criteria are recognised: contractor profit, and resource allocation. Also, two stakeholder perspectives are considered relevant: park-option advocates, and professional/academic view.

- Contractor profit: The contractor is an identified stakeholder that was not possible to reach for the purpose of the workshop. The project was built by two contractors. The first one was not able to continue the project due to financial problems and then the project was re-tendered. It is possible to presume that the first contractor did not make any profit with this project, which is considered a negative impact on the contractor. Moreover, 'profit' as an evaluation criterion is probably considered critical by the contractor. According to the evaluation team, this possible negative critical criterion is not considered relevant for judging the whole project because it is considered a self-managed impact, i.e. the stakeholder is the responsible for the impact with no or little relationship with the project as a system. First, this kind of projects is very normal in the construction industry, a small-mid size project of 98 dwellings; i.e. it should not have a high construction risk. Second, the contractor entered into a contract voluntarily, there was no enforcement for building the project. And third, the financial risk is an essential risk for any company that works in a market; it is the other side of profit. According to the evaluation team, this profit criterion could have been considered an important negative impact if the project would have been the cause of that impact. For

example, if there was any suggestion of illegal practices, unfair conditions, excessive power of one stakeholder, a complex project in an unstable scenario, etc. However, this is presumably not the case.

- Resource allocation: No stakeholder included a criterion to judge if the resources for developing social housing project were correct, optimal, efficient, etc.; or if the public resources could be spent on another social requirement. This is considered a strategic criterion that high public authorities should include in their project judgement. The stakeholders that could have included this criterion in their perspective are the Housing and Urbanism Ministry, an external evaluator, consulting group, or a strategic governmental authority such as the Treasury Ministry or the Social Development Ministry. To answer, whether allocating resources to the social housing project is positive (and a positive impact) the UK guideline of public investment 'Green Book' can be employed. The purpose of the Green Book is to ensure that no policy, programme or project is adopted without first having to answer these two questions: are there better ways to achieve this objective? Are there better uses for these resources? (HM Treasury, 2003). Answering 'no' to both questions is the goal for any pursued public investment. For the purpose of the evaluation, the Green Book questions could be transformed into two criteria: better use of the same resources means 'opportunity cost', and better ways to achieve the same purpose means 'efficiency'. In terms of opportunity cost, the social housing policy and its resource allocation is not generally discussed as it is one of the main social problems of the country. In term of 'most efficient solution', the use of a PPP model to procure the social housing projects has been a public discussion. This last point will be further developed in the PPP effect analysis.
- Park-option advocates: Another stakeholder that was not included in the evaluation was the park-option advocate. Employing the categories of the incentive-based stakeholder map (stakeholder identification section) a project-system could have opportunity-cost advocates. This is a stakeholder that formally presents an alternative option to the project that is being or was developed, arguing that the project should not be pursued. This differs from stakeholders that can legitimately be against the project. The project review shows that the project site had a different original purpose; it was planned for a green area (park). This might be the alternative option, however, a common practice in this context is that a change in the land use can occur due to the development of projects such as social housing. This green area alternative project is not considered relevant, as (1) the need for social housing is considered, at least, as relevant as the green area (within the norm), and (2) an explicit claim for a green area was not seen, meaning that this alternative option was just in the original planning with no committed supporters.
- Professional/academic view: Urban planners, academics, and/or professional views were not included in this test case. The reason is mainly due to the later development of the incentive-based stakeholder map (coded as TP5) that helps a broader stakeholder identification process. The incentive-based stakeholder map was employed in this test case only for presenting the stakeholder identification process. Nevertheless, according to the evaluator's review, there is a generally positive impression in the academia about urban integrated small-scale social housing projects. Moreover, the current discussion is how to improve the social housing policy to provide projects like the one that was selected for this evaluation. Therefore, it can be suggested that the judgement of external observers to this project should be positive, or at least, neutral.

PPP effect analysis

One of the conceptual components of the evaluation method is the idea that the impact of the PPP projects must be separated between (1) the impact of the public policy that supports the development of projects, (2) the impact of the project, and (3) the impact of the procurement method employed for delivering the project. The evaluation of a PPP project has an implicit focus on the attributes or factor that deal with the procurement method. Therefore, when evaluating the impact of a PPP project, this evaluation method proposes to make a separation between the stakeholders' evaluation criteria that are influenced by the fact that the project was procured using a PPP model; and the rest of the criteria that have no relationship with the procurement model and most probably would have not changed if the project would have been procured using a different approach. For example, an urban toll road might create city segregation which is a negative impact; however, it is possible to argue that the same negative impact would have happened with any procurement. Then 'city segregation' is an evaluation criterion that is not linked to the PPP model.

Under this proposed evaluation, the criteria gathered from the stakeholders during the workshops are categorised under (1) high effect of PPP, (2) medium/low effect of PPP, and (3) no effect of PPP. This classification was done by the researcher and is not intended to be a deterministic categorisation; the boundaries are fussy, and the main idea is to have a comparative diagram to understand the influence of the procurement model on the different impacts.

	High effect of PPP			Medium/low effect of PPP			No effect of PPP		
NGO (head director)	Surrounding physical habitat	Community habitat	City integration	Technical house quality					
	Self-organised community	Link between volunteers and families		Reduce or close slum					
NGO (technical department)	Localisation heterogeneity	Community organisation	House size standard	Project closeness to services	Level of densification	Time	House growth capacity	Integration of surrounding community	
	Capacity building of technical area								
Housing and Urban Ministry (technical entity)	Quality of the EGIS			Quality of the Contractor	Integration (location and densification)	Gap between used regulation and new one	Within the regulation	House use (not for renting)	House functionality
Council	Guidance of private entity			Efficiency in public resources	Efficiency in land use	Neighbourhood generation	Conflict between neighbours	Capacity to deliver social housing	Ownership sense form the city
							Community leader development	Urban design	
Community leaders	Community organisation			Time	House size and materiality	Each family satisfaction	Life change (calmness)	Co-property	
Surrounding neighbours				Organisation of new neighbourhood	Infrastructure quality	Quality of people (in the new project)	Privacy with the existing neighbourhood	Social housing land use	Drug dealing
							Safety	Level of integration	Disturbs during construction

Figure 7-7: Stakeholder criteria separated by the influence of the PPP model (Social housing test case)

[Source: Author's]

The holistic view of the project has a general good valuation when observing qualitatively all the criteria in Figure 7-7, as most of them were judged positively by their stakeholders. Time is the only criterion that is valued as negative and that is within the effects of the PPP. This criterion was discussed in the negative criteria analysis, concluding that the causes of the impact are diverse and cannot be assigned to a specific responsible.

The PPP effect analysis shows that the procurement method employed in this project has generated mostly positive impact to their stakeholders, having one negative impact (time) that is difficult to assess. Therefore, this analysis suggests the procurement method (PPP model) has generated positive impacts to the project-system.

Stakeholder judgement principles

The stakeholder judgement principles are the cornerstone of the conceptual framework 'Project Success Evaluation Pyramid Model'. Understanding a project as a system of stakeholders, these judgement principles are considered the key characteristics of a successful project-systems. This means that they provide good support for stating that a project-system is successfully working. These principles are willingness for repetition, no excessive negative outcome, and no evidently better project alternative.

One of the components of the workshops was to gather the stakeholders' judgement over those principles. The incentive-based stakeholder map (Figure 7-8) is employed to show the results of this judgement; generating a holistic view. Stakeholders that were not included in the workshop are also part of the diagram, to show the limitations of the results, and to avoid any misinterpretation of them. It is possible to suggest that all the stakeholders that were part of the evaluation through the workshops are willing to repeat the project, they think no stakeholder was excessively negatively impacted, and an evident better alternative option to the project was not possible to state; as the park alternative is not considered sufficient, nor evidently better.

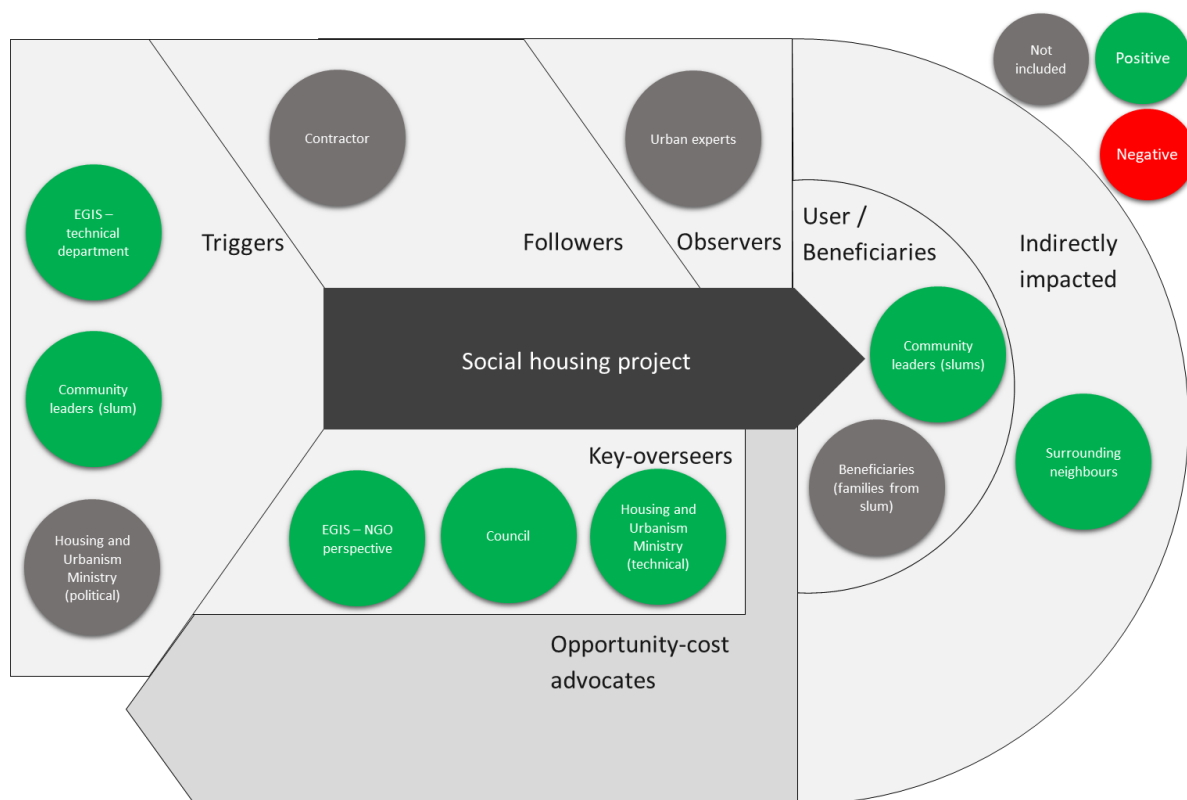


Figure 7-8: Incentive-based map with stakeholder judgement over the three judgement principles 'willingness for repetition', 'no excessive negative outcome', 'no evident better alternative' (Social housing test case).

[Source: Author's]

7.2.4. Test case 1 evaluation conclusion

Table 7-1 summarised the evaluation process for test case 1. It shows the identified stakeholders, the data gathered from the workshops and the analyses conclusions.

Table 7-1: Social housing test case evaluation summary

Project	
Social Housing PPP project	98-swelling social housing project for mostly informal settlement families; with an estimated budget of MMUSD 1.5. Practical completion in 2010
Identified and included stakeholders	
EGIS - NGO perspective (included)	Council (included)
EGIS - technical department (included)	Community leaders (slum) (included)
Contractor	Beneficiaries (families from the slum)
Housing and Urbanism Ministry (technical entity) (included)	Surrounding neighbours (included)
Housing and Urbanism Ministry (political entity)	Urban experts
Workshop data collection	
1) Stakeholder's definition	4) Evaluation criteria (critical and non-critical)
2) Stakeholder's aim towards the project	5) Assessment of the evaluation criteria
3) Key motivations and drivers of the stakeholder	6) Judgment of stakeholder principles
Workshop analysis	
Critical criteria analysis	All critical criteria are assessed as positive
Negative criteria analysis	Safety, time, and disturbs during construction were not considered highly relevant
Missing perspectives	Contractors profit negative impact is presumably a self-problem. The strategic view is considered positive in terms of opportunity cost, and inconclusive in terms efficiency (PPP analysis is missing)
PPP effect analysis	Most of the criteria considered under the PPP effect are positive, and the analysis of time (the only negative criterion) is not conclusive.
Stakeholder judgement analysis	All the reached stakeholders have a positive judgement towards the three principles. This is a strong argument to support a successfully working project-system

[Source: Author's]

Following the evaluation method and its procedure, the evaluation should have three main conclusions: (1) about the project-system, (2) about the procurement method, and (3) about the objective of the evaluation. This conclusion structure was developed as an observation in test case 3 (coded as HP6) and applied to all test cases.

The project-system social housing project has 10 identified stakeholders and 6 of them were included in the evaluation. The project-system generally meets the three principles that inform a successful project. First, there is a willingness for repetition from all the identified stakeholders. Second, the evaluation suggests that there is minimum fairness, which means that no stakeholder received any evident excessive negative impact from the project when analysing all the gathered criteria from the stakeholder workshops. Third, the alternative option to the project, suggested by the surrounding neighbourhoods, is a park that was initially planned for that site, but it was not implemented and is not sufficient argument against the social housing development.

The project was procured through a PPP model. The identified project impacts that were considered by the researcher as influenced or 'an effect' of the PPP model are all positive impacts except for the time of delivery, which is (at least) questionable as a direct effect of the PPP model. Even though the PPP analysis is not thorough (for this test case), it suggests a positive general impact.

The objective of this evaluation is the last conclusion. The project evaluation was done in this case for academic purposes; thus, it does not follow the four general aims which are: transparency and accountability, PPP legitimisation, industry learning, and government agency learning. However, it can be stated that the project evaluation provides lessons learned and it also should inform strategic decision-making.

7.2.5. Refinement insights from test case 1

During the application of the evaluation method in this test case, several issues were recognised by the researcher as insights for refining the evaluation method. These issues are part of the design process and are incorporated in the final version of the evaluation method described in Chapter 8. The refinement insights are:

- (SH1) the evaluation method is also applicable to PPP projects that do not have an evident profit driver;
- (SH2) stakeholders must provide reasonable expectations and not excessively optimistic ones;
- (SH3) the facilitator needs essential skills for conducting a workshop;
- (SH4) 'perceived value' is better understood as 'actual value' when explained to the stakeholders in the workshops;
- (SH5) re-adjust the workshop structure to be conducted under 1 hour;
- (SH6) development of the data visualisation templates,
- (SH7) development of lifecycle stakeholder diagram, and
- (SH8) development of data analyses.

SH1: The evaluation method is also applicable to PPP projects that do not have an evident profit driver

This test case demonstrated that the evaluation method was possible to be applied to a project with no evident profit driver incentive which is considered a key attribute of PPP projects. This refinement insight is acknowledged.

The existing PPP literature describes the different incentives that generate the profit driver of the private entity. Bundling is the incentive to address a whole-of-life approach due to contract arrangements (designer, contractor and facility manager under the same goal); and their impacts such as efficiency, time and quality occurs when the PPP is well designed (as explained in Chapter 2). However, the selected test case is developed by a not-for-profit organisation, which does not have an evident profit driver.

For conducting this test case, workshops were held with two stakeholders categorised under private sector: NGO (corporate level) and NGO (technical level). The NGO has the aim of eradicating slums from Chile and that is why it became a certificated private entity to develop social housing projects using public subsidy. When assessing the social housing project, none of the stakeholders made any reference to criteria such as efficient investment, Value for Money, budget compliance, profit, etc. Their concern was mostly on quality and satisfaction. After the data analysis was conducted and the conclusions were discussed, it is possible to suggest that the evaluation method was able to provide a reasonable conclusion about the project, regardless of the low relevance that the NGO put on financial criteria. This characteristic of the test case stresses the evaluation method. This test case shows the capability of the evaluation method to assess a pure social project.

SH2: Stakeholders must provide reasonable expectations and not excessively optimistic ones

The evaluation method has to include in its documentation that stakeholders must declare their 'expectations' regarding the project 'within a fair spectrum' and 'in the most objective way', preferably based on supporting documents such as contracts or existing written material. The 'expectation' concept is further explained.

The stakeholder workshops are for gathering a stakeholder perspective that includes the criteria for assessing the project and the judgements on those criteria. The judgement on each criterion should be supported with a declaration of what is the expectation and the perception over that criterion, e.g. if the selected criterion is time, then an expectation could be '4 years' and the perception (real value) could be '8 years'. In this case, the stakeholder will probably judge the 'time' criterion as a 'negative impact' supported by the difference between expectation and perception. Ideally, the stakeholder should provide convincing evidence that, for example, '4 years' was a real expectation because there are records that show the original timeline, or by quoting written material on the issue.

The expectations on every criterion need to be declared as a reasonable estimation that is possible to achieve. In this test case, the expectations of the NGO were to generate a much better result in terms of house size and quality that was not reasonable with the existing budget or the current social policy. This confusion between realistic expectations and 'dreamed' expectations took place in one of the workshops. The facilitator had to explain that the expectations must be calibrated with the real possibilities of normal work and resources. For the case of the NGO, they were very emphatic to say that the expectations will always be high and almost impossible to achieve. Here, the facilitator suggested that, according to this evaluation method, the expectations should be placed at a range in which normal results are achievable. The skills to manage this process appeared to be very relevant, and this issue suggests that it is not possible to follow the method without a facilitator guiding the stakeholder workshop.

SH3: The facilitator needs essential skills for conducting a workshop

The workshops conducted in this test case made evident that the facilitator guiding the workshops need many skills to keep the discussion about the stakeholder perspective within the boundaries of rationality. This insight is included in the evaluation method by recommending (in the documentation) that workshops must be conducted preferably in-person and with sufficient understanding of the theory behind the evaluation method.

The many components of the stakeholder workshop are not easy to understand, and the role of the facilitator is confirmed as relevant. The facilitator needs to (i) discuss and provoke the stakeholder to provide a reliable perspective. Therefore, it needs to synthesise the ideas, arguments, and examples that are provided by the stakeholder representative. Then, the facilitator needs to confirm what the stakeholder was trying to say and make the information to 'fit' into the evaluation structure. During the stakeholder workshops, it was common to receive misleading answers. For example, for the question about the main (core) functions of the stakeholders, the answer was about the main challenges that the stakeholder had to face. Consequently, the facilitator (in this case the same researcher) had to explain and suggest that the 'reasons for facing that challenge were because they have certain specific core functions'. The importance of the facilitator suggests that the workshop should be conducted preferably in-person, rather than through a skype/remote meeting.

SH4: 'Perceived value' is better understood as 'actual value' when explained to the stakeholders in the workshops

In the stakeholder workshop, 'perceived value' was replaced by 'actual value' so stakeholder could better understand the concept. This modification was included in the subsequent workshops for all test cases and employed in the final evaluation method.

During the interviews with experts and practitioners and during the stakeholder workshops the word 'perception', as the contrast to 'expectations', was not well understood. According to the participants, the word perception implies a subjective view of the facts, rather than the objective truth. However, according to the employed theory, it is impossible to claim that something is completely true, even with the best available quantitative support; that is why the word 'perception' was initially employed. Nevertheless, the word 'perception' was changed to 'actual value' as it is implicitly understood that the perception is based on reasonable support.

SH5: Re-adjust the workshop structure to be conducted under 1-hour

The time required for each workshop was estimated in 1 hour and throughout the workshops, several strategies were tested by the researcher for achieving an effective stakeholder workshop capable of capturing all the necessary information. A better structure was tested, applied to the other two test cases, and documented in the final evaluation method.

As it was stated in the prototyping exercise (coded as PE4), a 1-hour workshop was considered reasonable as it enables to gather enough information and it is an amount of time that most stakeholder representatives can accept. However, one hour does not allow any mistake, and the impacts of delays affect the last part of the evaluation, which is the answer to the judgement principles. Throughout the workshops, the researcher learned that the introduction must be very brief, only naming the four components of the evaluation, and then, start the data gathering straight away. Further information about the aim of the method and the theory behind is not necessary. As the method is new, the stakeholders can understand it along with the workshop. At the end of the workshop, a brief explanation of how the gathered information will be contrasted with the other stakeholder workshops gives a good overview of the theory behind the evaluation method.

The structure modification was considered appropriate for the other test cases and for the final evaluation method. As explained in earlier chapters, the three test cases represent the existing diversity of the PPP projects around the world. Therefore, this evaluation assumes some common restrictions such as stakeholders' lack of time.

SH6: Development of the data visualisation templates

The idea for presenting the data with colour-coded boxes to show the stakeholder criteria and their valuation of the criteria was designed after test case 1. The design is included in the final evaluation method, including, as examples, the results of the test cases.

Coloured boxes to represent the criteria valuation is an improvement of the original data collection database designed as a first draft. The new format is aligned with one of the purposes of the evaluation, which is the transparency and simplicity for general audience understanding. The basic structure of the data visualisation shows the stakeholders, their selected criteria, and the valuation of those criteria under the scale positive, neutral, or negative. This criteria representation is what was conceptualised as stakeholder impacts.

SH7: Development of lifecycle stakeholder diagram for stakeholder identification

For the stakeholder identification process, the lifecycle stakeholder diagram was developed to show that the stakeholders can be relevant in certain parts of the project lifecycle. This view contributes to the judgement whether to include them in the evaluation or if it is necessary to address the same stakeholder in different time frames (e.g. the same public authority from different governments periods). The lifecycle stakeholder diagram is applied to the three test cases and described in the final version of the evaluation method.

SH8: Data analysis development

Before gathering the data for test case 1, there was only a broad idea on how to analyse the data gathered from the stakeholder workshops. At this point, testing the data visualisation allowed to design structured analyses with specific purposes. Critical criteria analysis, theme analysis, negative criteria analysis, pro project analysis, and PPP effect analysis were developed. All these analyses are part of the full evaluation method documentation in Chapter 8.

Summary of test case 1 contribution to the evaluation method development

Test case 1 contributed to improving the data collection process by refining the workshop structure to fit in 1 hour, adopting the concept of 'actual value', refining the facilitator skills, developing the lifecycle stakeholder map, and including the requirement of realistic expectation from the stakeholders. The data collection improvements were applied to test case 2 and 3. To the data analysis process, test case 1 contributed mainly with the data visualisation and subsequent analyses which were applied later to test cases 1, 2, and 3 when the project evaluation results were developed.

7.3. Test case 2 - Transport PPP project

The result of applying the evaluation method to a transport PPP project is presented in this section. From this test case, eight refinement insights were synthesised by the researcher and incorporated to the final evaluation method.

The structure of the test case 2 is as follow. First, the context of the project is presented (overview). Second, the stakeholder identification process is described. Third, the analysis of the stakeholder workshops is presented using seven types of analyses namely, (1) consistency analysis, (2) critical criteria analysis, (3) theme analysis, (4) negative criteria analysis, (5) quantitative analysis (6) PPP effect analysis, and (7) stakeholder judgement analysis. Fourth, the conclusion of the evaluation is presented. Finally, the eight refinements insights are described.

7.3.1. Test case 2 overview

The context of the urban toll road PPP project

The urban toll road opened in 2005 is located in the east-west axis of a Chilean city mostly along its main river. The highway is 42 kilometres long and connects the high land in the east with the airport on the west side of the city and with other cities in the coastal region. It has a tag system, first in Chile to work with a free-flow method. The project is being developed through a PPP model, with an initial investment of USD 384 million.

The milestones of the project since the first draft of the highway layout started in 1958, in which a new metropolitan regulatory plan was presented, including the highway in the area near the river. In 1991 the PPP policy was presented and in 1995 the procurement of the highway was proposed. In 1996 the first expression of interest (EOI) call for the project was conducted. Between 1996 and 1999 the conceptualisation stage occurred, including (1) the postponement of the tendering process, (2)

negotiations with the different councils that were affected by highway layout, (3) the opposition of community organisations with legal actions, and (iv) the announcement of the winning bid. The construction stage started officially in 2001 with a layout change to build the highway under the river, mainly to reduce the intervention over the city. From 2005 to 2007 the project was opened to the public by stages. Improvements to the highway, better connections and design modifications to adapt to new needs are being part of the project until today (2018) with a recent opening of a tunnel that connects to the highway in a highly congested central area (Allard & Nazar, 2019).

The impact of the highway over the city is discussed by Allard and Nazar (2019), describing the urban tension between (1) creating a highway that can expand the urban area of the city and the freedom of its citizens, and (2) the opposite argument that the highway generates an induced demand by changing the citizens' behaviour to drive more distance, generating a bigger traffic problem. There is an agreement that the highway expanded the land market and the goods market within the city. The cost of congestion is the problem that urban theory is still discussing (van der Loop, Haaijer, & Willigers, 2016).

The project original budget was 400 MM approx. and end it up with a total cost of more than USD 1200 MM. The cost of the project and its benefits are discussed from partial views, mainly economical and urban. Allard and Nazar (2019) state that cost overrun, the high public subsidies and the general problems with the procurement process do not overweight the benefits that the project has provided to the city and to the national economy. This statement is presented as a 10-year overview since the project's inauguration.

The media has covered the project mainly because of the major disruptions in certain parts of the city, which are relatively minor compared to the size of the project, as it was built under and next to the river. The media also covered the project profitability, which is considered over the expectations for a PPP project in Chile (Sepulveda, 2016).

The context of the evaluation method tested in the test case

The toll road project was delivered, and it has been operating for more than 10 years. The prototype evaluation method (with some refinements from test case 1) was applied to this PPP project, which represents one of the most common types of PPP, contributing to the required diversity for testing the developed evaluation method. Some evaluation formalities were not included in this test case because the objective at this stage is to test and design the main components of the evaluation. Thus, steps such as the declaration of the evaluation requester, or the analysis of the best time-period for evaluation were not applied.

7.3.2. Test case 2 stakeholder identification

The identification process for this case test was conducted using two tools developed as refinements insights: the lifecycle stakeholder diagram (coded as SH7) and the incentive-based stakeholder map (coded as TP5).

Figure 7-9 shows the lifecycle stakeholder diagram applied to test case 2. The researcher judged in which part of the project lifecycle the stakeholders were relevant and selected a stakeholder representative for each of them. Some of the stakeholders were not reached, and they appear in Figure 7-10 as 'not included'. As explained in test case 1, this is not a problem for the design process because the objective is to test the effectiveness and efficiency of the evaluation method, rather than to achieve a full project description.

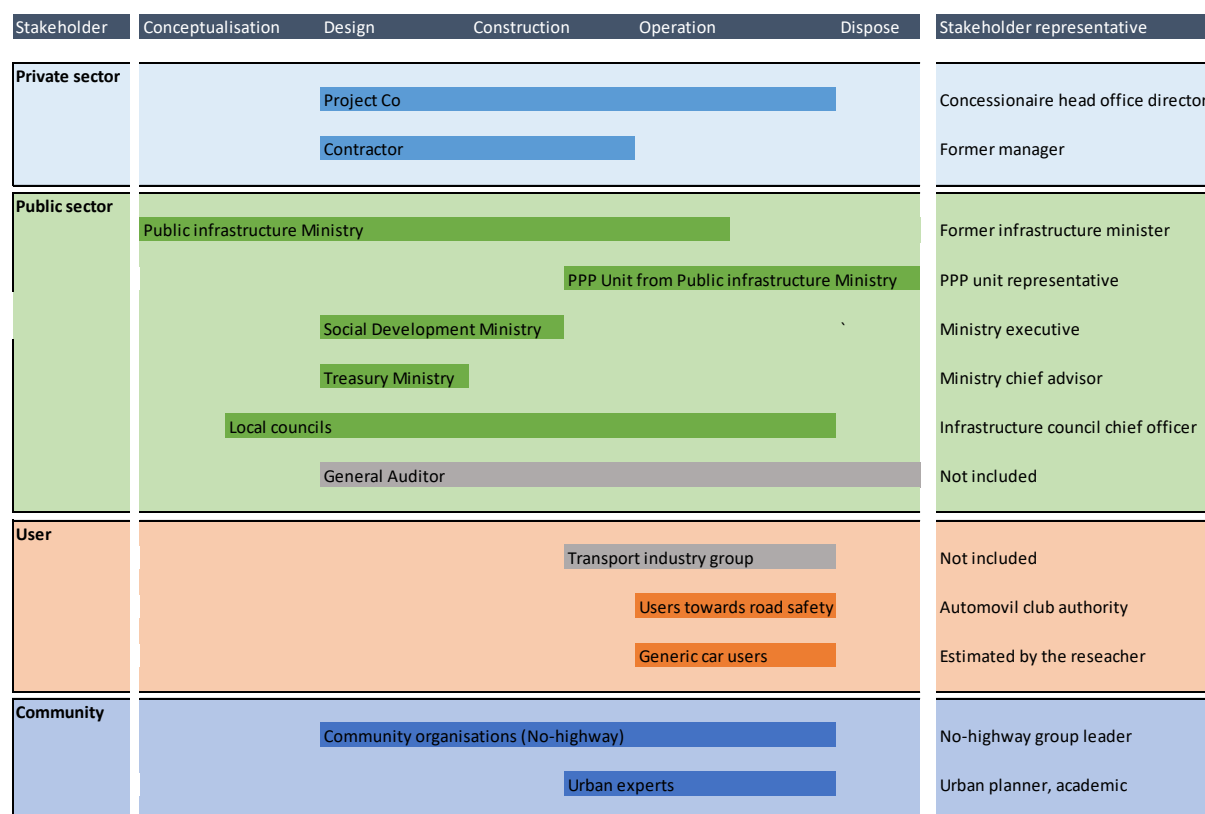


Figure 7-9: Project lifecycle stakeholder diagram (Transport test case)

[Source: Author's]

Similar to test case 1, Figure 7-10 shows the incentive-based stakeholder map, which is based on the project-system concept that understands a project as a set of implicitly connected stakeholders to pursue an endeavour.

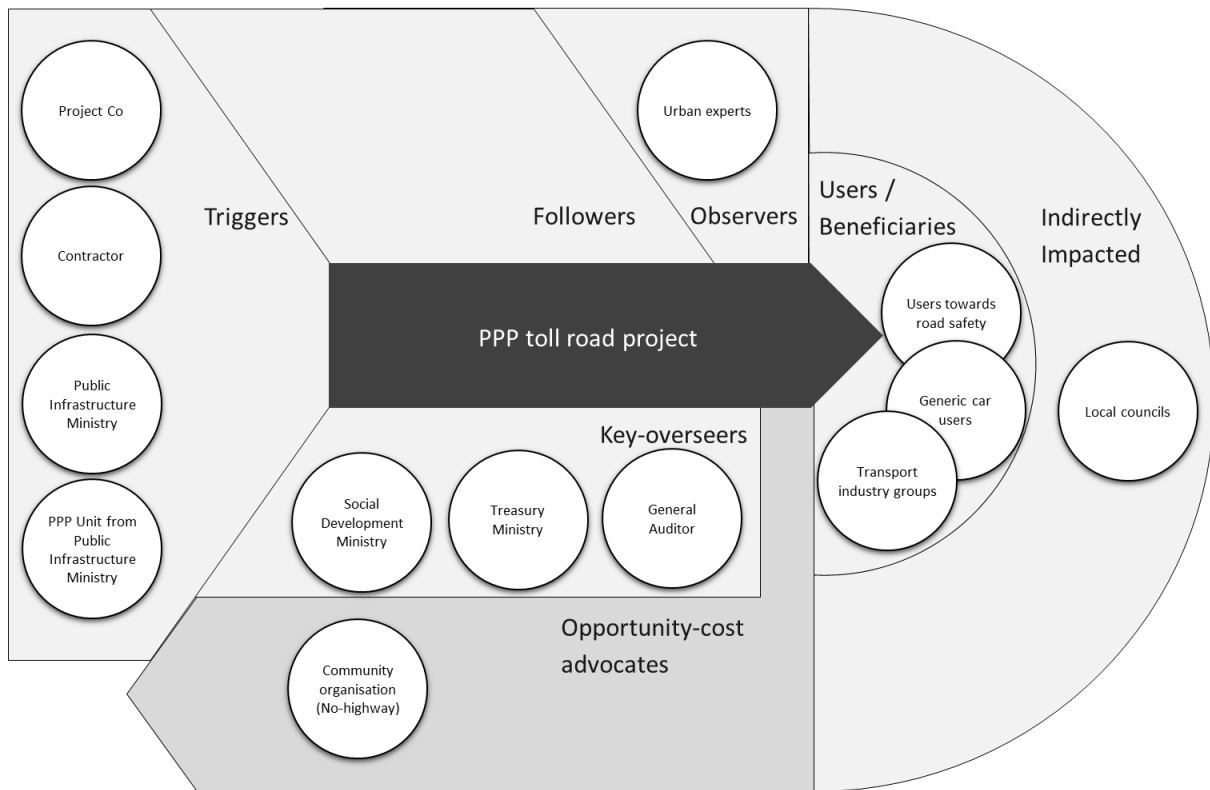


Figure 7-10: Incentive-based stakeholder map with the identified stakeholders (Transport test case)

[Source: Author's]

The next step was to conduct the stakeholder workshops using (1) the stakeholder workshop forms developed in the prototyping exercises (coded as PE8), and (2) the refinement insights from test case 1 to make the workshop more effective under 1-hour (SH3, SH4, and SH5). The analysis of the data obtained from the stakeholder workshops is presented in the next section.

7.3.3. Test case 2 stakeholder workshop analysis

The components gathered from the stakeholder workshops in test case 2 have no difference from the ones gathered in test case 1. The six components that formed a stakeholder perspective were:

1. Stakeholder's definition (who is the stakeholder?),
2. Stakeholder's aim towards the project (why is it relevant to the project?)
3. Key motivations and drivers of the stakeholder (what drives stakeholder's actions?)
4. Evaluation criteria (which criteria the stakeholder uses to assess the project?)
5. Assessment of the evaluation criteria (assess/valuate the selected criteria)
6. The judgment of stakeholder principles (principles defined in section 5.5: willingness for repetition, no excessive negative outcome, no evident better alternative)

The researcher guided the workshops taking a more evident facilitator's role (discussed in SH3). There was an improvement in this aspect from test case 1 because of the gained experience. However, test case 2 was more challenging, as the project is more complex. It has more exposure to the public, it has received more critics, and there are more political and ideological influences. There were stakeholders with completely opposite views and different assessments for the same (or similar) project outcomes.

For this test case, seven analyses were conducted to the data: (1) consistency analysis, (2) critical criteria analysis, (3) theme analysis, (4) negative and missing criteria analysis, (5) descriptive statistic

analysis (6) PPP effect analysis, and (7) stakeholder judgement analysis. The PPP effect analysis was further improved in this test case by using partially the refinement insight 'TP4 taxpayers' perspective'. This refinement included a new data collection component; thus, it was fully applied only in test case 3.

Consistency analysis

The stakeholders were analysed in terms of their consistency for judging the project. This analysis was done for this test case because the project was considered controversial. This is a review of the coherency that stakeholders should have when declaring their perspectives. For example, the evaluation criteria they use to assess the project should be aligned with at least one core value (motivator/driver) they declared. From the six components that were gathered in the stakeholder workshops, four were used to analyse the consistency of the stakeholder perspectives. The components are (1) stakeholder's purpose towards the project, (2) key mission/motivators/core values/drivers, (3) project success criteria, and (4) judgement principles. Figure 7-11 shows an example of a consistency diagram to show the first consistency analysis within one stakeholder.

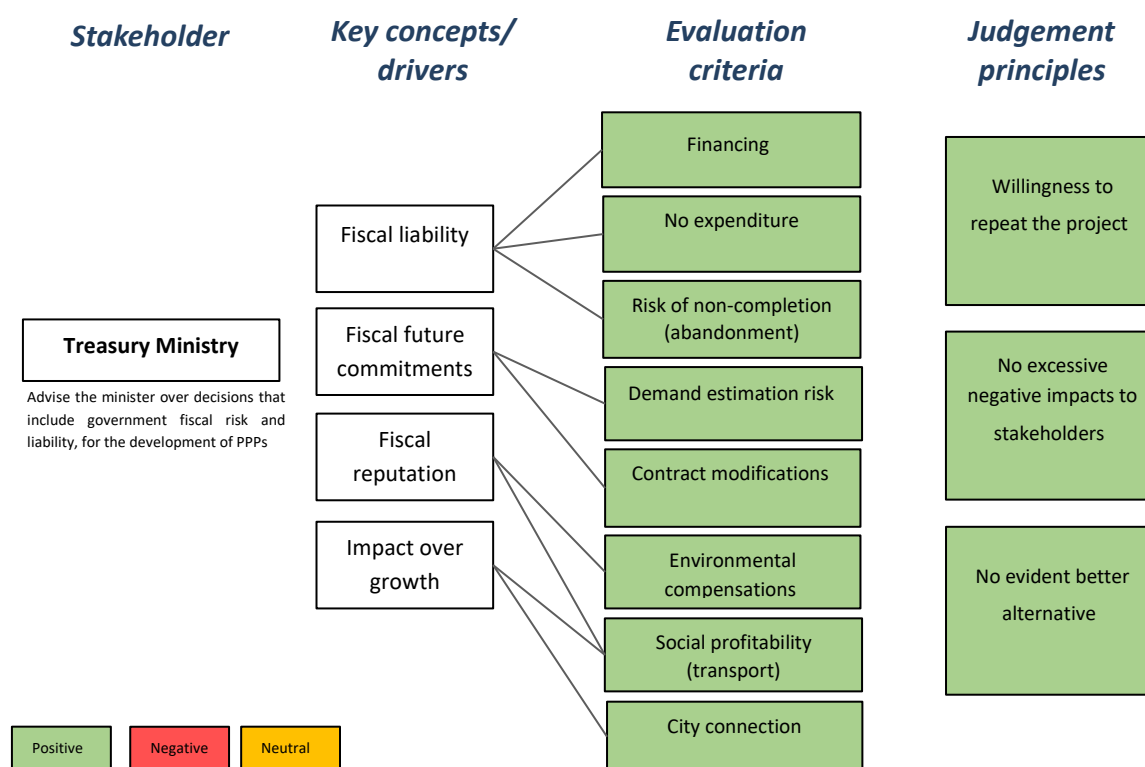


Figure 7-11: Example of consistency diagram for the Treasury Ministry (Transport test case)

[Source: Author's]

- The stakeholders' purpose towards the project: This is the first component that is gathered for each stakeholder and it is the reference for the consistency analysis; i.e. the consistency is judged by how the stakeholder is coherent with its purpose.
- key mission/motivators/core values/drivers: Stakeholders have key concepts that drive their functioning. For example, the private entity has a profit driver, the public entity has a social impact driver or 'efficient use of taxpayers money' driver. This component is employed as an overarching frame for the evaluation criteria. For each stakeholder, all their evaluation criteria should be linked to one or more key concepts. For this project, there are no inconsistencies in

the perspectives; all evaluation criteria are judged as reasonably aligned with the general concepts of each stakeholder.

- **Success criteria and judgement principles:** According to the evaluators, the valuation of the success criteria made by each stakeholder is consistent with the judgement they made over the judgement principles. This means that the data suggests the judgements are validly supported on their criteria valuation (i.e. general positive valuations of the criteria is aligned with a positive answer to the judgement principles).

Critical criteria analysis

Figure 7-12 shows the critical and non-critical criteria from all the stakeholders. Most of the critical criteria were judged as positive for each stakeholder, except (1) business model, which is considered neutral according to the Public Infrastructure Ministry, and (2) city integration, which is considered negative according to the activist group 'No-highway'. Both criteria will be developed in the negative impact analysis.

This analysis does not show clear support that the project-system is successfully working, as core expectations of different stakeholders are considered not positive. As explained in test case 1, the approach of this evaluation method is to consider a project-system to be successfully working unless there is a solid argument that suggests the opposite. In this case, the analysis does not provide any strong argument.

	Critical criteria		Non-critical criteria									
Project Co	Profit		Social profitability	Relationship with PI ministry	Collateral satisfaction	Time	User satisfaction	Highway connectivity with the system				
Contractor	Profit by construction	Profit as part of the consortium	Value of engineering component	Time	Quality of project team and partners	Public private relationship						
Public Infrastructure Ministry	Highway utilization	Business model	Infrastructure quality	Intervention during construction	Land value surrounding highway	Integration to the city	Private profit	Expectations over highway's impact				
PPP unit at the Public Infrastructure Ministry	Pay per use	Service availability	Service quality	Quality standard	Safety	Connectivity	Travel time	Benefit to users	Positive business model			
Social Development Ministry	Social NPV-IRR		Public participation	Sinergy with the city	Macro-design for all highways	Environmental costs	Renegotiations	Connection highway-city				
Treasury Ministry	Social profitability (transport)		City connection	Financing	No expenditure	Demand estimation	Impact on economic growth	Risk of non-completion (abandonment)	Contract modifications	Environmental compensations		
Local council	Quality of access (to highway)		Connection with other areas of the city	Maintain capital gain (property)	Value creation in surrounding area							
User towards road safety	Accident rate		Safety design and information for users	Car-pedestrian relationship	Relieve existing local traffic	Connectivity						
Generic user	Travel time reduction		safety	Connection with the city	Price							
Community organisation (no-highway)	Urban integration		Creation of social capital	Impact of participation	Effect over car use/transport	Beneficiaries	Effect over urban connectivity					
Urban experts	Project purpose	Adaptation to geographical characteristics	Generate landscape	Generate public spaces	Synergy with mobility	Public-private-people interaction	Experience the city	City connection				

Figure 7-12: Evaluation criteria for project stakeholders separated by critical and non-critical (Transport test case)

[Source: Author's]

Negative criteria analysis

Similar to test case 1, the negative criteria analysis is applied to this case for analysing in detail the main negative impacts that were gathered from the stakeholder workshops. Three criteria are recognised as probable deal-breaker for the project; these three criteria could support the argument

that the project-system is not successfully working. They are business model, urban integration, and connection to the city.

- Business model: The business model is considered a negative impact by the Public Infrastructure Ministry. The reasons are (1) private profits over the expectations, (2) excessive direct subsidies due to scope changes, and (3) the service-fee pricing system. This criterion is probably the strongest argument against the positive judgement of the project. The highway was the first urban PPP project in the country under the new 1991 PPP regulation. The result is a project with a high innovation, delivered with private resources in a time which public resources were not available for this kind of mega-project. However, the media has pointed out estimations of the incomes, suggesting the Project Co is recovering many times the investment and therefore, getting profits over the expectations. Due to the changes in the design and to make the project more desirable to the private sector, direct subsidies to the project were transferred, questioning one of the principles of the PPP which is to be a filter for white elephants (section 2.2). Finally, the fee price has a demand-based strategy, which means that it depends on the number of cars that use the highway. If the demand is low, the tag-price is low (per km), if there is traffic congestion, the price increases to reduce congestion by market principle. Due to the high traffic on the highway, the highest price has been employed excessively, affecting the users and their expectations about the private profit.
- Urban geographical segregation: According to the No-highway activist group, the main impact of the highway is the urban segregation that the highway generates, causing a 'scar' in the city that divides the city in north and south. According to this stakeholder's representative, the option would have been to do no project, which can be related and placed in the view of the compact and walkable city that is well recognised in the urban academic sphere. This negative impact needs to be considered a 'project impact' and not a 'procurement option impact'. This means that with any kind of procurement system, the impact would have been the same or similar. Two arguments can reduce the relevance of this claim for the do-nothing option. First, the general layout of the east-west highway was first introduced in 1958 in the metropolitan plan, together with the city encirclement and the north-south highway, which sounds 'at least' reasonable from a macro city perspective. Second, the claim of this division/segregation of the city due to the highway can be considered an increase of the natural segregation that the same river has over the city, dividing it into north and south. Of course, a river is not the same as a highway and this argument is not to diminish this difference, but it must be stated that rivers do divide cities.
- Connection to the city: The bad connection of the highway to the city (streets) was mentioned by several stakeholders as a negative impact. The highway design, which includes the exits to the city, could be part of the problem as they are not capable of managing the current demand level of the highway. Conversely, the problem could be in the metropolitan transport plan, which should address transport times and traffic jams taking into account all the available infrastructure of the city. The public policy should also estimate future demand and invest accordingly. It needs to be recognised that the project scope ends at the highway exits. Increased demand and growth of the city in urban, economic, and wealth terms has created traffic jams in many exits of the highway. Even though this could have been planned, forecasted, and mitigated, it is also possible to state that it is a problem in the entire transport system. The highway, with an explicit and clear scope, is only an important component of this transport system. The strategic transport policy of the city should be the main responsible for the demand that is being concentrated on the highway. However, the counterargument is that the traffic jam is due to the induced demand that any highway can create, and the

concentration and delays in pick hours are intrinsic results of any similar infrastructure (van der Loop et al., 2016).

Theme criteria analysis

This analysis was developed as a refinement insight of test case 1 (coded as SH8). It was applied only to this test case because it was not considered valuable; it repeats most of the conclusions from the other analyses and is not so clear. However, it was part of the development process, it was tested in this test case and it was also included as a possible analysis in the final evaluation method as part of the 'toolbox'.

The aim of this analysis is to create themes for classifying the gathered criteria (Figure 7-13). Thus, search for possible clusters, incompatibilities, conflicts, or inconsistencies between the different stakeholders. This includes different valuations of the same criteria for more than one stakeholder (e.g. negative to one stakeholder and positive to another). Each theme is analysed separately.

	Direct highway purpose		Connection to city (in transport)	Business / finance		Urban / social	Process	Indirect impacts	
Project Co	User's satisfaction	Social profitability	Highway connectivity with the system	Profit	Time		Relationship with PI Ministry	Collateral satisfaction	
Contractor	Value of engineering component			Profit by construction	Profit as part of the consortium		Quality of project team and partners		
				Time			Public private relationship		
Public Infrastructure Ministry	Highway utilization	Infrastructure quality	Integration to the city	Private profit	Business model		Intervention during construction	Land value surrounding highway	Expectations over highway's impact
PPP unit at the public infrastructure ministry	Service quality	Quality standard	Connectivity	Pay per use					
	Service availability	Benefit to users		Positive business model					
	Travel time	Safety							
Social Development Ministry	Social NPV-IRR		Connection highway-city	Renegotiations		Public participation	Sinergy with the city	Environmental costs	
Treasury Ministry	Social profitability (transport)		City connection	Financing	No expenditure	Macro-design for all highways		Environmental compensations	Impact on economic growth
				Risk of non-completion (abandonment)	Contract modifications				
				Demand estimation					
User towards road safety	Accident rate	Safety design and information for users				Car-pedestrian relationship			
	Connectivity	Relieve existing local traffic							
Generic car users	Travel time reduction	Safety	Connection with the city	Price					
Community organisation (No-highway)	Effect over car use/transport	Beneficiaries	Effect over urban connectivity			Creation of social capital	Urban integration	Impact of participation	
Urban experts	Project purpose		City connection			Generate landscape	Generate public spaces	Public-private-people interaction	
						Experience the city	Capacity to adapt to geographical characteristics		
						Synergy with mobility			

Figure 7-13: Stakeholder criteria categorised by theme (Transport test case)

[Source: Author's]

The 'direct purpose' of the project is to generate an east-west transport route for vehicles in the Chilean city. The criteria related to this main purpose is generally valued as positive by all stakeholders. However, two criteria are considered relevant to analyse. First, the No-highway stakeholder argues that the project does not promote public transport and generates induced demand. This perspective has been mentioned before and it is supported by urbanists that aim for a compact and walkable city. However, a counterargument for the public transport claim is the fact that the metro line goes parallel to the highway, especially in the central district of the city. It is possible to suggest that the highway and metro line complement the private and public transport, or at least it is a matter of discussion. Second, user satisfaction is considered by the concessionary a negative impact on them because there is highway overload. However, they claim this issue is beyond the scope

of the project (highway), and it is a city-scale problem to be solved by a strategic public policy and not by a single project.

The city connection theme refers to how the highway connects with the existing transport infrastructure of the city. Even though the judgement of the different stakeholder regarding connection to the city is diverse, there is sufficient support to state that the connection of the highway to the city is an overall negative impact. However, is not completely clear if the problem remains in the project design, or in the public policy that supports the project.

In the business and finance theme, the business model and the service fee price are considered negative impacts for the public sector, and for the users respectively. This issue is discussed in the negative criteria analysis.

The urban and social impact theme is mostly positive except for the activist group No-highway which valued 'urban integration' criterion as a negative impact. This criterion was already discussed under the negative criteria analysis.

The process theme that refers to the design and construction phase is mostly positive, mainly due to the strategy of building the highway under the river that reduced the disturbance to the city, and the satisfactory processes in terms of public-private relationship and environmental compensations. The impact of participation is considered a neutral impact for the No-highway group as (in their view) they were able to change the highway layout and have an influence on the project design, however, not up to the desired point of cancelling the project.

The indirect impact theme has mostly positive impacts according to their stakeholders with an increased in the land value, economic growth, and a general collateral satisfaction perceived by the concessionaire. One negative impact was that (according to the Public Infrastructure Ministry) excessive expectations were generated for how the highway was going to solve the transport problem of the city, which was not achieved completely, disappointing users and beneficiaries.

Descriptive statistical analysis (Quantitative analysis)

A quantitative analysis shows the percentage of positive and negative impacts, critical and non-critical within the project system. The assumptions for the analysis are: (1) all stakeholders are equally important in the project-system, (2) all criteria are equally supported with reasonable documentation, and (3) all criteria represent a legitimate view of a stakeholder. This quantitative view complements the main qualitative analysis showing that most of the criteria and most of the critical criteria are considered positive impacts. However, further conclusions cannot be obtained with this analysis.

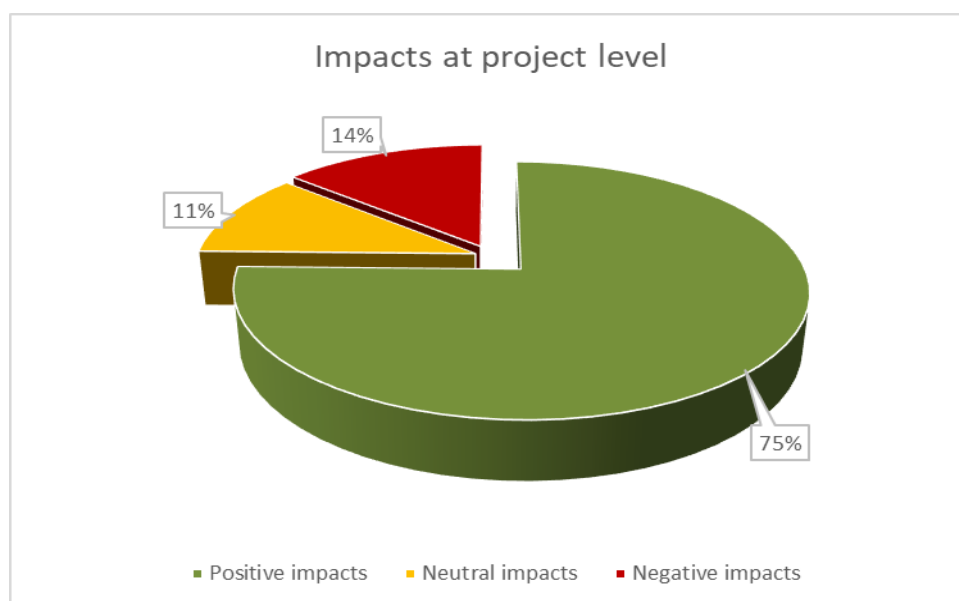


Figure 7-14: Descriptive statistical analysis of impacts at the project level (Transport test case)

[Source: Author's]

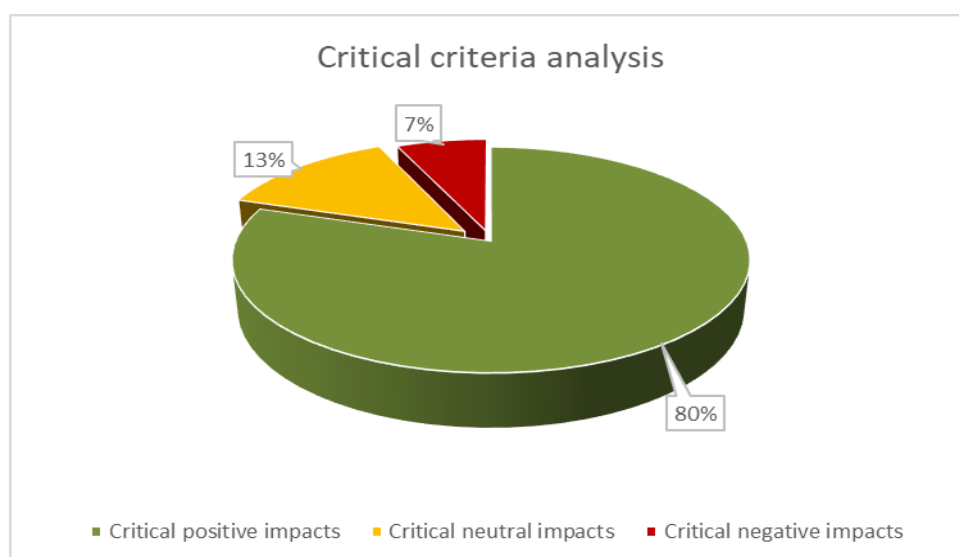


Figure 7-15: Descriptive statistical analysis of critical impacts at the project level (Transport test case)

[Source: Author's]

PPP effect analysis

The PPP effect analysis was also applied to test case 1. However, for test case 2, an improvement was made based on the refinement insight TP4 taxpayers' perspective. Apart from dividing the impacts between PPP related and non-PPP related (as in test case 1), a circle chart to show the number of criteria affected by the PPP model is also included to help its evaluation (Figure 7-17); and a structured argumentation is made to judge if the identified PPP problems can outweigh the (presumed) PPP benefits.

PPP projects are usually compared with its theoretical traditionally procured option (publicly financed, designed and maintained). This is related with the third judgement principle 'no evident better alternative', supporting a thesis that the PPP model was the best procurement choice. In this test case,

stakeholder representatives did not suggest that a possibly better option was to do the project through Traditional Procurement. However, for analysing the PPP model, this needs to be considered. According to the different stakeholders and the reviewed documentation, the capital investment for the project was not even considered to be leveraged with public resources. It cannot be known with certainty, but PPP experts suggested that this kind of investment was not possible to do in the Chilean context of the 90' and early 2000'. The option of taking a loan for investing in this kind of infrastructure was not present. The review shows that an alternative option was discussed for doing a smaller project with a Traditional Procurement (freeway). However, there is no certainty about the real scope, their effectiveness to the transport system, how it was going to be funded, or if it was a real option. It is possible to assume that the comparison is between doing the mega-project through a PPP or no project.

To ex-post evaluate the impact of the PPP model within the project, a stakeholder approach is followed. According to the evaluation procedure, the criteria gathered from the stakeholders are categorised (Figure 7-16) under (1) high effect of PPP, (2) medium effect of PPP, and (3) no effect of PPP. This classification was done by the researcher and is not intended to be a deterministic categorisation, the boundaries are fussy, and the main idea is to have a comparative diagram to understand the influence of the procurement model on the different impacts.

	High effect of PPP		Medium-low effect of PPP		No effect of PPP				
Project Co	Profit	Relationship with PI ministry	User satisfaction		Social profitability	Collateral satisfaction	Time	Highway connectivity with the system	
Contractor	Profit as part of the consortium	Quality of project team and partners	Profit by construction	Public private relationship	Time	Value of engineering component			
Public Infrastructure Ministry	Private profit	Business model	Expectations over highway's impact	Highway utilization	Infrastructure quality	Intervention during construction	Land value surrounding highway	Integration to the city	
PPP unit at the Public Infrastructure Ministry	Service quality	Quality standard	Pay per use		Safety	Connectivity	Travel time	Benefit to users	
	Positive business model	Service availability							
Social Development Ministry	Renegotiations		Public participation		Sinergy with the city	Macro-design for all highways	Environmental costs	Social NPV-IRR	Connection highway-city
Treasury Ministry	Financing	No expenditure	Risk of non-completion (abandonment)	Environmental compensations	Impact on economic growth	Social profitability (transport)	City connection		
	Demand estimation	Contract modifications							
Local council			Maintain capital gain (property)	Value creation in surrounding area	Quality of access (to highway)	Connection with other areas of the city			
User towards road safety	Car-pedestrian relationship		Accident rate	Safety design and information for users	Relieve existing local traffic	Connectivity			
Generic car user			Travel time reduction	Price	Connection with the city	Safety			
Community organisation (No-highway)	Impact of participation				Creation of social capital	Urban integration	Effect over car use/transport	Beneficiaries	Effect over urban connectivity
Urban experts	Public-private-people interaction		Generate public spaces	Adaptation to geographical characteristics	Generate landscape	Synergy with mobility	Project purpose	Experience the city	City connection

Figure 7-16: Stakeholder criteria categorised by PPP effect (Transport test case)

[Source: Author's]

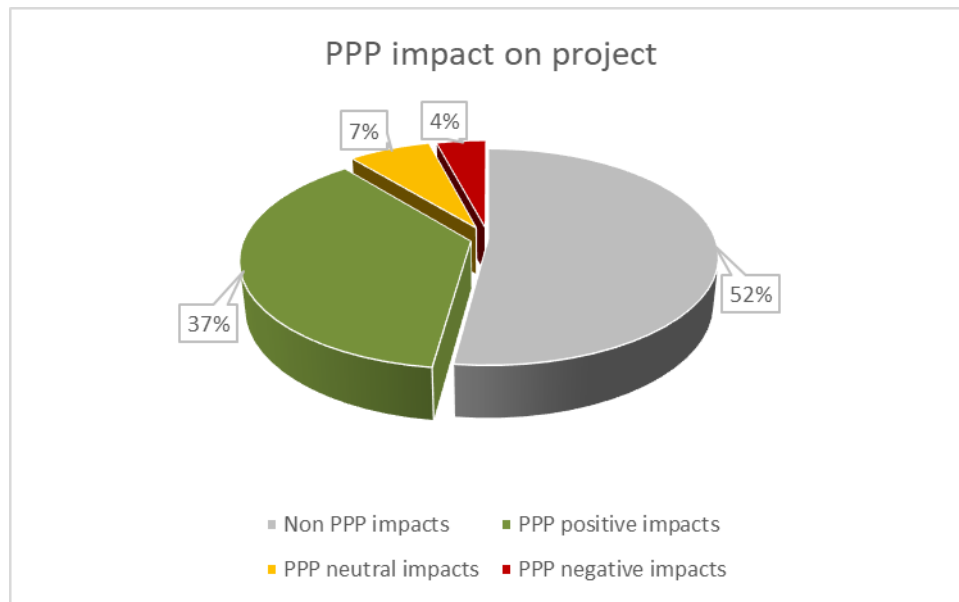


Figure 7-17: Descriptive statistic analysis of the PPP effect on the project (Transport test case)

[Source: Author's]

Nearly half of the gathered criteria are considered an effect of the PPP model. Most of the criteria are valued as positive impacts by the stakeholders. Five criteria that are considered an effect of PPPs are also critical: profit, pay per use, business model, service availability, travel time reduction. All valued as positive impacts, except 'business model'. The business model, in this case, refers to the procurement method and contract development that has resulted in (1) high direct public subsidy, (2) private profits over the expectations, and (3) incorrect service-fee price strategy. This criterion can be a deal-breaker to the argument of a successfully working project-system.

To analyse the business model deal-breaker criterion, this part of the analysis aims to judge if the (presumed) problems of the PPP model outweigh the benefits (if they exist) of employing this procurement method. The following equation is used from the refinement insight coded as TP4.

Equation 7-1: PPP problems and benefits

$$\text{value of (PPP benefits)} > \text{value of (PPP problems)}$$

From the gathered criteria, the main benefit of the PPP model is infrastructure availability. The study suggests that this project (with its magnitude) would have not been a feasible option for a Traditional Procurement method, involving public funding. It can be considered the first mega intervention of the city after the metro system. Therefore, it is possible to suggest that the decision was (1) through a PPP model, or (2) a mitigation strategy (no project option).

The judgement that needs to be placed is: does the availability of the project outweigh the negative (or insufficient) business model?

Table 7-2: Summary of PPP benefits and PPP problems (Transport test case)

PPP benefits	PPP problems
Infrastructure availability	Business model (direct subsidy, service-fee price strategy, private profits over the expectations)

[Source: Author's]

In this case, the non-availability of the infrastructure is considered a major problem compared to the negative business model because the project is understood as an alternative route, based on voluntary payments that add kilometres of road to the city. Theoretically, it is an optional route, paid only by the users that are willing to pay for the service. However, in practice, this is questioned because it has become the main access to the city from a large portion of the metropolitan area. Nevertheless, the argument of the highway being an additional alternative and voluntary to pay reduces the relevance of the negative business model impact. It is not certain that the negative business model impact is a deal-breaker, having the consequence of not building the infrastructure; as mentioned before, assuming that the project was not feasible with public funding. Therefore, this evaluation suggests that the identified PPP benefit outweigh the identified PPP problem under the explained assumptions.

This part of the analysis (taxpayers' perspective) was conducted with partial information, as it was implemented with the data that was already gathered for this test case. In test case 3, additional data was gathered to provide better support for judging the PPP benefit/problem comparison. For example, test case 2 has no information regarding the maintenance of the infrastructure, which could have been an additional benefit, considering that PPPs (theoretically) generate incentives for higher quality and more effective whole-of-life maintenance.

Stakeholder judgement principles

The stakeholder judgement principles were developed in the conceptual framework 'Project Success Evaluation Pyramid Model'. They are now applied to this test case.

The judgement for the three judgement principles is as follow. All the stakeholders, except the opportunity cost advocates (No-highway group), are willing to repeat the project; they think that no excessive negatives outcomes occurred to none of the stakeholders, and a better alternative was not possible to identify. Opportunity cost advocates (No-highway group) is the only identified stakeholder that thinks the project should not be repeated, that it generates excessive negative outcomes to the neighbours, and a that a better option would have been to do no project. The arguments that the community groups presented are considered legitimate and with a base on logic reasoning. However, they are not enough to invalidate the perception of all the other stakeholders that think differently and judge the outcomes (mostly) positively.

Figure 7-18 presents the incentive-based stakeholder map to show the judgement of the stakeholders regarding the three judgement principles. In this case, the three principles have the same responses from the stakeholder, so they are presented in the same diagram: willingness for repetition, no excessive negative outcome, and no evident better alternative principle.

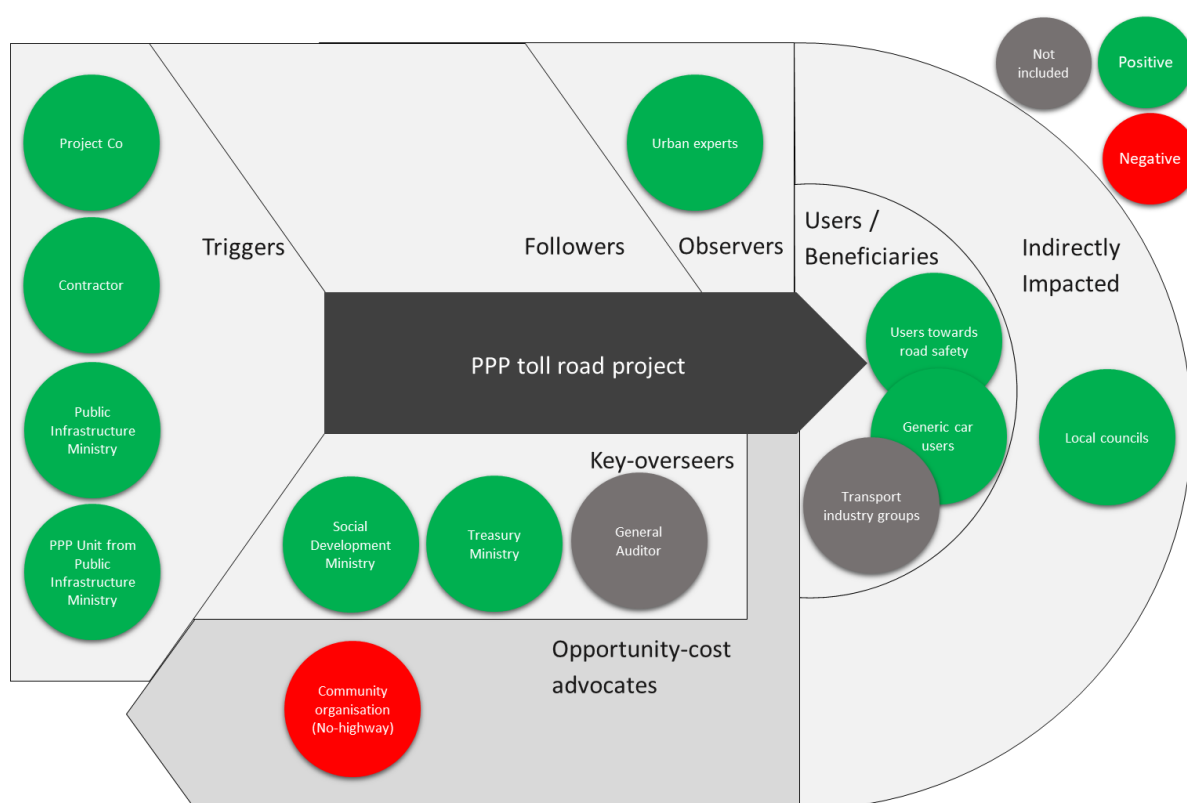


Figure 7-18: Incentive-based map with stakeholder judgement over the three judgement principles (Transport test case)

[Source: Author's]

According to most stakeholders, the urban toll road project met the three principles of successful project-systems. The researcher's analysis of the arguments and criteria that led to their judgements does not disprove the perception of the stakeholders.

7.3.4. Test case 2 evaluation conclusion

Table 7-3 summaries the evaluation data and the analyses that were conducted. This is the support for the evaluation conclusion.

Table 7-3: Transport test case evaluation summary

Project	
Urban toll road PPP project	East-west city highway in a Chilean city with a cost USD 500 million, operating since 2005.
Identified and included stakeholders	
Project Co (included)	General Auditor
Contractor (included)	Transport industry group
Public Infrastructure Ministry (included)	Generic car users (included)
PPP unit from the Public Infrastructure Ministry (included)	Users towards road safety (included)
Social Development Ministry (included)	Activist group (No-highway) (included)
Treasury Ministry (included)	Urban experts (included)
	Local councils (included)
Workshop data collection	
1) Stakeholder's definition	4) Evaluation criteria (critical and non-critical)
2) Stakeholder's aim towards the project	5) Assessment of the evaluation criteria
3) Key motivations and drivers of the stakeholder	6) Judgment of stakeholder principles

Workshop analysis	
Consistency analysis	Within the stakeholder perspectives, a search for inconsistencies between stakeholder definition, key concepts, selected criteria, and judgements was conducted. No inconsistencies were found
Critical criteria analysis	12 critical criteria were gathered, 'business model' was assessed as neutral and 'urban integration' as negative
Theme analysis	6 themes were created and analysed separately. 'Direct purpose of the project' and 'process' had an overall positive impact; 'city connection' and 'business and finance' had mostly a negative assessment
Negative and missing criteria analysis	The business model criterion is further discussed in PPP effect, connection to the city is inconclusive on who is responsible (project or public policy), and urban segregation is questioned by the researcher.
Quantitative analysis	No major conclusions could be obtained
PPP effect analysis	A discussion lead to suggest that the benefit of 'infrastructure availability' unweights the 'business model' problem
Stakeholder judgement analysis	All the reached stakeholders have a positive judgement towards the three principles, except for 'opportunity cost advocates' (No-highway group)

[Source: Author's]

The evaluation method structure (coded as HP6) has three parts: (1) about the project-system, (2) about the procurement method, and (3) about the objective of the evaluation.

The project-system 'urban toll road project' has 13 identified stakeholders and 11 of them were included in the evaluation. The project-system generally meets the three principles that inform a successful project-system. First, there is a willingness for repetition from all the identified stakeholders. Here, the No-highway group was the only stakeholder that did not want to hypothetically repeat the project; they provided legitimate support, but according to the evaluators, it is not enough to counterbalance all the other stakeholders. Second, the evaluation suggests that there is minimum fairness, which means that no stakeholder received any evident excessive negative impact from the project. Third, the no-project alternative option to the project, presented by the no-highway community group, is not considered evidently better compare to the existing project and it can be a matter of discussion. However, it can be suggested that by most of the stakeholders, the 'do-nothing option' is not a better alternative.

The project was procured through a PPP model. The identified project impacts that were considered by the researcher as influenced or 'an effect' of the PPP model included an insufficient business model as the main issue to analyse. Here, the aim is to judge if the most recognisable benefits of the PPP model outweigh its problems. The evaluation suggests that the benefits (service availability) outweigh the identified PPP problems under business model (Direct subsidy, excessive private profit, service-fee system) mainly because the infrastructure is considered 'voluntary' and 'additional' to the transport system, which could be a discussable assumption.

The objective of this evaluation is the last conclusion. The project evaluation was done in this case for academic purposes; thus, it does not follow the four general aims which are: transparency and accountability, PPP legitimisation, industry learning, and government agency learning. However, it can be stated that the project evaluation provides lessons learned and it should also inform strategic decision-making.

7.3.5. Refinement insights from test case 2

The application of the evaluation method in this test case generated refinement insights that are synthesised into 8 observations made by the researcher.

- (TP1) some stakeholders do not have representatives that can fully represent them,
- (TP2) stakeholders' representatives could have no support to answer the judgement principles,
- (TP3) consider the ex-ante evaluation to have objective data,
- (TP4) development of the taxpayers' perspective analysis,
- (TP5) development of incentive-based stakeholder map for the stakeholder identification,
- (TP6) two judgement principle must be reframed, and
- (TP7) suggesting criteria to stakeholders in the workshop.

TP1: Some stakeholders do not have representatives that can fully represent them

Some stakeholder representatives cannot provide a judgement without discussing it internally with the stakeholder. This insight is included in the final version of the evaluation method under the stakeholder identification process.

The stakeholder representation of (for example) public entities is difficult to address because it involves personal opinions that in certain occasions need to be discussed internally within the stakeholder before making any judgement. For example, the Department of Finance is not able to make a judgement without making a complete study that supports it, aligned with their internal procedures. An evaluation based on perceptive criteria which are answered through an 'expert opinion' is not considered valid for public declarations. This means that if this ex-post impact evaluation aims to solve a public request beyond the academic sphere; a deep analysis of each stakeholder must be done, creating formal reports signed by the proper authorities. This will be a way to validate the holistic approach of the ex-post evaluation.

TP2: Stakeholders' representatives could have no support to answer the judgement principles

A stakeholder representative does not necessarily have all the information for answering to the stakeholder judgement principles. This capacity should be part of the evaluation record for transparency issues. A modification is included in the final version of the evaluation method.

One stakeholder was not able to properly respond to the principle 'no excessive negative impacts on any stakeholder'. The answer was that as a public agency they were not capable of answering because they could not be sure about the status of all the other stakeholders. This issue was not considered before, and it should be included in the evaluation procedure. The facilitator should ask for the judgement principles using phrases such as 'at the best of your knowledge...' or 'if you are able to answer...' and state the real answer also in terms of the stakeholder's confidence for answering.

TP3: Consider the ex-ante evaluation to have objective data

The ex-ante evaluation for contrasting the perception that is gathered from the stakeholder workshop was mentioned by one of the stakeholder representatives as something very important. Therefore, a deeper review of the existing documentation is included in the final version of the evaluation method (workshop preparation section), to identify key information before starting the stakeholder workshops.

The link that should exist between the ex-ante and the ex-post evaluation was already discussed in Chapter 6 (refinement insights MF3). However, this link was noticed with no practical implications. Now, the test cases provide a ground for using the ex-ante evaluation (1) in the stakeholder workshop

for discussing and arguing with the representatives, and (2) for the data analysis process in which the perspective (represented with criteria) is analysed by the evaluators for reaching a final judgement.

TP4: Development of taxpayers' perspective analysis

The evaluation process systematically addresses all the stakeholder that were identified from the private sector, public sector, user/beneficiaries, and community/society. However, the taxpayer, which is the essential trigger for conducting an evaluation, is not well represented. A taxpayers' perspective analysis was developed and included in the final version of the evaluation method, together with the other analyses developed in test case 1 (SH8). This analysis includes a data collection component, so it was fully applied only in test case 3.

To address the taxpayers' perspective in the evaluation method, a new component is included: qualitative Value for Money approach. This component has the aim of assessing the Value for Money of the public investment specifically for the PPP method. The current use of the Value for Money indicator is the quantitative comparison between the best private bid and the public-sector comparator, measured in Net Present Cost (NPV). This is done ex-ante for decision-making purposes, but it is rare to find a confirmation of this initial estimation. Moreover, this estimation could not be confirmed in the future because the PSC would probably not change as the estimation is a Traditional Procurement option that never occurred.

The proposed qualitative Value for Money assessment is a systematic judgement over a comparison between the PPP benefits and PPP problems for a single project. The information presented is not quantifiable, but a judgement can be made with clear and transparent arguments. The analysis focuses on the main concern about PPPs, which is, do the PPP benefits outweigh the PPP costs and problems?

Equation 7-2: PPP benefits and PPP problems

$$PPP\ benefits > PPP\ problems?$$

[Source: Author's]

The main concern of the taxpayer is, how much resources does the PPP project incurs? According to the PSC comparison, PPPs never have extra cost. However, this way of isolating the problem is an answer to the existing questioning of PPPs and the Value for Money assessment that employs PSC.

Another issue to consider within the PPP benefits are the avoided problems from the Traditional Procurement. Then, the judgement that needs to be made is: does the extra cost of the PPP outweighs the PPP benefits (including the avoided TP problems), discounting the other PPP problems. Equation 7-3 disaggregates the PPP problems and benefits.

Equation 7-3: Disaggregation of PPP problems and benefits

$$PPP\ problems = PPP\ extra\ cost\ problem + PPP\ other\ problems$$

$$PPP\ benefits = PPP\ evident\ benefits + avoided\ TP\ problems$$

[Source: Author's]

Due to the unquantifiable nature of the benefits and problems, the concept of value is introduced.

Perceived value is a personal (or agreed in a group) valuation of anything, based on moral and cognitive principles. Value is difficult to quantify but it is possible to rank it for available options. A judgement can be made on what delivers more or less value than another thing; i.e. what is more valuable. Value is presented in this way because the main problem of the PPP decision-making is the impossibility to quantify (in reliable terms) the benefits and problems of a PPP project. This issue is

discussed by Barton, Aibinu, and Oliveros-Romero (2019). According to the authors, Value for Money needs to be separate the notion of 'value' and 'money'. The value of something (e.g. a project option) is measured by employing an ordinal scale. This means that, according to the decision-maker, one project option can provide more or less value than the other options. Then, money, which is the quantified price or cost of the options, provides the counterpart of a 'Value for Money judgement'. This framework is useful for the proposed VfM approach.

In order to make a structured logic thinking and decision-making, the judgement is presented in a mathematical form, assuming that the components cannot be quantified.

To deliver a project using a PPP model, the following inequation must stand:

Equation 7-4: Value of disaggregated PPP benefits and problems

$$v(\text{PPP extra cost}) < v(\text{PPP benefits}) - v(\text{PPP problems}) + v(\text{avoiding TP problems})$$

[Source: Author's]

The idea behind the conceptual equation is that good Value for Money is possible to achieve if for the taxpayer the extra cost of a PPP is worth its benefits (discounting its problems). It is a judgement that tries to synthesise the main concern of the theoretical taxpayer: does the presumably PPP extra cost outweighs the PPP problems/benefits? PPP benefits and avoided TP problem are presented separately because it clarifies the identification of both; however, they complement each other in most of the cases and the same attribute should not be counted more than once.

The value of the PPP extra cost can be counter intuitive as it means that 'the extra cost has a value by itself'. However, this is aligned with the employed Value for Money understanding from Barton et al. (2019), as it represents the marginal value of an option, which is part of the Value for Money analysis. For example, with two options of projects, one more expensive than the other, the question to ask is, does the extra cost of option 1 worth the extra value it provides? Here it is assumed that option 2 is providing less value but it also costing less money. For the case of the PPP analysis, the question would be 'does the extra cost of the PPP worth the extra value of using a PPP model (judged using benefits and problems)?'.

Due to the nature of each different project, the use of the PPP method can provide different benefits, problem and avoided TP problems. Moreover, depending on the project's nature, the valuation of these benefits and problems will be different in each case. Theoretically speaking, an imaginary project can have only one exceptional benefit capable of outweighing all other problems and costs. Consequently, it will be possible to make a case for having done the project as a PPP even with all the problems and including the extra cost. For example, if the value of the infrastructure availability (to have the facility operating) is larger than the value of many problems including a non-competitive tendering process, then, the PPP option is providing more Value for Money to the taxpayer.

The valuation of each benefit and problem depends of which option is being compared. According to Heald (2003), the real decision-making is not between the PPP and the Public Sector Comparator (PSC) in terms of efficiency but in terms of funding availability. If the availability of public funding allows the PSC to be built; then, that is the option to compare with the PPP. If there is no available public funding, the fall-back position needs to be considered instead of the PSC. The Fall-Back Position (FBP) is the option that should be done with the real available funding, which is probably a mitigation strategy, temporary improvement, or a time buying by outsourcing the missing service. For example, the value of the PPP benefit in term of urgency (infrastructure available earlier) will be much higher if the PSC has no funding availability, because the PSC will never occur. Then, the correct comparison is between the PPP projects and FBP that could be a temporary solution or similar. With no funding, the issues is

not an efficiency analysis between the public or private option, it is an effectivity analysis whether the project is important and urgent or not (Oliveros-Romero & Aibinu, 2019).

The improvement to the PPP ex-post impact evaluation method by using this theory is operationalised by (1) presenting a form to the stakeholders in which all the possible PPP benefits, PPP problem, and avoided TP problems are described; gathered from the existing literature. (2) The stakeholder's representative discusses and selects which of the attributes or issues happened (at best of his/her knowledge) in the project that is being assessed. (3) The data is analysed by assessing coincidences between stakeholders, the validity of the judgement (according to their perspective), and inconsistencies. The final report should include this information as clear as possible to suggest with reasonable support the existence of certain PPP benefits and problems.

The evaluation should then suggest and argue if the inequation is valid. This means to judge if the value of the PPP extra cost is less than the value of the benefits/problems that the PPP model provides. Even though this can be quantified by putting scores and weights, a driver of this proposed evaluation is to avoid quantitative decision-making and use qualitative reasoning for making a judgement.

TP5: Incentive-based stakeholder map for the stakeholder identification

To improve the stakeholder identification process, and additional tool to the existing lifecycle stakeholder map (SH7) was created. The new tool called incentive-based stakeholder map is included in the final version of the evaluation method and applied to all test cases (1,2, and 3) only for presenting the project evaluation. This tool was not applied to the stakeholder identification process of test case 3 because the stakeholders were identified before the incentive-based stakeholder map was fully developed.

A project-system was defined in Chapter 5 as *“the set of different stakeholders implicitly connected to develop a temporary endeavour, which is undertaken to create a unique product, service or result”*. Consistent with this definition, the incentive-based stakeholder map shows how the different stakeholders interact with the project and the aim of the project, classifying them according to their general incentives. Seven groups of stakeholders are present in this framework (Figure 7-19): triggers, followers, key-overseers, beneficiaries, indirectly impacted, observers, and opportunity-cost advocates.

Similar to the impact of the 5 Porter's forces (Porter, 2008) or the business canvas model (Osterwalder, 2004), the incentive-based stakeholder map should inform and support decisions by helping to clarify and understand reality. This framework complements the existing literature and improves the rigour of the stakeholder analysis; the aim is not to replace existing frameworks such as the power-interest matrix, the rainbow diagram, etc.

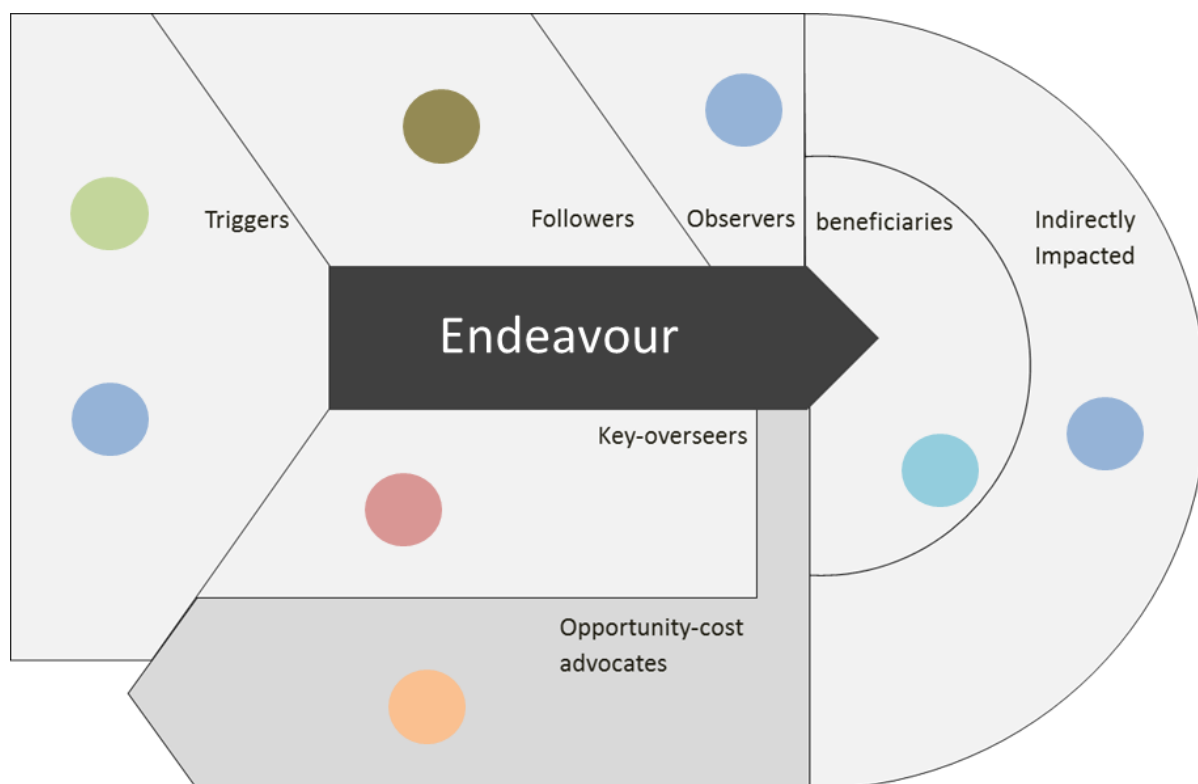


Figure 7-19: Incentive-based stakeholder map

[Source: Author's]

Triggers

Any endeavour, action, or temporary organisation for a specific purpose (general definition of project) will, undeniably, have stakeholders that will start this process; they will have the first drafted idea, the first approval, the first motivation to generate an endeavour, either to solve a problem, or to take an opportunity. These stakeholders are called 'triggers', and their power should be enough to think that they can (1) deliver the project, (2) support the project until the end, or (3) be capable of enforcing someone else to deliver and/or accomplish the project's purpose.

Important attributes to consider when analysing a project as a system is the recognition of triggers' efforts for the accomplishment of the project. Their power and interest should generally prevail when contrasted with other stakeholders unless a strong argument can support others' view over the project.

An example of a project trigger would be the innovation department in a business organisation, a think-tank in a public policy development project, the public authority in a public project, a community organisation in a social project, or the entrepreneur in a new business.

Key-overseers / key-supporters

After the project has started, there are certain stakeholders which need to be 'in favour' or need to approve the project explicitly or implicitly; at the start and during the project's lifecycle. In general terms, key-overseers will always have enough power to cancel, interrupt, modify, or change the course of the project, and that is why they are 'key'. They do not necessarily need to be interested in the project, but they need to have enough power to be considered in this category.

In a business, the key overseers will be, for example, the justice system, the tax agency, the local council, or even key suppliers. In a public project, the Treasury Ministry will always be a key-overseer.

Followers

This stakeholder category is part of the project-system because of their indirect/independent relationship with the endeavour. These stakeholders need to be considered because they are either affected by the process of the project (construction, implementations, temporal deliverables), or they provide or affect the project without real interest. Suppliers and maybe workers can be in this category; a competitor to the purpose of the project (e.g. trying to deliver the same outcome), or a provider of an alternative solution.

Users / Beneficiaries

The beneficiaries are the stakeholders that should receive the positive impact that is explicit in the project's purpose. For example, in a business they should be the client; in a public policy, the subgroup of citizens; or in a public project, a specific community, among others.

Indirectly impacted

Mainly because of the purpose of the project, there are stakeholders that receive or are affected positively or negatively by the project's outcome. In a public infrastructure project, the surrounding community could be in this category, and also the extended society as legitimate owners of the public resources (taxpayers' money).

Observers

Stakeholders that have an overarching view of the project could be considered because (1) they provide the closest view to an impartial analysis, (2) they understand the project with a wider (maybe worldwide) view, and (3) they have might forecast the long-term implications of the project. Even though their influence and relationship to the project is not as close as other stakeholders (such as the triggers, followers or beneficiaries), they can influence the agenda in which the project is immersed, affecting future developments that will be aligned with the project.

International organisations, national media, national authorities in an independent local project, professional organisations, universities, academia, etc. They could all be observers and part of the project-system. They usually do not have power in the short term, but they might in the long term.

Opportunity-cost advocates

A project is assumed to be part of a larger structure in which many projects, actions and options can be delivered within the same closed environment. A business organisation has many business opportunities, but they have limited resources. Similarly, a public agency has many options for solving an existing social problem and many social problems that need to be solved.

Opportunity-cost advocates are the organised groups that can present a real alternative to the project that is being developed (based on rationality and study). They can oppose to the project (because they argue that the 'do nothing' option is better), or they can claim for a big modification in the project to increase its value. Also, they can advocate for a different project that (1) can solve the problem in a better way, (2) solves a more important problem, or (3) mitigates the negative externalities of the developed project. Opposition or dissatisfaction with the project can be in any stakeholder. However, the opportunity-cost advocates are the ones that present a feasible and clear alternative to the project.

The incentive-based stakeholder map is considered a main development within the testing process and also applicable to other projects beyond this research.

TP6: Two judgement principle needed to be reframed

The stakeholder judgement principles proposed in the conceptual framework Project Success Evaluation Pyramid Model were reframed for clarity purposes to the actual willingness for repetition, no excessive negative outcome, and no evident better alternative. These modifications were updated throughout the thesis from the conceptual framework in Chapter 5 to the final version of the evaluation method in Chapter 8.

The reframing of the second and third judgement principle is the only modification (refinement insight) that affected the conceptual framework 'Project Success Evaluation Pyramid Model'. Initially, the principles were expressed as questions. The second principle was, 'are the stakeholders willing to exchange roles with any other stakeholder assuming they will receive their same impacts? This question was confusing and too embodied in the theory that supports the principles. During the test cases, and for clarity purposes, this principle was changed to 'no excessive negative outcome to any stakeholder', which, in fact, expresses even better the purpose of the principle. The willingness to exchange roles is just a theoretical exercise to prove that if everyone is willing to accept being any other stakeholder, then it means that presumably there are no excessive negative outcomes. The third judgement principle was written as, 'is the project the best possible option within the existing original project alternatives? it was reframed to 'no evident better alternative'.

TP7: Suggesting criteria to stakeholders in the workshop can improve criteria identification

From test case 1, three refinement insights addressed the way to conduct stakeholder workshops: (SH3) about the skills of the facilitator, (SH4) about the of actual value, and (SH5) about the workshop structure and timing. To further improve the stakeholder workshop structure two strategies were tested: (1) to suggest criteria to the stakeholder's representative, and (2) to expect an independent identification of criteria from the stakeholder's representative. After testing both, the solution adopted was to provide an independent list of concepts that are mostly generic and can fit any stakeholder's purpose. This improvement was included in the last version of the evaluation method and applied in test case 3.

The first option which was a fix list of success criteria for the stakeholder to approve or rank was considered inaccurate because it contains a flaw; it drives the stakeholders to the common criteria (time, cost, purpose), possibly hiding the real deal-breakers. Conversely, the second strategy was suggesting the criteria, which can trigger a facilitator bias that can modify the stakeholder perspective. Asking the stakeholder to independently think and identify the criteria can generate the risk of missing important criteria, exacerbated if the stakeholder was not prepared for the workshop. The researcher observations suggest that the risk of missing criteria can be better addressed in the subsequent evaluator analysis (missing criteria analysis); compared to the risk of having a facilitator bias. The reference list of concepts was seen as a middle ground option the stakeholder can use to help the identification process, reducing the influence over their decision.

Summary of test case 2 contribution to evaluation method development

Test case 2 applied the evaluation method including some improvement from the previous test case for the data collection process: a better stakeholder workshop structure (coded as SH3 and SH5), and the use of 'actual value' instead of 'perceived value' (coded as SH4). Test case 2 contributed to the improvement of the evaluation method with two analytical tools: the incentive-based map and taxpayers' perspective analysis; both major developments. Additionally, it contributed with the

reframing of two judgement principles and with a further improvement of the stakeholder workshop structure. Test case 3 will fully apply the taxpayers' perspective analysis (TP4).

7.4. Test case 3 – Health PPP project

This section presents the results of applying the evaluation method to a PPP public speciality hospital in Australia. Speciality means that it has a specific target within the population (e.g. Children, cancer, mental health, pregnancy, etc.). Six refinements insights were synthesised from this test case.

The structure of the test case 3 project evaluation is as follow. First, the context of the project is summarised (overview). Second, the stakeholder identification process is described. Third, the analysis of the stakeholder workshops is presented using five types of analyses, namely, (1) critical criteria analysis, (2) negative impact analysis, (3) PPP effect analysis, (4) taxpayers' perspective (Value for Money), and (5) stakeholder judgment analysis. Fourth, the evaluation conclusion. Finally, the refinements insights are described.

7.4.1. Test case 3 overview

An Australian public speciality hospital announced in 2003 a redevelopment of its building located in an inner-city suburb. The new redevelopment employed a Public-Private Partnership model of procurement. The selected new site of construction was another inner-city area, next to the general public hospital. The business case was submitted under the guidelines of 'Partnerships Victoria', the PPP unit of the state government of Victoria, which is part of the Victorian Department of Treasury and Finance.

The objectives of the project are [extracted from Frost (2008)]:

- Provide a modern facility that supports the delivery of accessible, cost-effective and high-quality patient services to Victorian women and their babies
- Provide a new facility that is operationally efficient, capable of achieving service plan targets and sustaining service levels into the future
- Provide an environment that enhances patient safety and improves clinical outcomes through the provision of comprehensive tertiary care for women, improved access to emergency services, specialist inpatient services, advanced diagnostics, intensive care and specialist clinical services
- Provide a facility with flexible infrastructure capable of adapting to new technologies, clinical practice changes, changes in government policy and funding arrangements
- Provide appropriate facilities for (and the integration of) teaching, training and research within clinical areas and between the speciality hospital and general hospital
- Achieve a successful relocation with no interruption of the ongoing delivery of services
- Procure the new hospital in a manner which delivers Value for Money to the State
- Complete the project within the budget and other parameters agreed by the government
- Achieve efficiencies in capital and recurrent costs for both the speciality hospital and the general hospital through the best use of jointly available resources

The original investment was AUD 250 million in capital funds from the government for the redevelopment of the hospital. It is a 25-years contract with a design, build, maintain, and transfer (DBMT) PPP project. The direct transfer payments from the government to the private entity supposedly started after the hospital was commissioned and estimated to be in total AUD 1073 million in nominal terms, equivalent to AUD 421.5 million in net present value terms at 30 June 2007.

A project agreement between the government and the private entity stated a construction stage of approximately 3 years, with a completion date June 2008, and an ongoing delivery of cleaning, car

parking, help desk, security and building maintenance services (the services), excluding clinical functions—for a period of 25 years, from completion of the facility to June 2033.

The Victorian General Auditor assessed this project when it was at its construction stage. The focus of the report was on the correct compliance of the processes, which include mostly the PPP unit guidelines and the public investment procedures. So far, there are three hospitals that have been procured under the PPP unit guidelines, two speciality hospitals and one general hospital. The three hospitals have created a learning curve for the public sector.

The context of testing the evaluation method

The hospital was delivered in 2008 and therefore, it has 10 years of operation. The prototype evaluation method (with some refinements) was applied to the PPP project. Some evaluation formalities were not included in this test case because the objective at this stage is to test and design the main components of the evaluation. Thus, steps such as the declaration of the evaluation requester, or the analysis of the best time-period for evaluation were not applied.

7.4.2. Test case 3 stakeholder identification

This process is presented using the tools that were developed for this purpose: the lifecycle stakeholder diagram (coded as SH7), and the incentive-based stakeholder map (coded as TP5).

Figure 7-20 shows the stakeholders that were part of the evaluation. Some stakeholders were not possible to reach, and they are categorised as ‘not included’. Similar to the two previous test cases, this is not considered a problem for the design process because the aim is to develop the evaluation method. However, any extrapolation of the project’s contextual results should be aware of this data limitation.

Stakeholder group	Conceptualisation	Design	Construction	Operations	Disposal	Stakeholder representative
Private sector	Project Co					Project Co head director
	Builder					Not included
	Operation and management company					Not included
Public sector	Hospital Operator (public)					Hospital executive director
	Department of Health & Human Services					DHHS PPP contract administrator
	Hospital PPP contract management					Hospital PPP contract administrator
	Department of Treasury and Finance					DTF executive
	Council					Not included
	General Auditor					Not included
User/beneficiary	Medical staff					Former hospital junior staff
	Patients					Former hospital patient
	Private hospital					Private hospital director
Community	Surrounding neighbours					Not included
	Health specialists and consultants					Not included

Figure 7-20: Project lifecycle stakeholder diagram (Health test case)

[Source: Author's]

Additionally, the incentive-based stakeholder map (Figure 7-21) shows the stakeholders in the project-system, categorised under their main motives and drivers.

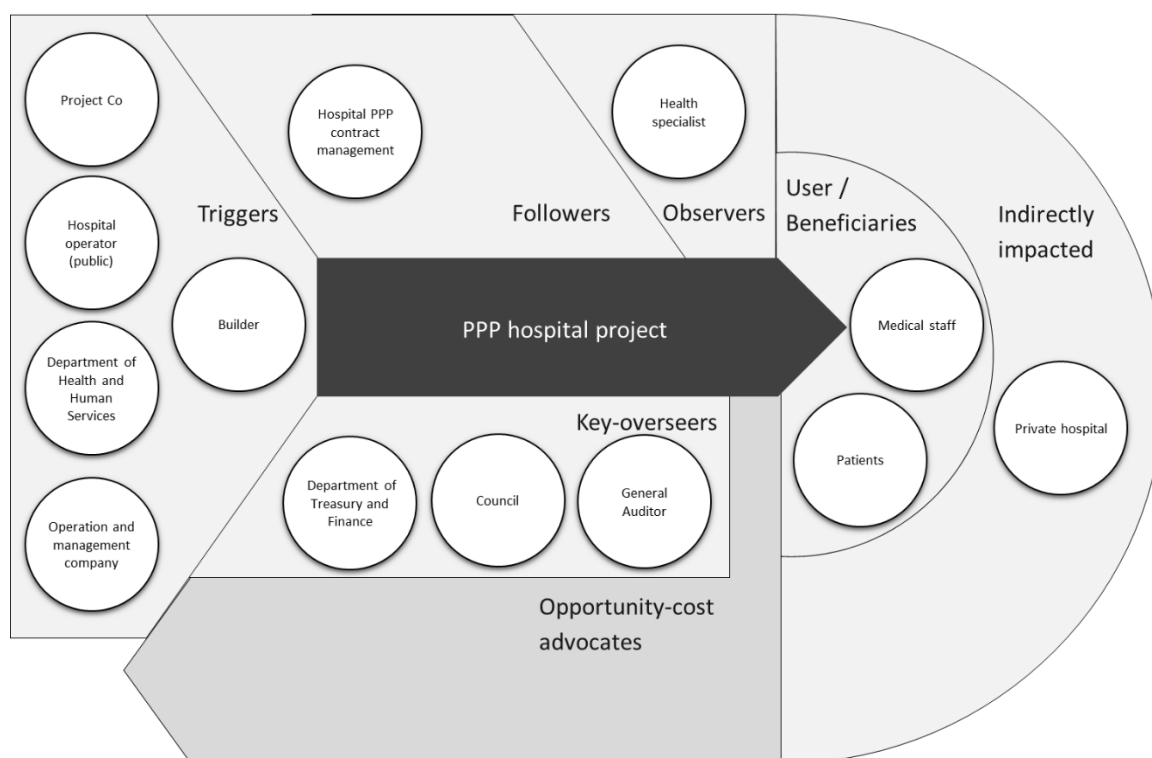


Figure 7-21: Incentive-based stakeholder map (Health test case)

[Source: Author's]

The next step in the evaluation method was to conduct workshops with each of the stakeholder representatives mentioned in Figure 7-20. Their stakeholder perspectives were gathered using a stakeholder workshop form (PE8 in Chapter 6) which was modified with the refinement insights TP4 taxpayers' perspective and with TP6 that reframes two judgement principles. The description and analysis of this data are presented in the next section.

7.4.3. Test case 3 stakeholder workshop analysis

For this test case, seven components were gathered per stakeholder workshop (test cases 1 and 2 had only 6 components).

1. Stakeholder's definition (who is the stakeholder?),
2. Stakeholder's aim towards the project (why is it relevant to the project?)
3. Key motivations and drivers of the stakeholder (what drives stakeholder's actions?)
4. Evaluation criteria (which criteria the stakeholder uses to assess the project?)
5. Assessment of the evaluation criteria (assess/valuate the selected criteria)
6. Identification of PPP benefits, PPP problems, and avoided Traditional Procurement problems that are present in the evaluated project.
7. The judgment of stakeholder principles (principles defined in section 5.5: willingness for repetition, no excessive negative outcome, no evident better alternative)

The second last component regarding the identification of PPP benefits and problem is incorporated into this test case as a result of the refinement TP4 developed after test case 2. Test case 3 is the only test case in which the taxpayers' perspective (TP4) insight was fully applied in a data collection and data analysis process.

The seven workshop components were analysed employing five types of analysis: (1) critical criteria analysis, (2) negative impact analysis, (3) PPP effect analysis, (4) taxpayers' perspective (Value for Money), and (5) stakeholder judgment analysis.

Critical criteria analysis

Figure 7-22 shows the critical and non-critical criteria from all the stakeholders. For this project, all the critical criteria are valued positive impacts for their stakeholders. This analysis suggests that the core expectations of the stakeholders are met. This result is a strong argument in favour of a positive final judgement about the question, in this project-system successfully working? In practice, this is a holistic view of the project that shows the core impacts of every stakeholder as positive.

	Critical criteria		Non-critical criteria					
Project Co	Net present value		Facility performance	Satisfaction of State of Victoria	Prestige	Stakeholder relationship	Risk assessment	
Hospital operator (public)	Service quality and safety	Patient experience	Project's purpose achievement	On budget, on time	Regulation compliance	PPP governance and relationship	Meet functional requirements	Whole-of-life cost lifecycle
Department of Health and Human Services	Private entity's proactivity		Value for Money	Maintenance	Contract compliance	Hospital operator understanding of PPP	Relationship DHHS- FM - HO	
Hospital PPP contract management	Relationship public-private		Agreed performance indicators	User/customer experience	Accreditation	Narrower focus of work (outsourcing)	Future proofing	Contract design
Department of Treasury and Finance	Value for Money	On budget	On time	Robustness of project plan (min. changes)	Service availability	Economic impact at city level	Risk management and mitigation	
Medical staff	Quality of medical senior staff	Training quality	User/friendly infrastructure	Accessible infrastructure	Quality of equipment	Building appropriate to purpose	Medical logistic	
Patients	Medical quality	hospital emergency response	Access to medical technology	quality of emergency service	Access to specialists			
Private hospital	Business strategy		Financial indicator	Maintenance of PPP (infrastructure)	Relationship with RWH	Clinical synergy	Infrastructure lease	Commercial viability

Figure 7-22: Evaluation criteria for project stakeholders separated by critical and non-critical (Health test case)

[Source: Author's]

Negative criteria analysis

Similar to test cases 1 and 2, this analysis focusses only on the negatively valued impacts. The reason is that this evaluation method presumes that the project-system is successfully working until the opposite is suggested. Therefore, negative impacts are analysed individually.

- **Risk assessment:** According to the Project Co, the risk allocation process had a big pitfall, as the technological risk of updating medical equipment was assigned to them. This risk is part of the Project Co main responsibility, which is to provide the maintenance of the hospital, excluding the core medical service. However, according to them, this risk is difficult to assess and is strongly linked to external variables such the health policy or the existing technology improvement. This issue has not been discussed with the hospital operators and it was not mentioned by another stakeholder. Nevertheless, it is a reasonable claim that is not considered critical to this PPP project, but it is a relevant lesson learned.
- **Future-proofing:** Stated originally by a Victorian General-Auditor report and recognised by the hospital PPP contract manager, one issue in the project was the non-construction of the extra floor that was already part of the design for future expansions. According to the General-

Auditor report, to build the extra floor together with the rest of the hospital would have been a better option than building it in the future while the hospital is operating. The latter option is more expensive to build, and it creates a bigger intervention to the hospital. This is considered a lesson learned and it has already been applied to subsequent hospitals that were built with all the conditions to make future expansion easier. For example, installing the lifts with the extra floors included.

- Contract design: According to the hospital PPP contract manager, the contract design is considered insufficient and has led to many negotiations between the public operator and the Project Co. According to them, there are many issues that the contract does not address, leading to interpretations and to act in 'good faith'. However, it is stated that there is a good relationship between the private and public sector, recognising the proactivity of the private company.

Neutral impacts are not analysed in detail, as in this case they are not considered a strong argument for changing the judgement of the overall project.

Missing criteria analysis

The purpose of this analysis is to identify any stakeholder or criterion that was not included by the stakeholders that are considered part of the project-system. This identification is done by the evaluators (researcher).

- Contractor's perspective: The contractor was not possible to reach as a stakeholder and it is not included in this evaluation. No information about profit or benefits was found from the contractor's perspective, as it merged into a bigger company. However, a media search did not arise any problem regarding the construction of the hospital, and Lend-lease (the new company) has been entering PPP project with the Victorian government after the end of this project. This information, at least, suggests that an overall impact of the project on the builder is unlikely to be negative. Nevertheless, the specific criteria in which there could be negative impacts are unknown.
- Council: The urban planning office of the Council was also not possible to contact, but similar to the contractor view, nothing suggests that the perspective of the council could be negative. The hospital is part of the health district plan in which the city is involved, and the hospital is an infrastructure that usually local government are likely to pursue and appreciate. Maybe the relocation from one neighbourhood to another could be an issue, but, even so, it is also unlikely that this relocation (if negative) could be considered critical by the council.
- Surrounding community: The surrounding community of the new hospital was not identified for this test case. Here it is possible to, at least, suggest three positive aspects. Any health infrastructure, in theory, can provide value to the community as it can be used in case of emergency. Second, public participation regulation in Australia offers stages for the community participation, which in many countries (for example Chile) is not a formal process and not enforced by law. Third, a media search did not show any major problem with the hospital's construction or operation, meaning that a negative impact on the surrounding community is unlikely. However, similar to the previous missing perspectives, the specific criteria in which the community could negatively assess the project are not known.

PPP effect analysis

This analysis was applied to test cases 1 and 2. It aims to separate the impacts of the project from the impacts of the procurement method. The criteria gathered from the stakeholders are categorised (Figure 7-23) under (1) high effect of PPP, (2) medium effect of PPP, and (3) no effect of PPP. This classification was done by the researcher and is not intended to be a deterministic categorisation, the boundaries are fussy, and the main idea is to have a comparative diagram to understand the influence of the procurement model on the different perspectives.

	high effect of PPP		medium/low effect of PPP				no effect of PPP		
Project Co	Net present value	Satisfaction of the State	Facility performance	Prestige	Stakeholder relationship	Risk assessment			
Hospital operator (public)	Project's purpose achievement	PPP governance and relationship	On budget, on time	Whole-of-life cost lifecycle	Meet functional requirements	Patient experience	Service quality and safety	Regulation compliance	
Department of Health and Human Service	Private entity's proactivity	Contract compliance	Hospital operator understanding of PPP	Maintenance	Value for Money	Relationship of public and private partners			
Hospital PPP contract management	Relationship public-private	Narrower focus of work (outsourcing)	Agreed performance indicators	Accreditation	User/customer experience	Future proofing			
		Contract design							
Department of Treasury and Finance	Value for Money		On budget	On time	Robustness of project plan (min. changes)		Risk management and mitigation	Economic impact at city level	Service availability
Medical Staff			User/friendly infrastructure	Accessible infrastructure	Quality of equipment	Building appropriate to purpose	Quality of medical senior staff	Training quality	Medical logistic
Private hospital		Maintenance of PPP (infrastructure)	Financial indicator		Infrastructure lease	Clinical synergy	Business strategy	Commercial VIABILITY	Relationship with RWH
Patient representative							Medical quality	Hospital emergency response	Access to specialists
							Access to medical technology	Quality of emergency service	

Figure 7-23: Stakeholder criteria categorised by PPP effect (Health test case)

[Source: Author's]

The project has 48 impacts from which only 11 are not considered an effect of the PPP model. This means that the procurement method has a relatively big influence on the impacts of the project.

Contract design was already analysed as a negative impact. Now, this criterion is considered an effect of the PPP model. However, considering that (1) the criterion is not critical, and (2) negotiations and good faith mitigated its severity; it is not considered sufficient to counterbalance the other positive PPP impacts. Future-proofing and risk assessment criteria were also analysed before, and they are not considered decisive for concluding an overall negative impact of the PPP model.

None of the stakeholders mentioned that an alternative option to the project was a Traditional Procurement. However, this analysis must be included in the PPP effect analysis. The Victorian Auditor-General Office (VAGO) reported that the project followed correctly the stages of procurement method selection, in which the Traditional Procurement option and the PPP option were compared. At this point, the ex-ante Value for Money analysis (reported with the project business case) concluded that the PPP option should deliver more Value for Money than the Public Sector Comparator. This analysis was conducted before the project was built, for selecting the procurement method.

Taxpayers' perspective (Value for Money approach)

This type of analysis was developed as a refinement insight of test case 2 (TP4), partially applied on test case 2, and fully applied on test case 3. The 'taxpayers' perspective' is essentially a Value for Money analysis assuming the main interest of the taxpayer is to ensure the selected option (financially) is the one that provides more Value for Money.

The analysis consists of identifying the PPP benefits and PPP problem that are present in the project, and then judge if the problems outweigh the benefits. A list of PPP problems and benefits mentioned in the existing PPP literature was built for using it in the stakeholder workshops. Every stakeholder with a reasonable understanding of the PPP model (excluding, for example, junior medical staff) was asked to recognise the benefits and problems that were present (real) in the evaluated project. Table 7-4, 7-5 and 7-6 shows the recognised problems and benefits from the stakeholders. For this identification, the benefits of the PPP model are complemented with the category 'avoided problems of Traditional Procurement'. The idea of this additional component is to state what problems were theoretically avoided if the project was to be traditionally procured.

The benefits of employing a PPP model are listed in Table 7-4. Each benefit is under a category, and they are identified by the stakeholders that have knowledge about the PPP model of the project.

Table 7-4: Identified benefits of the PPP model (Health test case)

Category	Attribute	PPP unit (treasury)	Project Co	Hospital PPP operator	Health Ministry	Private hospital	Total
Finance	Better risk allocation	x	x	x	x	x	5
Quality	Incentive for high-quality maintenance	x	x	x	x	x	5
Quality	Clear performance standards	x	x	x	x	x	5
Quality	Better residual infrastructure after transfer to public	x	x	x	x	x	5
Finance	Funding availability		x	x	x	x	4
Efficiency	Operational efficiency due to bundling		x	x		x	3
Efficiency	On-time	x	x	x			3
Efficiency	On budget	x	x	x			3
Efficiency	Integration of design, construction, financing, operational, maintenance, and refurbishment responsibilities	x	x			x	3
Efficiency	Lower estimated lifecycle cost compared to PSC	x	x			x	3
Finance	Payment after commissioning		x		x	x	3
Quality	Incentive for availability	x	x	x			3
Efficiency	Improvement of construction quality for long-term use.	x	x				2
Finance	Design and fit-for-purpose risk on private sector	x	x				2
Finance	The potential for lower cost of assets				x	x	2
Finance	funding certainty for maintenance	x			x		2
Quality	High level of innovation	x	x				2

Quality	Transparency	x	x	2
Finance	Off-sheet balance treatment		x	1
Finance	Economically feasible projects		x	1
Finance	Impact on the economy (due to the PPP method)			0
Finance	Availability of funding for using the PSC budget on a different project			0
Finance	Urgency (due to funding availability)			0

[Source: Author's]

The recognised problems of delivering the hospital with a PPP model are shown in Table 7-5.

Table 7-5: Identified problems of the PPP model (Health test case)

Category	Attribute	PPP unit (treasury)	Project Co	Hospital PPP operator	Health Ministry	Private hospital	Total
Problem	Need for additional skills in negotiation, tendering and management	x		x		x	3
Problem	Excessive renegotiations			x			1
Problem	Loss of public budget future flexibility (due to future service fee transfers)	x					1
Problem	Rigidity in service provision			x			1
Problem	Higher tendering costs	x					1
Problem	Higher return rates over capital investment	x					1
Problem	High (out of the norm) private profits						0
Problem	Risk of non-well defined functional specifications						0
Problem	Stakeholder rejections due to private involvement (ideological reasons).						0

[Source: Author's]

Finally, the avoided problems from doing the project through a Traditional Procurement in the Australian context are identified in Table 7-6.

Table 7-6: Identified 'avoided Traditional Procurement problems' (Health test case)

Category	Attribute	PPP unit (treasury)	Project Co	Hospital PPP operator	Health Ministry	Private hospital	Total
Avoided TP problem	Low quality maintenance	x	x	x			3
Avoided TP problem	The incentive for developing new infrastructure to the detriment of investing in infrastructure maintenance	x	x	x			3
Avoided TP problem	Limited risk appetite due to the political impact			x		x	2

Avoided problem	TP	Deficient project selection			0
Avoided problem	TP	Inefficient adjustment	service-fee	price	0
Avoided problem	TP	Regulation corruption	problems	and	0

[Source: Author's]

This identification process provides an overview of the real effects the PPP model has over a specific project. The confirmation of the benefits and problems that are recognised by the stakeholders is not part of the scope of this study, but it could certainly contribute.

The judgement about the impact of the PPP from the perspective of the taxpayer is presented as a Value for Money dilemma. For a taxpayer, a PPP should be a better Value for Money option if the following inequation stands.

Equation 7-5: Value of disaggregated PPP benefits and problems

$$v(\text{PPP extra cost}) < v(\text{PPP benefits}) - v(\text{PPP problems}) + v(\text{avoiding TP problems})$$

[Source: Author's]

Where 'v' represents the 'value of', which means how valuable is a component of the equation for the evaluator. The idea behind the conceptual equation is that good Value for Money is possible to achieve if the taxpayer values less the PPP extra cost (if there is any) than the PPP benefits minus the problems. It is a judgement that tries to synthesise the main concern of the theoretical taxpayer: Do the PPP benefits outweighing the PPP problems?

From Table 7-4, Table 7-5, and Table 7-6, the main problems and benefits were gathered to present a 'case for judgement'. Not all the issues were selected, as they need to be (1) generally agreed between the stakeholder, (2) in the area of expertise of the stakeholders, and/or (3) with no evident conflict of interest. Table 7-7 shows the selected issues from the stakeholder identification tables.

Table 7-7: Selected PPP benefits and PPP problems (Health test case)

PPP benefits	PPP problems
Better risk allocation	Additional negotiation skills
High-quality maintenance compared to TP	Loss of public budget flexibility
Certainty on funding for maintenance	The rigidity of service provision
More efficient maintenance	Higher tendering costs
Clear performance standards	Higher capital return rates
Better residual infrastructure	Insufficient contract design (from PPP analysis)
Lower estimate lifecycle cost (compared to PSC)	

[Source: Author's]

For analysing the Value for Money dilemma, the main benefits and problems are selected from table 7-7. The idea is to reduce the comparison (problem/benefits) to the PPP effects that could be considered deal-breakers. The main PPP benefits are (1) maintenance of the infrastructure, (2) clear performance standards, and (3) a lower lifecycle cost estimation. Infrastructure maintenance is a recognisable PPP benefit in term of quality and efficiency. Moreover, this benefit is even more valuable when contrasted with perceived poor maintenance of traditionally procured comparable projects in the Australian context.

The main recognised problems of this PPP are (1) the long-term contract, and (2) the extra costs of negotiation, tendering; and (3) a higher capital interest rate. The problem of a long-term contract is recognised in the literature as the risk of having no future flexibility in the service provision. In this case, this risk has not happened (yet) and according to the stakeholders, there is no evidence of this

problem for the hospital. Thus, it is possible to state that this problem is not considered determinant for judging this case. The second problem is the higher tendering and negotiation costs, which are usually marginal costs. Then, the main problem to analyse is the third one: the higher interest rate for capital investment. One argument for reducing the severity of this problem is the fact that the lifecycle cost estimation was audited by the Auditor's Office, stating that the evaluation was done according to the guidelines. However, Value for Money analyses have been criticised in the past, and for this PPP project, the estimated savings were less than 1% (for the lifecycle). Nevertheless, the worst-case scenario in which the higher return are real, they should be compared to the value of having a highly maintained infrastructure (main recognised benefit), which influences the quality of the service provision and the quality of the residual infrastructure.

After analysing the main PPP problems and benefits, one benefit and one problem are selected as deal-breakers for making a final judgement. The main benefit of employing a PPP model in the hospital was the higher infrastructure maintenance. The main problem of the PPP model is the high capital return rate. The high maintenance is considered unquantifiable, but almost certain considering the expertise of the stakeholders. The problem of the high capital return rate is considered uncertain and with the counterargument that if the PSC estimation is correct, the capital return rate should be within reasonable expectations. As a final judgement, this evaluation suggests that the benefit of high maintenance should outweigh the problems of high capital return rates.

Consistent with the original question, the value of the PPP extra cost here is suggested to be zero. Then, as the benefits are judged to outweigh the problems, this PPP hospital delivered a better Value for Money for the taxpayer than if traditionally procured.

Despite the subjectivity of the analysis, based mainly of the evaluator's logic, morality, and project knowledge; the value of the analysis is on its transparent process. The analysis starts by acknowledging that all competent stakeholders provided their judgement regarding the existing PPP problems and benefits. Then, the evaluator explained how the identified problems and benefits could be evaluated in terms of severity and relevance. Finally, for this case, the analysis resulted in a dilemma between the main benefit and the main problem. This transparency is what gives legitimacy to the analysis, as the evaluator is forced to show how he/she reasoned and judged. It differs from a quantitative pairwise comparison in which the weights of each criterion could be easily manipulated in order to get a final positive score. Even though a problems/benefits table with scores for each effect could help to show the bigger picture; a qualitative approach like the one proposed here is considered preferable.

Stakeholder judgement principles

As it has been explained in the two previous test cases, the stakeholder judgement principles are (1) willingness for repetition, (2) no excessive negative outcome, and (3) no evident better alternative. Each stakeholder was asked about the three judgement principles at the end of the stakeholder workshop. These judgements are used for concluding the evaluation and assess the consistency of their judgements.

The judgement for the three principles for successful project-systems is as follows. All the stakeholders are willing to repeat the project, they think no stakeholder was excessively negatively impacted, and an evident better alternative option to the project was not possible to state. Figure 7-24 shows the incentive-based stakeholder map for the three principles.

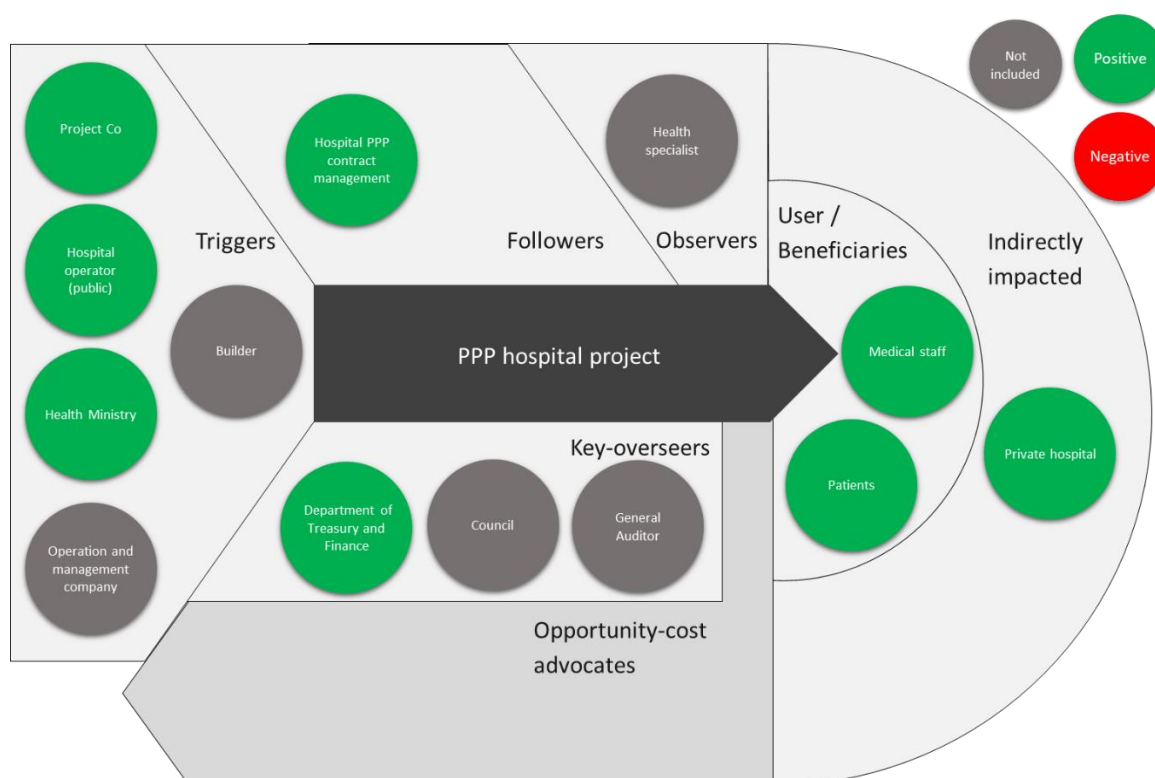


Figure 7-24: Incentive-based map with stakeholder judgement over the three judgement principles (Health test case)

[Source: Author's]

Many stakeholders have no judgement in this analysis because stakeholder workshops were not conducted with them. However, some relevant perspectives were also discussed under the 'missing criteria analysis' and nothing suggests a possible negative impact over them.

Five analyses were conducted to the data gathered from the stakeholder workshops: (1) critical criteria analysis, (2) negative impact analysis, (3) PPP effect analysis, (4) taxpayers' perspective (Value for Money), and (5) stakeholder judgment analysis. The next section provides an overall conclusion of the project evaluation.

7.4.4. Test case 3 evaluation conclusion

Table 7-8 summaries the evaluation data and the analyses that were conducted. This is the support for the evaluation conclusion.

Table 7-8: Health test case evaluation summary

Project	
Health infrastructure PPP project	Main Australian hospital relocation completed in 2011 with a budget of AUD 421.5 million under a Design-Build-Maintain approach
Identified and included stakeholders	
Project Co (included)	Council
Builder	General Auditor
Operation and Management company	Medical Staff (included)
Hospital Operator (public) (included)	Patients (included)
Department of Health and Human Services (included)	Private Hospital (included)
Hospital PPP contract management (included)	Surrounding neighbours
Department of Treasury and Finance (included)	Health specialists and consultant

Workshop data collection	
1) Stakeholder's definition	5) Assessment of the evaluation criteria
2) Stakeholder's aim towards the project	6) Taxpayers' perspective (Value for Money analysis)
3) Key motivations and drivers of the stakeholder	7) The judgment of stakeholder principles
4) Evaluation criteria (critical and non-critical)	
Workshop analysis	
Critical criteria analysis	10 critical criteria were gathered, all assessed positively by their stakeholders
Negative criteria analysis	Risk allocation, contract design, and future-proofing are negative impacts. They are considered valid, but none of them is critical
Missing criteria analysis	The contractor, the council, and the surrounding community perspectives were not reached. However, no support for evident negative impacts can be suggested
PPP effect analysis	The negative impacts that are an effect of the PPP model are not considered decisive to conclude an overall negative impact of the PPP model
Taxpayers' perspective (Value for Money)	Gathering benefits and problem, the judgement can be 'overly' simplified by stating that an (almost certain) better maintenance benefit outweighs an (uncertain) negative higher capital return rates
Stakeholder judgement analysis	All the reached stakeholders have a positive judgement towards the three principles. This is a strong argument to support a successfully working project-system

[Source: Author's]

Following the evaluation method and its procedure, the evaluation should have three main conclusions: (1) about the project-system, (2) about the procurement method, and (3) about the objective of the evaluation.

The project-system 'public hospital' has 14 identified stakeholders and 8 of them were included in the evaluation. The project-system generally meets the three principles that inform a successful project. First, there is a willingness for repetition from all the identified stakeholders. At this point, it must be noticed that no 'opportunity cost advocates were found'. Second, the evaluation suggests that there is minimum fairness, which means that no stakeholder received any evident excessive negative impact from the project when analysing all the relevant criteria gathered from the stakeholder workshops. Third, it was not possible to identify and evident alternative option to the project, which also supports its success.

The project was procured through a PPP model. The identified impacts that were considered by the researcher as influenced or 'an effect' of the PPP model included no critical negative criteria, and the negative impacts were not considered enough to make a general disapproving PPP model. Additionally, a taxpayers' perspective analysis was conducted to determine if the identified PPP benefits outweighed the identified PPP problems. The analysis suggests that the main benefit (infrastructure maintenance) do outweigh the main problem (risk of higher capital return rate). It must be stated that the evaluation could continue to confirm the reliability of these two effects, making an even more accurate judgment.

The objective of this evaluation is the last conclusion. Similar to the previous two test cases, the project evaluation was done in this case for academic purposes; thus, it does not follow the four general aims which are: transparency and accountability, PPP legitimisation, industry learning, and government agency learning. However, it can be stated that the project evaluation provides lessons learned and it also should inform strategic decision-making.

7.4.5. Refinement insights from test case 3

The application of the evaluation method in this test case generated six insight for refining the evaluation method coded as hospital project (HP) from 1 to 6.

- (HP1) snowball stakeholder identification can be employed,
- (HP2) stakeholders that change their representation through time,
- (HP3) government disclosure regulations can restrict the development of the stakeholder perspective,
- (HP4) use key motivations and drivers of the stakeholder (workshop component) to help the criteria identification process,
- (HP5) stakeholders can lack sufficient knowledge to judge a selected criterion, and
- (HP6) development of the evaluation conclusion structure.

HP1: Snowball stakeholder identification can be employed

The evaluation method originally proposed a systematic stakeholder identification process with a broad identification, exclusion and inclusion criteria, and a final selection. However, the test cases employed a snowball process, which consists on asking for contacts to the contacted stakeholder. This alternative is included in the final evaluation method as a possible way to do the stakeholder identification process.

As the testing of the evaluation method was done for academic purposes, the stakeholder identification on the three test cases was difficult to do, as there is no authority to engage stakeholders. The proposed stakeholder identification was a systematic process of identification, assessment, and selection of stakeholders only if they comply with certain characteristics such as 'evident representation of an identified stakeholders, no conflicts of interest, and/or with the desired position at the time of the evaluation (avoiding former representatives). Instead, a snowball process was used mainly because it was the only way to reach the stakeholders in all the test cases. The researcher has low power to force authorities or representatives to join an evaluation process for academic purposes. This has two implications. First, the stakeholder representatives are chosen not by their capacity to represent the stakeholder's interest, but also by the disposition and the possibility to be contacted. Second, there was no selection process, as every contacted stakeholder was included in the test cases for capturing as much information as possible. However, according to the researcher, there are no negative implications for the development of the method. Moreover, the stakeholder representatives that participated in the test cases provide reasonable and (most likely) authentic representation.

HP2: Stakeholders that change their representation through time

One of the challenges of the stakeholder identification process is to identify different representatives that can represent the same stakeholder in different periods. This issue is partially solved using the lifecycle stakeholder map developed in test case 1 (SH7). However, an explanation needs to be included in the final version of the evaluation method to acknowledge this ambiguity.

Stakeholder representatives can address different periods of time and they might not have the same perspective towards the project. It is important to declare that a stakeholder has no 'objective' perspective; no one can claim to be the perfect representation of a stakeholder in cases other than singular stakeholders (e.g. owner) or highly hierarchical ones (e.g. small company). In 20 years of PPP contract, the public sector changes several times, the users might also change, the contractor's representative might change between construction and operation stage, etc. This challenge was already flagged in the prototype exercises presented in Chapter 6 (PE1).

A timeline and the judgement of the evaluators is necessary for deciding whether a stakeholder needs to be separated in two or more periods of time, or if the stakeholder must be represented by more than one representative in a jointed stakeholder workshop. For test case 3, the hospital's consortium was bought by another company during the operational stage, meaning that there are two companies that can have a different perspective over the same project in different periods of the project lifecycle.

HP3: Government disclosure regulations can restrict the development of the stakeholder perspective

Another refining insight that refers to the stakeholder representative is their restriction for providing information due to disclosure regulation. This problem must be acknowledged and solved by the evaluator using, for example, well-supported information to develop a 'missing perspective' in the missing perspective analysis. This means that the stakeholder cannot provide a public declaration, but the evaluator can infer the perspective based on the available information.

Refinement insight TP1 (test case 2) acknowledged the fact that some representatives cannot provide a judgement without agreeing internally with the stakeholder. Aligned with this problem, public authorities (finance department, sectorial minister, state or local government, etc.) could also have restrictions when representing a public agency because it might be risky to represent the interest of the government. To declare a formal judgement could have negative political impacts. However, for the evaluation to be legitimate, the perspectives should be included if they are considered relevant. To enforce a judgement could be one option if the evaluation is conducted by a formal authority, or through the missing perspective analysis.

HP4: Use key motivations and drivers of the stakeholder to help the criteria identification process

In the stakeholder workshop, the criteria identification from the stakeholder's representative is difficult to conduct. Therefore, the key stakeholder motivations and drivers could help in the criteria identification process. This suggestion is included in the final version of the evaluation method (Chapter 8).

The identification of evaluation criteria is difficult because stakeholders have to synthesise their view and present it into a set of criteria. During test case 3, a new strategy was tested to help stakeholders in this identification process. After finishing the criteria identification, the motivations and drivers were reviewed with the stakeholder to test if (1) all drivers are represented with at least one criterion, or (2) any identified criteria could be part of a driver that was not identified before. This re-check helps the stakeholder to clarify their perspective and to be consistent with the stakeholder's purpose. Another way to help the stakeholder is to ask for criteria in which there is no valuation; i.e. if all the impacts are positive, ask for anything negative that might be missing, or vice-versa.

HP5: Stakeholders can lack sufficient knowledge to judge a selected criterion

Some stakeholders are not capable of judging a criterion they already identified for assessing the project. For example, the user's representative (patients) have no information to judge a critical criterion like medical quality. However, is something they have a legitimate interest. Thus, the evaluation should address this lack of information and describe the level of understanding of every stakeholder. This should be part of the data analysis.

PPPs involve usually a new infrastructure (construction), and then an operational phase in which the services are provided. Stakeholders such as users have problems to visualise the project as a PPP and they understand the project just as a service. For these cases, more guidance needs to be done by the facilitator, to extract more directly the information they need. For example, the facilitator could redefine the project as the user experience, and then ask about parts of the service provision,

assuming that the project lifecycle is not able to be judged by this kind of stakeholder. This interference of the evaluator must be declared in the project evaluation (report).

HP6: Development of the evaluation conclusion structure

The way to conclude the evaluation was the last thing to develop, as it was designed using all the information from the three test cases. It was applied retrospectively to all test cases and incorporated to the final version of the evaluation method. The proposed structure of the conclusion has three components: (1) about the project-system, (2) about the PPP model, and (3) about the evaluation objective.

The first component of the evaluation, about the project-system, is based on the stakeholder judgement principles that were proposed in the conceptual framework PSEPM. The aim of this first evaluation component is to judge the successfully working project-system. The first principle 'willingness for repetitions' is based mainly on the stakeholder judgement regarding this principle contrasted with the criteria analysis; i.e. the evaluator verifies that stakeholders judged their willingness for repetition with a coherent argument (perspective). The second principle of minimum fairness (no excessive negative outcome) comes from the criteria analysis, searching for an excessively negative impact to at least one stakeholder. The evaluator's judgement must be also contrasted with the stakeholders' judgements on this principle. The third principle is the search for evident alternatives project options; their validity, their possible acceptance, and their rationality. The conclusion must include all the relevant information analysed during the previous processes (stakeholder identification and stakeholder workshop analysis). The main analytical support for concluding the project-system evaluation is the critical criteria analysis, the missing criteria analysis, and the negative criteria analysis.

The second component of the evaluation conclusion is about the PPP model. The aim of this component is to judge if the PPP model has a net positive effect on stakeholders' impacts. It should explain how relevant the procurement method is within the project-system and if there is any negative effect that can support a negative judgement of the PPP model. Additionally, it must recognise if the PPP benefits outweigh the PPP problems. The main analytical support for concluding the PPP model evaluation is the PPP effect analysis and the taxpayers' perspective analysis.

The aim of the project evaluation is the third evaluation conclusion component. The purpose of the evaluation and the use of its content and judgement must be clearly stated. In stage A (Chapter 6), experts provide feedback regarding the aim of ex-post evaluation. Four general aims were described for the evaluation: transparency and accountability, PPP legitimisation, industry learning, and government agency learning. To declare the purpose (using the four general aims) contributes to the effectiveness of this evaluation method.

Summary of test case 3 contribution to evaluation method development

Test case 3 is the last test case in which the evaluation method was applied. The data collection process (stakeholder workshops) included the refinements from the previous test cases (1 and 2). The data analysis process included the refinements of test cases 1, 2 and 3. The main contribution of test case 3 to the final evaluation method is the improvement of the criteria identification during the stakeholder workshop (SH4), and the design of the evaluation conclusion structure (HP6), which is the last component of the evaluation method.

7.5. Analysis of the testing on the three test cases

The iterative process for designing and refining the ex-post impact evaluation method was conducted using the three test cases. The evaluation method was applied to each test case, resulting in project evaluations for each of them. Each test case had four parts for presenting the project evaluation: (1) project overview, (2) stakeholder identification, (3) stakeholder workshop analysis, (4) evaluation conclusion. Additionally, a fifth part was included with the refinement insights obtained from the project evaluation.

This section analyses the three cases in terms of their results (project evaluations) and in terms of their contribution to the development of the evaluation method (refinement insights). Aggregated data of the three project evaluations and a summary of the refinement insights is presented.

Table 7-9 shows, in quantitative terms, the amount of data that was gathered from the stakeholder workshops; supporting its relevance for the design process. Overall, the evaluation method was tested and refined using three test cases, 25 stakeholder workshops were conducted, and 173 criteria were identified.

Table 7-9: Summary of the three test cases

Attribute/project	Test case 1: Social housing (SH)	Test case 2: Transport project (TP)	Test case 3: Hospital Project (HP)
Stakeholder identification			
Identified stakeholder	10	13	15
Stakeholder workshops	6	11	8
Criteria identification during stakeholder workshops			
Identified criteria	47	73	53
Positive impacts	43	41	45
Neutral impacts	1	7	5
Negative impacts	3	10	3
Judgements of stakeholders from stakeholder workshops			
Willingness for repetition	6/6	10/11	8/8
Excessive negative outcome	6/6	10/11	8/8
No evident alternative option	6/6	10/11	8/8
Taxpayers' perspective component during stakeholder workshops			
	Not applicable	Partially applied	Fully applied
PPP benefits			7
PPP problems			6

[Source: Author's]

Table 7-10 summaries the analyses that were applied to each test case. The reasons for not applying certain analyses were (1) no available data, (2) not considered useful for the project evaluation, or (3) another analysis was providing similar or equal information. Some analyses that were not employed in the test cases, but they are included in the final version of the evaluation method. This is not considered a problem because the aim is to present coherent results when applying an evaluation method that has multiple options. It was not considered rigorous to use all the available (proposed) analytical tools because it would have distorted the project evaluations that have the intention of showing a valuable and clear assessment of a real project.

Table 7-10: Summary of data analyses applied to test cases

Workshop analysis	Test case 1: Social housing (SH)	Test case 2: Transport project (TP)	Test case 3: Hospital Project (HP)
Consistency		X	
Critical criteria	X	X	X
Theme		X	
Negative criteria	X	X	X
Missing criteria and evaluation team	X	X	X
PPP effect	X	X	X
Quantitative		X	
Taxpayers' perspective			X
Pro-project bias			
Conflicting criteria			

[Source: Author's]

Table 7-11 summarises the refinement insights from the three test cases. The way they are incorporated into the evaluation process and how they applied within the testing process is described in the last column.

Table 7-11: Design inputs for evaluation method development and refinement

ID	Description	Action
Test case 1: Social housing project (SH)		
SH1	The evaluation method is also applicable to PPP projects that do not have an evident profit driver	Acknowledged
SH2	Stakeholders must provide reasonable expectations and not excessively optimistic ones	Included in the final version
SH3	The facilitator needs essential skills for conducting a workshop	Applied to test case 2 and 3, and included in the final version
SH4	'Perceived value' is better understood as 'actual value' when explained to the stakeholders in the workshops	Applied to test case 2 and 3, and included in the final version
SH5	Re-adjust the workshop structure to be conducted under 1 hour	Applied to test case 2 and 3, and included in the final version
SH6	Development of the data visualisation templates	Applied to test case 1, 2 and 3; and included in the final version
SH7	Development of lifecycle stakeholder map	Applied to test case 1, 2 and 3; and included in the final version
SH8	Development of data analyses	Applied to test case 1, 2 and 3; and included in the final version
Test case 2: Transport project (TP)		
TP1	Some stakeholders do not have representatives that can fully represent them	Acknowledged
TP2	Stakeholders' representatives could have no support to answer the judgement principles	Acknowledged
TP3	The ex-ante evaluation must be considered to have objective data	Acknowledged
TP4	Development of the taxpayers' perspective analysis	Applied to test case 2 (partially) and 3, and included in the final version
TP5	Development of Incentive-based stakeholder map for the stakeholder identification	Applied to the existing data of test case 1, 2 and 3; and included in the final version
TP6	Two judgement principle needed to be reframed	Applied to test case 3, and included in the final version
TP7	Suggesting criteria to stakeholders in the workshop can improve criteria identification	Applied to test case 3, and included in the final version
Test case 3: Hospital project (HP)		
HP1	Snowball stakeholder identification can be employed	Applied to test case 1, 2 and 3; and included in the final version
HP2	Stakeholders that change their representation through time	Acknowledged
HP3	Government disclosure regulations can restrict the development of the stakeholder perspective	Acknowledged

HP4	Use key motivations and drivers of the stakeholder to help the criteria identification process	Applied to test case 3 (partially), and included in the final version
HP5	Stakeholders can lack sufficient knowledge to judge a selected criterion	Acknowledge
HP6	Development of the evaluation conclusion structure.	Applied to test case 1, 2 and 3; and included in the final version

[Source: Author's]

Two refinement insights from test case 2 are highlighted for their relevance within the research and in the design of the evaluation method. The taxpayers' perspective (TP4), and the incentive-based map for stakeholder identification (TP5). These two components, developed during the testing process, are relevant because they provide conceptual knowledge on how to deal with the complexities that were originally mentioned for PPP projects. The taxpayers' perspective analysis contributes to the theoretical understanding of the taxpayer regarding the use of PPP, and the incentive-based map provides an applicable tool for understanding a project-system.

All the refinement insights from the test cases contributed to improve the ex-post impact evaluation method which will be fully described in the next chapter.

7.6. Summary

This chapter describes the process of testing and refining the ex-post evaluation method in three test cases: a social housing project in Chile, a transport project in Chile, and a public hospital in Australia. The three are PPP projects that are completed or in operation.

The social housing project was assessed employing the developed evaluation method. It considered 6 stakeholder perspectives, and 47 impacts were gathered. The project was judged as meeting the three principles of a successful project-system. The refinement insights from this test case are:

- (SH1) the evaluation method is also applicable to PPP projects that do not have an evident profit driver;
- (SH2) stakeholders must provide reasonable expectations and not excessively optimistic ones;
- (SH3) the facilitator needs essential skills for conducting a workshop;
- (SH4) 'perceived value' is better understood as 'actual value' when explained to the stakeholders in the workshops;
- (SH5) re-adjust the workshop structure to be conducted under 1 hour;
- (SH6) development of the data visualisation templates,
- (SH7) development of lifecycle stakeholder map, and
- (SH8) development of data analyses.

The transport project was employed to test again the evaluation method. 11 stakeholders were included, 73 impacts were gathered, and general approval of the project can be suggested except for one stakeholder that was completely against the project. The project was judged as meeting the three principles of a successful project-system. The procurement method was analysed separately from the project-system and two issues were analysed to judge the success of the PPP model: the deficient business model, and the benefit of infrastructure availability. The evaluation suggests that the PPP benefits outweigh the PPP problems. The refinement insights from this test case are:

- (TP1) some stakeholders do not have representatives that can fully represent them,
- (TP2) stakeholders' representatives could have no support to answer the judgement principles,
- (TP3) Consider the ex-ante evaluation to have objective data,
- (TP4) development of the taxpayers' perspective analysis,

- (TP5) development of Incentive-based stakeholder map for the stakeholder identification,
- (TP6) two judgement principle must be reframed, and
- (TP7) suggesting criteria to stakeholders in the workshop.

A public hospital was the final test case in which the evaluation method was applied. 8 stakeholders workshops were conducted and 53 impacts were gathered. The project was judged as meeting the three principles of a successful project-system. The procurement method was analysed, and the taxpayer perspective (Value for Money approach) was included for the first time in the process. It was judged that the PPP benefits outweighs the PPP problems. 6 refinement insights were synthesised:

- (HP1) snowball stakeholder identification can be employed,
- (HP2) stakeholders that change their representation through time,
- (HP3) government disclosure regulations can restrict the development of the stakeholder perspective,
- (HP4) use key motivations and drivers of the stakeholder to help the criteria identification process,
- (HP5) stakeholders can lack sufficient knowledge to judge a selected criterion, and
- (HP6) development of the evaluation conclusion structure.

The project evaluations (contextual results) are considered a contribution to knowledge in their specific contexts (transport, social housing, health), however, they must be considered academic exercises and not official information.

Chapters 6 and 7 have shown the development of the evaluation method. Chapter 8 will show the outcome of the process, which is the procedure for conducting the PPP ex-post impact evaluation.

CHAPTER 8

THE ARTEFACT - PPP EX-POST IMPACT EVALUATION METHOD

8.0. Introduction

This chapter describes the final deliverable, an ex-post impact evaluation method for PPP projects. Employing Design Science methodology, this is the artefact that was designed and tested. This is a self-contained chapter and can be employed independently for applying the evaluation method and it summaries all the development from the previous chapters. The aim of this chapter is to provide the general audience with a tangible artefact that solves a real problem: A procedure to conduct a PPP ex-post impact evaluation. Therefore, this chapter repeats and summarises most of the research described in the previous chapter and tries to avoid any reference to the literature.

This chapter is divided into three parts. The first part summarises the design process in section 8.1. The second part is the final evaluation method explained through five sections (8.2 to 8.6). Finally, section 8.7 provides the final feedback from the experts regarding the evaluation method.

8.1. Overview of the PPP ex-post impact evaluation method

This evaluation is to be applied to PPP projects, after they are operative or completed. It aims to assess the impacts of the PPP project on their multiple stakeholders. For doing so, it captures the perspectives of each stakeholder and then judges if the project is working successfully as a project system in which all stakeholders interact.

The evaluation is based on the conceptual framework Project Success Evaluation Pyramid Model which employs the Project Success concept. Successful projects are the ones that can be successfully judged from the different stakeholder perspectives. This concept is applicable to PPP projects because it addresses one of the complexities of PPP projects which is the multiplicity of views. PPPs are megaprojects that are financed with private resources, are supposed to deliver a public good or service, have long-term contracts, provide gains and losses in political terms, and they are tinted with ideological principles. In other words, they have many impacts and judgements surrounding them.

To address the multiplicity of views, this evaluation method, first, defines an impact as a rational judgement of a stakeholder over a self-identified criterion. Second, this method provides a systematic and guided procedure to gather all the different perspectives from the relevant project stakeholder in terms of impacts. The overall judgement of the project, which contains the multiple views, is addressed by understanding a project as a system of stakeholders that interact (voluntarily and involuntarily) to deliver an endeavour. This evaluation method aims to assess if the project-system worked successfully, establishing certain principles in which all the stakeholders and their views are contained. These principles are the theoretical baseline for judging a PPP project while incorporating the multiple stakeholder perspectives.

The final evaluation method is summarised in Figure 8-1; it has five main components: (1) evaluation setup, (2) Stakeholder workshops preparation, (3) stakeholder workshop, (4) analysis and judgement, and (5) evaluation conclusion. This structure of the evaluation follows the logic of other existing methods such as the guideline for implementing a Performance Management System (Neely et al., 1995); the PPP procurement guidelines from the European PPP expertise center (2011), and Gertler et al. (2016) book for conducting an impact evaluation.

PPP ex-post impact evaluation procedure

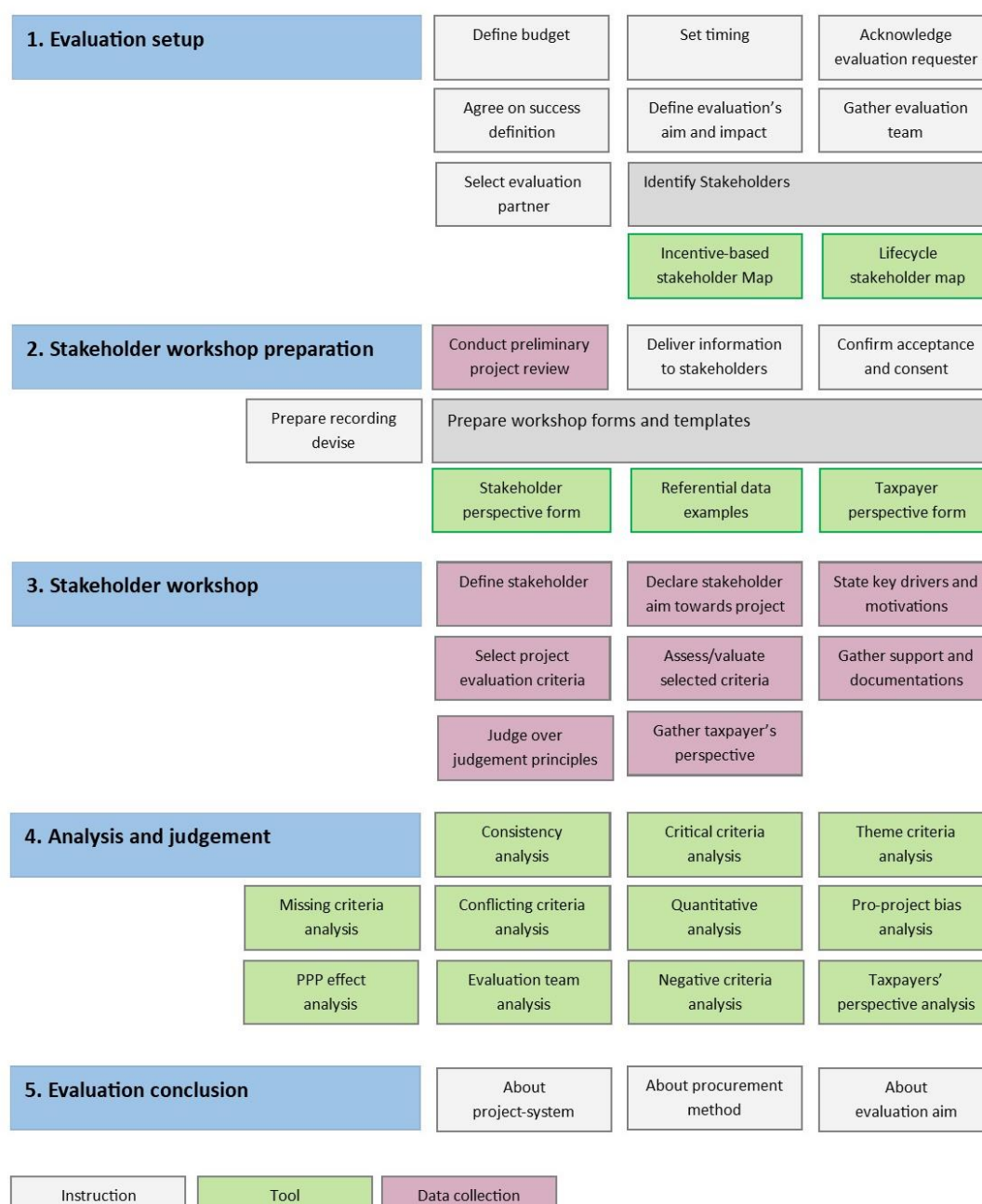


Figure 8-1: Final evaluation method

[Source: Author's]

8.2. Evaluation setup

The first of this evaluation is to define, at an early stage, the characteristics and structure of the evaluation. In the final evaluation report, the timing, the budget, the evaluation team, the evaluation partner, the requester, and the use of the results need to be decided, declared, and explained. The preparation process also involves the stakeholder identification in which the lifecycle stakeholder diagram and incentive-based stakeholder map are recommended.

Budget

The budget can influence the quality and expertise of the consulting group that conducts the research, the fieldwork options (on-site interviews or video/audio interviews), hours of document analysis, and the incorporation of key policy-makers and consultants for the data analysis. Generally, by increasing the magnitude of the evaluation, it increases the cost of it, but also the reliability of its results.

Timing

How much time is available for conducting the evaluation? Does the evaluation need to respond to any milestone? Depending on the available time, the evaluation can change in scope and rigor. Therefore, this variable should be taken into account for scheduling the planning, data collection, data analysis, and final report.

The stage of the project in which the evaluation is going to be conducted is a relevant information to declare and support. The moments in time can be after the construction stage, after the contract has ended, or at a certain time after starting the operational stage (searching for stable operation). A possible bias can be declared here if the evaluation can be influenced by the time of applications, such as the inauguration of the infrastructure, a related controversy, a political problem, or an election year. Figure 8-2 exemplifies how to synthesise the decision for choosing a moment for conducting the evaluation.

Milestone	Pros	Cons	Results
End of contract			
Concession renewal			
Start of service contract after PPP			
Other			

Figure 8-2: Possible milestones for assessment application form (to be used by the evaluator)

[Source: Author's]

Evaluation team

Depending on the scope and magnitude of the evaluation, there are several roles that need to be included within the evaluation team: (1) evaluation manager, (2) a data collection team, (3) a form design and data analysis team, and (4) a data and policy analysis.

Evaluation partners

There are public entities, private entities and universities that can act as evaluation partners. The prestige, the access to information, or the time restriction can influence the decision of who to select as an evaluation partner.

Requester (sponsor)

The requester of the evaluation needs to be declared in the final report, together with the amount of money conferred, and the total cost of the evaluation. This is a common practice in audit reports, such as the Victorian auditor general. For example, the audit report for public hospitals 2016-17 states 'We used the results of these audits in preparing this report. The cost of preparing this report was AUD 377 400, which is funded by Parliament' (Victorian Auditor-General's Office, 2017).

The role of the requester is important also for the evaluation governance. Ideally, the requester should have enough power to, for example, enforce the contribution of the entities that need to provide information. The requester could have an important impact on the evaluation results, because the

official declaration of some stakeholders might be difficult to obtain, especially in the case of public entities. In this case, the requester should have a supporting role or be part of the evaluation team.

Results and rationale

Depending on the requester, the timing, and the scope of the evaluation, the results should have certain impacts and objectives. Generally, there are four main objectives for a PPP ex-post impact evaluation: (1) accountability, (2) PPP legitimacy, (3) industry learning, and (4) public sector learning. The way that the results of the evaluation will be employed must be declared to establish the evaluation's relevance and impact. Figure 8-3 presents a prototype form for declaring and defining the use and the impact of the different deliverables of the evaluation.

Reason for the assessment	Why the evaluation should be conducted?
Objectives of the assessment	Who is going to be affected and how? What it will impact and how?
Assessment results	Who might use the components of the evaluation and how? (criteria data, stakeholder identification, overall judgement, etc.)
Judgement principles	Why each judgement principle is important in this project evaluation?

Figure 8-3: Prototype of the rationale summary table (to be used by the evaluator)

[Source: Author's]

Agreement of theoretical success definition

This evaluation employs multiple stakeholder perspectives, which are gathered and analysed systematically in order to reach a judgement about the PPP project impact. Therefore, this evaluation employs stakeholder judgement principles, which are attributes or characteristics that support the thesis that the project-system, in which all stakeholders are interrelated, was developed in a successful way. These principles help to judge the sum of all stakeholder perspectives employing a common ground.

The evaluation should explain and declare what are the principles that will be employed for assessing the multiple stakeholder perspectives. Three principles are presented as possible options, meaning that they can be used partially, or to complement any other method.

- **Willingness to repeat the project:** Are the stakeholders willing to repeat the same project and receive the same outcomes assuming that they could choose from the same original options? This willingness to repeat the project supports the assumption that the stakeholder can accept the gaps in expectations when they are compared with the benefits they received.
- **No excessive negative outcome:** Does any stakeholder received an evident negative outcome that could have been avoided? An excessive negative outcome is an impact that is considered 'out of the norm' which means, out the normal spectrum of positive and negative impacts that stakeholder can get for entering a project. This requirement provides a minimum of 'fairness' in which an excessive and evident negative outcome cannot be accepted for considering a project as successful.

- **No evident better alternative option:** In the case that the project does not want to be repeated and it has evident negative outcomes; do the stakeholders think that the project is still the best option from the original alternatives that were available at the conceptualisation of the project (including the 'do nothing' option)? Not having a better alternative is a strong argument for judging a project as successful. It helps to modify (calibrate) the expectations with the real available options and the estimated results.

8.2.1. Stakeholders identification

The project is a system of stakeholders that are interrelated to drive an endeavour (section 5.6). Therefore, the stakeholder identification can be conducted using (1) the incentive-based stakeholder map, and (2) a lifecycle stakeholder diagram. Additionally, the topic of identifying stakeholder representatives is also discussed.

Incentive-based stakeholder map

Figure 8-4 shows the incentive-based stakeholder maps which represents a project-system in which stakeholders are identified according to their main roles and incentives towards the endeavour. The aim of this diagram is to help to identify stakeholders and their role in the system. The categories of the map have a general aim and description, but the boundaries are not discrete, which means that a stakeholder can be in more than one category or change category depending on the evaluator. This diagram should also help to analyse possible conflicts of interest in which certain views should or should not prevail. The diagram recognises that the interrelations between stakeholders can be voluntary or mandatory and can be explicit (through a contract) or implicit. Table 8-1 describes the aim and function of each stakeholder's category within the incentive-based stakeholder map.

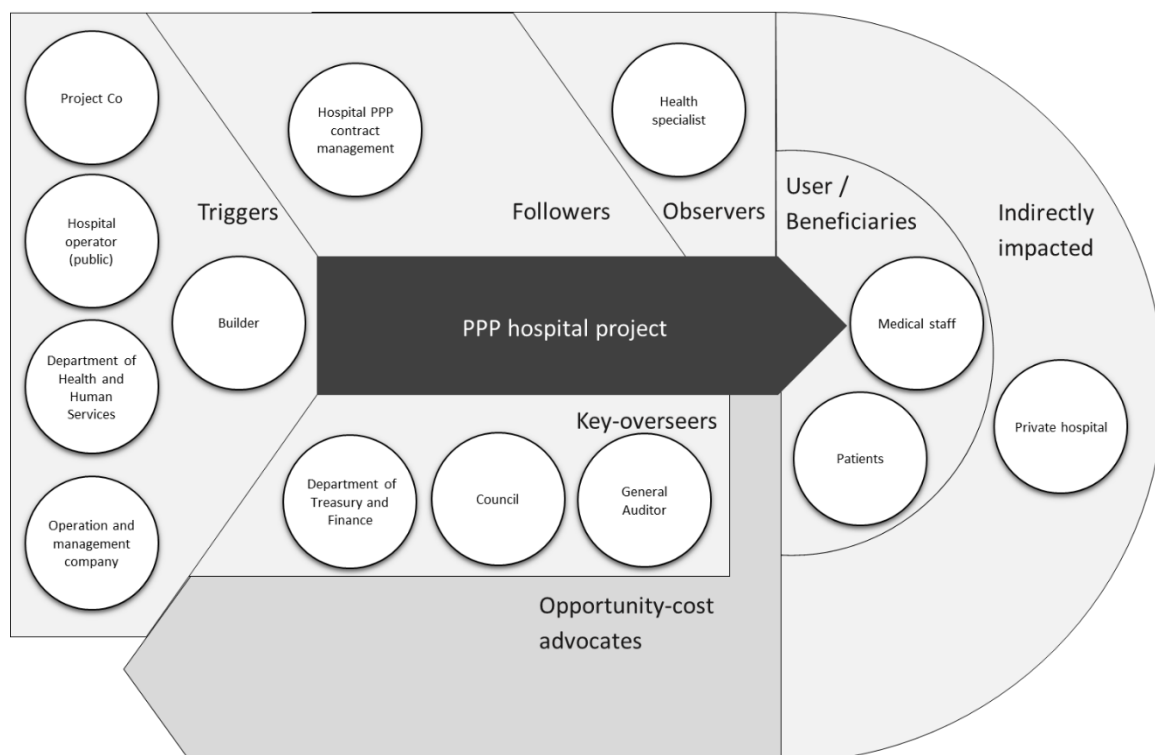


Figure 8-4: Example of Incentive-based stakeholder map applied to a PPP hospital project

[Source: Author's]

Table 8-1: Incentive-based stakeholder map categories

Stakeholder type	Description	Power/interest	Examples in a PPP
Triggers	They start the process, are in favour of the project, should be the ones doing most of the effort, and they should be accountable for the final outcome	Their power and interests should prevail in the general cases, unless a strong argument can support a different 'view' or if there is a major interest in the project	A public agency, geographical public authority, private entity, NGO or think tank, etc.
Key-overseers / Key-supporters	They need to be 'in favour' or need to approve the project explicitly or implicitly. Their support is essential for the project feasibility	They do not necessarily need to be interested in the project, but they should have enough power to be considered in this category	Treasury, overarching authority (presidency). Environmental authority, tax system, justice service, etc.
Followers	They are affected by the process of the project (construction, implementations, temporal deliverables), or they provide or affect the project	They usually do not have power, and their interest depends on each case. Big decisions are not taken by them, because of the lack of power or interest	Suppliers, workers, competitors (for the purpose of the project), alternative solution developers, control groups, indirectly affected by the construction stage, etc.
Beneficiaries	They should receive the positive impact that is explicit in the purpose of the project	In a PPP, beneficiaries do not have a voluntarily impact of the project (as in a business), because PPPs usually provide services that are considered rights (health, education, transport, etc.)	In a business is the client; in a public policy are a subgroup of citizens; in a public project are a specific community, etc.
Indirectly impacted	Because of the purpose of the project, there are stakeholders that receive or are affected positively or negatively by the project outcome	They usually do not have power nor interest, but the impact they receive should be relevant when addressing the project scope	The surrounding community, local or extended market, public institutions (affecting their performance), etc.
Observers	They have an overarching view of the project. They can provide an impartial, wider (maybe worldwide) and long-term analysis of the project; probably better than other stakeholders	They should not have enough power, but they can influence future decisions and investment agenda	International organisations, biased and unbiased media, national authorities in a local project, professional organisations, universities, researchers, etc.
Opportunity-cost advocates	They are organised groups that can present a real alternative to the project that is being developed, including the 'do nothing' option. They differ from stakeholders that can be against the project without a real solution	They have a high interest in the project and their power can be (1) able to change the project, (2) to reconsider the project, or (3) to achieve their proposed option	Environmental groups, think tanks, community groups, other government agencies that share a budget, political parties, etc.

[Source: Author's]

Lifecycle stakeholder diagram

To identify stakeholders throughout the project, the lifecycle stakeholder diagram is proposed. It has two axes, one in which stakeholders are categorised between the public sector, private sector, user/beneficiaries, and community/ society. The other axis is the project lifecycle divided into conceptualisation, design, construction, operation, and dispose. The aim is to identify stakeholders that are relevant at certain stages of the project. Figure 8-5 exemplifies this diagram.

Stakeholder group	Conceptualisation	Design	Construction	Operations	Disposal	Stakeholder representative
Private sector						Project Co head director
						Not included
						Not included
Public sector						Hospital executive director
						DHHS PPP contract administrator
						Hospital PPP contract administrator
						DTF executive
						Not included
						Not included
User/beneficiary						Former hospital junior staff
						Former hospital patient
						Private hospital director
Community						Not included
						Not included

Figure 8-5: Example of a lifecycle stakeholder diagram applied to PPP hospital project

[Source: Author's]

This diagram helps to identify and trace how in the project's lifecycle stakeholders change, their expectations change, and their leader and representatives change. The changes of a stakeholder through time must be described and supported, highlighting the reasons for including a stakeholder with more than one representative.

Stakeholder representation

A stakeholder is considered a group of people because they share similar attributes and have 'stakes' in the same project. However, they are not a homogeneous group and they can have large differences in their judgement, in the understanding of the project, and in their role on it. The ideal scenario would be to have an agreement with all the people in a stakeholder group, but in the real world, this is almost impossible. A representative of the stakeholder must be appointed, and it could be a single person or a small group. The selection of the representative can have a misinterpretation risk if the representative has internal conflicts, or his/her personal incentives overcome the stakeholder incentives.

The identification process must be declared in the evaluation report, including the identification strategy. Two options are suggested:

- A broad review of the stakeholders and the possible representatives. Then, the use of exclusion criteria to argue why certain representatives should not be part of the evaluation

process. The reasons could conflict of interest, lack of representation for the entire stakeholder, etc.

- A snowball could be employed, suggesting that the representatives were searched through another stakeholder. This method is not recommended but can provide access to stakeholders that are difficult to contact. The evaluation team should judge between having a non-optimal representative or not having access to a stakeholder.

8.3. Stakeholder workshop preparation

The workshop with the stakeholders is the main data gathering process of the evaluation. It is conducted by a facilitator employing a set of preform templates to guide the workshop. However, the main outcome of the workshop is to obtain a reliable and clear perspective of the stakeholder; this means that the facilitator can alter the procedure to achieve the objectives of the workshop.

Five elements are important to consider when preparing a workshop: (1) a preliminary review of the project, (2) delivering preliminary information of the workshop to the stakeholder, (3) recording, (4) acceptance of the stakeholder, and (5) evaluation forms.

A preliminary review of the project

The role of the facilitator is not just to guide and gather the information and judgements from the stakeholders. The facilitator should have a complementary role by guiding the conversation towards the objective, which is to get a sincere, complete and thorough perspective of the stakeholder. Therefore, the preparation of the workshop includes a preliminary review of the project and their existing documents, approvals, audits, media coverage, and other kinds of information related to the project that can help the facilitator to manage the workshop in the best interest for the evaluation. Conflicts and controversies found in the review should be analysed and considered when discussing the stakeholder's perspective. It must be noticed that the facilitator bias should be reduced to the minimum; this is the influence of the facilitator on the stakeholder perspective. The role of the facilitator is to add points to be discussed, which could have not been recognised or remembered by the stakeholder representative. During the workshop, the facilitator should try to present, in the most impartial way, any facts that he/she thinks it may affect the perspective. For example, an existing audit can flag design problems that affect the service provided in the infrastructure. This problem should be posed as 'is the design a criterion to consider?' from your perspective, is the design an issue? why? The presentation of an audit to a stakeholder can drive the workshop towards accepting all the criteria of the audit as theirs, which would be a manipulation of the stakeholder's perspective. That should not occur as the idea is to extract the most honest judgement from the stakeholder.

Preliminary information of the workshop

Delivering to stakeholders a preliminary information about the workshop can help to make it more effective, giving time to the stakeholder's representative to think in advance about the criteria and the aim of the stakeholder.

Evaluation forms (templates)

The workshop employs three main forms (1) stakeholder perspective form, (2) examples of aims and criteria form, and (3) taxpayers' perspective form. Depending on the workshop's format, they can be printed or digitally used with the appropriate devices. Their use is explained in the stakeholder workshop section and the templates are in appendix B.

Recording

Recording the session is a good backup alternative in cases where the filled forms missed important information, or if it is necessary to search for comments that can affect the subsequent analysis.

Acceptance

A consent form is appropriate in any case and should be designed accordingly with the evaluation objective, and aligned with the institution that supports the evaluation (General-Auditor, NGO, Ministry, etc.)

8.4. Stakeholder workshop

The stakeholder workshop is the main data collection process. It has the aim of gathering eight components: (1) stakeholder's definition, (2) stakeholder's aim towards the project, (3) key motivations and drivers of the stakeholder, (4) evaluation criteria, (5) assessment of the evaluation criteria, (6) Support and documentation, (7) judgment of stakeholder principles, and (8) taxpayers' perspective over Value for Money.

(1) Stakeholder's definition

The representative should declare the scope of the stakeholder he/she is representing and be specific of what is not within his/her/their representation. In a governmental agency, the extent of the stakeholder can be the government's view, an entire ministry, a sub-department, a specific team, or even just a specific operating team within the organisation. In a community group, the representative should clarify the geographical boundaries, time boundaries (maybe the stakeholder was active only during a certain period), and maybe state any possible differentiation with a similar stakeholder. For example, a contractor can be differentiated between the contractor's owner and the contractor's personnel, assuming that both can be a relevant stakeholder for the evaluation.

(2) Stakeholder's aim towards the project

The stakeholder must agree with the facilitator on which is the aim of the stakeholder toward the project. The objective is to answer why the stakeholder is important for the project, and what is the contribution or the effect over the project. The general aim of a stakeholder could be just for a single project (e.g. the Single-Purpose-Vehicle in a PPP project) or applied to many projects. For example, a community group is not a stakeholder just for a single infrastructure project, but also for any kind of intervention or change that affects the community. The stakeholder's aim must be declared, explaining how it is aligned with the evaluated project.

(3) Key motivations and drivers of the stakeholder

The idea of this component is to go deeper in the detail of the stakeholder's description. The facilitator and the representative should discuss and agree on the key concepts that drive the stakeholder on its normal performance and functioning. For example, the public authority can have a core goal of efficiency of public expenditure, quality of life, or compliance. Another stakeholder can have identity with the community, community problem solutions, management of shared assets, etc. Conversely, a private entity can have core concepts that drive their functioning such as profit, prestige, or survival. These concepts need to be well defined, as they are the source for assessing the consistency of the perspective; the alignment between what I am, why I am here, what is my goal, and then, what I expect from the project. Consistency analysis (described in section 8.4) could suggest that the representative is assessing the criteria using their personal judgement, to the detriment of the stakeholder's view he/she represents. It can also help to assess the relevance of certain criterion

depending on how loose or erratic is the connection with the key concepts. Appendix D contains the stakeholder workshop form that includes the key motivations and driver concepts.

(4) Evaluation criteria

The stakeholder representative, with the help of the facilitator, will agree on what are the evaluation criteria that the stakeholder will use for assessing the project that is being evaluated. The evaluation criteria should include processes and outcomes, trying to analyse the entire project lifecycle. The form to fill in with the stakeholder is presented in Appendix D.

(5) Assessment of the evaluation criteria

After defining the criteria that represent the stakeholder perspective towards the project, they need to be assessed using three pieces of information: an expectation, a perception (actual value), and the judgement of an impact. The reasoning behind this structure is to try to systematically address an evaluation criterion by gathering (1) what was the expectation over that criterion (if there was any), (2) what is the real value perceived at that moment, and then, (3) what is the impact, measured in terms of that criteria (positive, neutral, or negative).

The information gathered for the expectation and perception could be presented in qualitative or quantitative terms. If the criterion is quantitative such as cost or profit, it is expected to have a quantitative expectation and perception. However, if the criterion is qualitative, then qualitative arguments can be used to describe it. A criterion such as 'relationship between the public entity and the project Co (consortium)' can have an expectation defined as 'regular according to contractual terms', and then a perception such as 'proactive relationship above the contractual terms'. This is evidently stating that the perception was over the expectation, and consequently (or most probably) a positive impact will be judged in this case. Moreover, if there is any qualitative or quantitative criterion that cannot be presented in real terms due to, for example, confidentiality issues, symbols such as 'expected 100 and perceived 100+' can be employed. The idea is to have a clear argument that supports a judgement on that criterion.

(6) Support and documentation

In cases where the criteria are controversial or with opposite views (between two stakeholders), the reasons that support the differences between expectation and perception can be registered with documentation and references. This documentation could also be part of the preliminary study conducted by the facilitator and the evaluation team. Additionally, depending on the scope and relevance of the evaluation, it is possible to go back to the stakeholders down the track to re-review certain judgements and ask them to back up their statements.

(7) Judging on stakeholder principles

The stakeholder principles discussed and defined at the evaluation setup are the third component of the stakeholder workshop. The stakeholders need to judge if they agree on the proposed principles according to the best of their knowledge. The meeting of three principles described in this procedure is the support for the thesis that a project-system is successfully working.

The first principle applied to the stakeholder workshop is, 'is your stakeholder willing to repeat the same project and receive the same impacts that are described in the previous stage of this workshop?' The idea of this question is to guide the stakeholder in making a judgement that employs the same criteria that were selected by him to evaluate the project. It must be understood that it is a hypothetical question that helps to analyse the success of the project and is not a real question for future decisions. The stakeholder can have a negative opinion towards repeating the project because

he/she would improve some things before repeating it. However, this is not the aim of the question, the idea is to guide a mental cost-benefit analysis in which all the positive, neutral and negative impacts can be mentally weighted, judging if the sum of all the impacts is worth the project (if there is a final positive outcome).

The critical criteria, which are the main deal-breakers for a stakeholder, are obtained by agreeing on which criteria would make a change in the judgement (passing from repeating to not repeating). The intention of this mechanism for obtaining the critical criteria is to close the 'cognitive gap' (or psychological gap) between the respondent and the situation that is being responded. To ask directly for the critical criteria is a different exercise than asking which criteria will make the willingness for repetition judgement to change; as it places the respondent in 'decision-making position'. This step has a psychological support, but in the end, is the facilitator and his/her skill which can alter, *in situ*, the way the critical criteria are identified. For example, if the stakeholder is not sure on the answer, maybe it is possible to go back to the criteria, revise them and then select the critical ones without any connection to the judgement principle. Moreover, answering the question can lead to discovering hidden criteria that were not rationally mentioned before, and even to change one of the key concepts or motivations of the stakeholders.

Even though the process is presented as a sequential procedure, it is possible to go back and forth to modify or add information, to improve the quality of the stakeholder perspective statement.

The second principle to ask is 'is there any stakeholder that received an excessive (over the norm) negative impact?' This question tries to find fairness within a project-system. This means that, in order for the project to be considered successful, none of the stakeholders that are part of the same project-system should gain a negative impact that is 'over the norm'. Bankruptcy, jail, life-threatening issues, political instability, war, etc. are impacts that can be judged as 'out of the norm'. Negative impacts 'within the norm' would be for example a negative profit for a private entity (part of the risk of entering a business), expropriations with compensations, unforeseen environmental issues possible to be mitigated or reduced, or loss of political capital for unpopular decisions.

The third question for the stakeholder to answer is 'is there any evident better alternative that should have been done instead of the selected option?' This principle is supported on the idea that a successful project is relative to the potential success of other alternatives. If there is no evident alternative to replace the chosen one, then, there is a strong argument in favour of considering the project as successful. This should also address the possibility that even the best option could be badly implemented, and then the theoretical discussion would be if an alternative option would have had the same poor level of implementation. Another issue to consider is that this is intended for big, major, essential changes in the project, as it is easy to find improvements to the actual project that can be considered an alternative option. For example, the same project, but delivered on time should not be considered an alternative option of the same infrastructure. An additional floor for future expansion should also not be considered an alternative project.

(8) Taxpayers' perspective over Value for Money

The last component of the workshop is the specific study of the procurement model over the project. It is named taxpayers' perspective because it is not possible to have a representative of a taxpayer. However, it is possible to assume that the taxpayer has a legitimate concern for ensuring that every PPP should provide better Value for Money than the Traditional Procurement alternative. Therefore, that last component of the workshop is to recognise PPP benefits and problems according to the following theoretical development.

The qualitative Value for Money assessment is a systematic judgement over a comparison between the PPP benefits and PPP problems for a single project. The information presented is not quantifiable, but a judgement can be made with clear and transparent arguments. The question is presented as 'do the PPP benefits outweigh the PPP costs and problems?'

Equation 8-1: PPP problems and benefits

$$PPP\ benefits > PPP\ problems?$$

[Source: Author's]

This PPP ex-post impact evaluation method operationalises this theory by, first, presenting a form to the stakeholders in which all the possible PPP benefits, PPP problem, and avoided TP problems are described; gathered from the existing literature. Second, the stakeholder's representative discusses and selects which of the attributes or issues happened in the project that is being assessed. And third, the data is analysed by assessing coincidences between stakeholders, the validity of the judgement (according to their perspective), and inconsistencies within the stakeholders. The final report should include this information as clear as possible to demonstrate the existence of agreed benefits and problems.

8.5. Stakeholder workshop analysis

After the stakeholder workshops have concluded, the gathered information should be synthesised and analysed. Several types of analysis can be conducted, however, is not mandatory to apply all the options mentioned here, as it depends on the scope of the evaluation and, more importantly, on the information gathered, which is the ultimate result of the project. The evaluator team should decide which analysis is pertinent to pursue the aim of the evaluation.

Ten analyses are described in this procedure:

1. consistency analysis,
2. critical criteria analysis,
3. theme analysis,
4. negative criteria analysis,
5. missing criteria and evaluation team analysis,
6. conflicting criteria,
7. PPP effect analysis,
8. quantitative analysis,
9. pro-project bias analysis, and
10. taxpayers' perspective analysis.

They should be applied to illustrate a holistic view of the project-system that includes all the stakeholders and to support the final judgement of the evaluation team.

(1) Stakeholder consistency analysis

This analysis should verify that the declaration of each stakeholder is coherent and aligned between (1) the declaration of the stakeholder (scope), (2) the purpose of the stakeholder towards the project, (3) the motivators and drivers of the stakeholder, (4) the selection of evaluation criteria, and (5) the valuation of the selected criteria. The way to present this information is through a network diagram that connects each node. This analysis is based on the Investment Logic Map (ILM) employed by governmental agencies in Australia (Doty, 2008). It can be done by analysing each attribute and connecting it to the logical sequential attribute. A specific driver will be connected to one or more evaluation criteria, and the purpose can be disaggregated in order to show that between the purpose

and the key concepts there is a logical connection. Any conflict of interest or relevant bias should be described, highlighting how it can influence the evaluation.

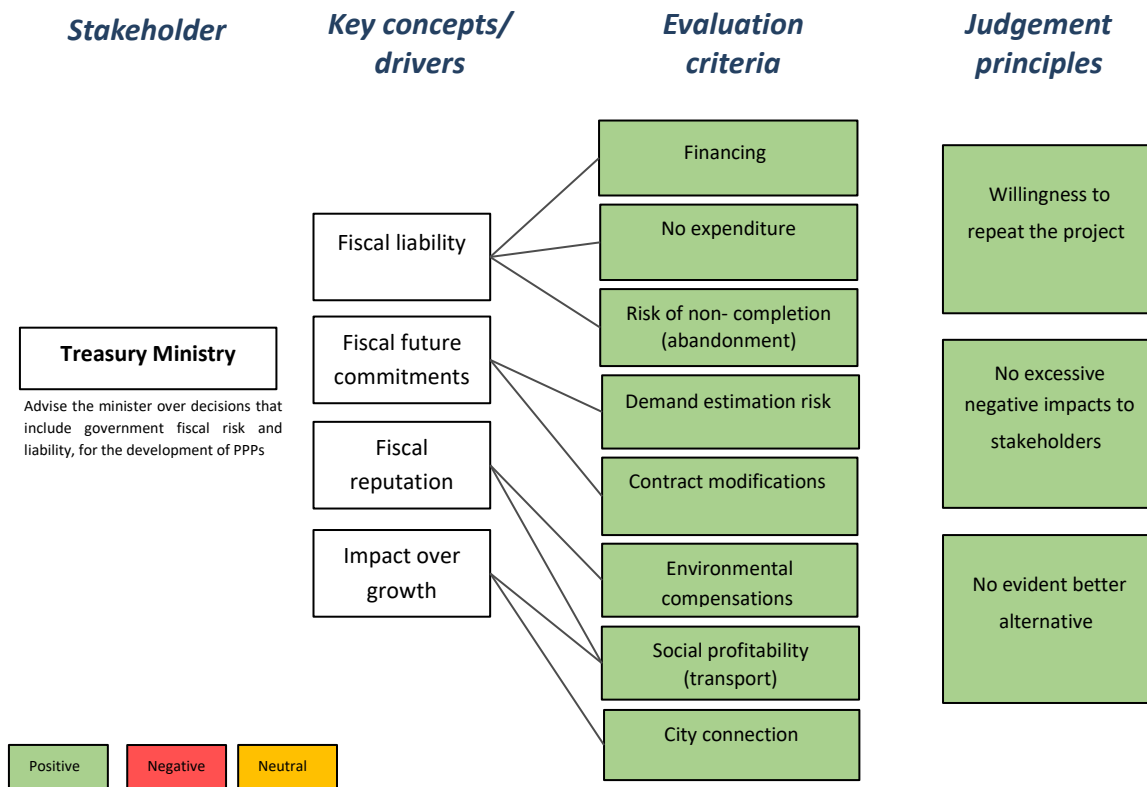


Figure 8-6: Example of a network diagram to describe stakeholder consistency analysis

[Source: Author's]

Figure 8-6 shows how the concepts can be connected. The principle of transparency and clarity of information should prevail. It needs to show any issue that can concern the general public. This analysis must be easy to understand, using colour-coded connections, different sizes, etc.

The answer of each stakeholder to the judgement principles (willingness for repetition, no excessive negative impacts, no evident better alternative) should also be analysed for consistency. The evaluation team must decide if the stakeholder judgements are reasonably supported on the criteria selection and valuation. In the example provided, there is no evident conflict as the aim of the stakeholder is consistent with the drivers, and in turn, each driver is represented by one or more evaluation criteria. Additionally, this stakeholder has a positive judgement on the three judgement principles, which is consistent with having only positive judgements on all the selected criteria.

(2) Critical criteria analysis

To judge if the project has met their main goal, this analysis focuses on the critical criteria of all the stakeholders. This analysis presents the core criteria and the deal-breaker that should be employed to judge if the project is successfully working as a system.

	Critical criteria		Non-critical criteria									
Project Co	Profit		Social profitability	Relationship with PI ministry	Collateral satisfaction	Time	User satisfaction	Highway connectivity with the system				
Contractor	Profit by construction	Profit as part of the consortium	Value of engineering component	Time	Quality of project team and partners	Public private relationship						
Public Infrastructure Ministry	Highway utilization	Business model	Infrastructure quality	Intervention during construction	Land value surrounding highway	Integration to the city	Private profit	Expectations over highway's impact				
PPP unit at the Public Infrastructure Ministry	Pay per use	Service availability	Service quality	Quality standard	Safety	Connectivity	Travel time	Benefit to users	Positive business model			
Social Development Ministry	Social NPV-IRR		Public participation	Sinergy with the city	Macro-design for all highways	Environmental costs	Renegotiations	Connection highway-city				
Treasury Ministry	Social profitability (transport)		City connection	Financing	No expenditure	Demand estimation	Impact on economic growth	Risk of non-completion (abandonment)	Contract modifications	Environmental compensations		
Local council	Quality of access (to highway)		Connection with other areas of the city	Maintain capital gain (property)	Value creation in surrounding area							
User towards road safety	Accident rate		Safety design and information for users	Car-pedestrian relationship	Relieve existing local traffic	Connectivity						
Generic user	Travel time reduction		safety	Connection with the city	Price							
Community organisation (no-highway)	Urban integration		Creation of social capital	Impact of participation	Effect over car use/transport	Beneficiaries	Effect over urban connectivity					
Urban experts	Project purpose	Adaptation to geographical characteristics	Generate landscape	Generate public spaces	Synergy with mobility	Public-private-people interaction	Experience the city	City connection				

Figure 8-7: Example of critical criteria analysis applied to a transport project

[Source: Author's]

(3) Theme criteria analysis

The criteria presented by all stakeholders can be separated into themes and analysed as different components of the project. This approach can help to analyse specific parts of the projects that have not worked well. Additionally, it is possible to seek 'agreements' between stakeholders in certain areas.

	Direct highway purpose		Connection to city (in transport)	Business / finance		Urban / social	Process	Indirect impacts	
Project Co	User's satisfaction	Social profitability	Highway connectivity with the system	Profit	Time		Relationship with PI Ministry	Collateral satisfaction	
Contractor	Value of engineering component			Profit by construction	Profit as part of the consortium		Quality of project team and partners		
				Time			Public private relationship		
Public Infrastructure Ministry	Highway utilization	Infrastructure quality	Integration to the city	Private profit	Business model		Intervention during construction	Land value surrounding highway	Expectations over highway's impact
PPP unit at the public infrastructure ministry	Service quality	Quality standard	Connectivity	Pay per use					
	Service availability	Benefit to users		Positive business model					
	Travel time	Safety							

Figure 8-8: Example of theme criteria analysis applied to a transport project

[Source: Author's]

(4) Negative criteria analysis

This evaluation method has an 'optimistic' starting point. Similar to the presumption of innocence, in which a person is declared innocent until the contrary is proven; this evaluation presumes that the project is successful until the opposite is suggested. Therefore, a thorough analysis of the neutral and negative impacts is conducted. The goal is to qualitatively analyse each of these criteria, discussing its relevance, its support and logic, its agreement throughout the different stakeholders, and if it can be considered a deal-breaker which would be a reason to judging the project system as 'not successfully working'.

(5) Missing criteria and evaluation team analysis

Evaluators and auditors can also address specific areas of the project (financial, environmental, etc.). The evaluator's perspective can be a legitimate and useful perspective to add to the stakeholder network of the project-system. Moreover, an expert's perspective can address additional issues that were not recognised by the other stakeholders. For example, if according to the experts, the geopolitical impact is not 'in the interest' of any stakeholder as an evaluation criterion, the expert team can flag this issue, analyse it, and value the geopolitical impact over the country. They can also suggest that a specific stakeholder should have this criterion as part of their perspective because it should be their responsibility and it should be under their radar. This evaluation team perspective is under the category of 'observers' in the incentive-based stakeholder diagram (section 8.1.1).

(6) Conflicting criteria analysis

Another analysis is the search for conflicts between stakeholders. One of the reasons why PPPs are difficult to evaluate is because of the conflict between the different stakeholder views, i.e. the same outcome could be a positive impact to one stakeholder and a negative impact to another one. The question to answer is 'is there any lack of coherency between the stakeholders' perspectives?' This analysis is the one that needs more judgement and deep understanding from the evaluation team because it can create a loss in legitimacy if it is not done thoroughly. First, for the conflicting criteria and valuations, the reasons behind the conflict need to be presented and exposed. Second, the documentation that supports the valuations (judgement) could be analysed, assuming that more documentation should lead to better reliability of the valuation. Third, scenario analysis could help on speculating the different options, 'if A happened, then B should have happened...'. Fourth, a judgement from the evaluators is expected to decide whether one stakeholder is not valuating the criteria accordingly, or if there is a misconception of the outcome. If in a conflicting issue, both stakeholder perspectives are legitimate, then the power of one stakeholder should prevail above the other. At any point, it is possible to go back to the stakeholders for more information regarding the reasons, arguments, documents, or data that supports the judgement about a certain criterion. If it is a non-quantitative criterion, then a larger argument must be gathered.

A guide for analysing conflicting the criteria is through a theoretical exercise: The addition of all the expectations should not make the project unfeasible. This means that a project should, at least theoretically, be capable to meet all the different expectations of the stakeholders. For example, if a hospital is expected to have the newest technology (a health authority perspective), but it is also expected to be within a maximum budget (national budget authority perspective) that makes it the technology impossible to buy, then the project is not 'feasible' and there is a conflict between the expectations.

(7) Quantitative analysis

Several quantitative analyses can be conducted as a support to the holistic view that involves all stakeholders. Trends and evident results can be shown using quantitative approaches, including:

- Positive impacts/total impacts: This ratio shows the percentage of positive impacts in relation to the total gathered criteria. At first glance, stakeholders and criteria are not weighted by any relevance. This ratio can present a clear result, such as a big percentage of positive or negative valuations.
- Critical positive impacts/total critical impacts: This ratio makes the difference of including only the critical criteria that were gathered from each stakeholder. This can show a better view of the 'big relevant picture'. It can also present clear evidence to support the second judgement principle; none of the stakeholders was excessively impacted negatively.
- PPP effect over the project: After selecting the criteria that are considered affected by the procurement method, a ratio between the 'PPP criteria and total criteria can show the level of involvement that the procurement method has over the impacts of the projects, and consequently, discuss the overall impact of the procurement method on the stakeholders.
- Level of conflict between stakeholders: After assessing each of the criteria, the ratio between (1) the criteria in which there is a conflict with other criteria and (2) all the criteria, can show how aligned or how 'controversial' is a project measured by the inconsistencies within the stakeholders.

Quantitative analysis should complement the qualitative ones, as any weighting process is, in the end, considered arbitrary and lacks reliability.

(8) Pro-project bias analysis

A specific bias needs to be addressed when analysing the results of the evaluation. Pro-project bias is the bias of judging positively the outcomes of the project because a negative result might influence the continuity of the project or could lead to a search for accountability in which the responder could be involved. This bias is difficult to mitigate as the evaluation method is built on the principle that all perspectives are valid. However, to flag this bias in the workshops and in the report improves the quality of the evaluation and the reliability of its report. The inclusion of measurable information and documentation when possible can also mitigate this bias.

(9) PPP effect analysis

It is not possible to determine with certainty the influence of the procurement method over the impact of the project on each stakeholder, which are measured by the recognition of evaluation criteria from each of the stakeholders. However, the evaluation team can draw over the theoretical conjectures of how the procurement method could affect those impacts (Figure 8-9). This analysis is based on the principle that is not possible to analyse separately the impact of the project from the impact of the procurement method. Therefore, for addressing one, the other must also be addressed. A reasonable option is to show which are the impacts that could have been affected by the procurement method, and try to explain the reasons behind it, comparing with other procurement methods. This means that if one impact is clearly influenced by the procurement method, it must be specified how the criteria could have changed positively or negatively if another procurement method was employed instead.

	High effect of PPP		Medium-low effect of PPP		No effect of PPP				
Project Co	Profit	Relationship with PI ministry	User satisfaction		Social profitability	Collateral satisfaction	Time	Highway connectivity with the system	
Contractor	Profit as part of the consortium	Quality of project team and partners	Profit by construction	Public private relationship	Time	Value of engineering component			
Public Infrastructure Ministry	Private profit	Business model	Expectations over highway's impact	Highway utilization	Infrastructure quality	Intervention during construction	Land value surrounding highway	Integration to the city	
PPP unit at the Public Infrastructure Ministry	Service quality	Quality standard	Pay per use		Safety	Connectivity	Travel time	Benefit to users	
	Positive business model	Service availability							
Social Development Ministry	Renegotiations		Public participation		Sinergy with the city	Macro-design for all highways	Environmental costs	Social NPV-IRR	Connection highway-city
Treasury Ministry	Financing	No expenditure	Risk of non-completion (abandonment)	Environmental compensations	Impact on economic growth	Social profitability (transport)	City connection		
	Demand estimation	Contract modifications							
Local council			Maintain capital gain (property)	Value creation in surrounding area	Quality of access (to highway)	Connection with other areas of the city			
User towards road safety	Car-pedestrian relationship		Accident rate	Safety design and information for users	Relieve existing local traffic	Connectivity			
Generic car user			Travel time reduction	Price	Connection with the city	Safety			
Community organisation (No-highway)	Impact of participation				Creation of social capital	Urban integration	Effect over car use/transport	Beneficiaries	Effect over urban connectivity
Urban experts	Public-private-people interaction		Generate public spaces	Adaptation to geographical characteristics	Generate landscape	Synergy with mobility	Project purpose	Experience the city	City connection

Figure 8-9: Example of PPP effect analysis applied to a transport project

[Source: Author's]

This workshop does not include the possibility of stakeholders to judge if the impacts they received are linked to the procurement method. However, this could also be added if the stakeholders have relevant expertise over the procurement method.

(10) Taxpayers' perspective analysis

As explained in the stakeholder workshop section (8.3) an additional component to assess the taxpayer view has the aim of answering to a missing perspective that cannot be solved using the general approach of the evaluation method. The taxpayers' perspective is difficult to obtain because no one can really represent *in situ* the taxpayers' interest. Nevertheless, it should be a very important perspective as is the one that, in theory, makes PPPs feasible in the first place. Therefore, a more theoretical approach is taken to address this perspective. In the stakeholder workshop, the PPP benefits, the PPP problems, and the presumably avoided Traditional Procurement problems were identified. Table 8-2 shows an example of identified benefits, including all the relevant stakeholders. This information is analysed to judge if the benefits of the PPPs can outweigh the problem due to the use of the PPP model. Is the PPP model giving more Value for Money than the traditional option? The judgement should qualitatively make a case in which the benefit could outweigh the problems or vice versa. The theoretical support for this judging process is explained.

Table 8-2: Example of PPP benefits recognised by the stakeholders of a health project

Category	Attribute	PPP unit (treasury)	Project Co	Hospital PPP operator	Health Ministry	Private hospital	Total
Finance	Better risk allocation	x	x	x	x	x	5
Quality	Incentive for high-quality maintenance	x	x	x	x	x	5
Quality	Clear performance standards	x	x	x	x	x	5
Quality	Better residual infrastructure after transfer to public	x	x	x	x	x	5
Finance	Funding availability		x	x	x	x	4
Efficiency	Operational efficiency due to bundling		x	x		x	3
Efficiency	On-time	x	x	x			3
Efficiency	On budget	x	x	x			3
Efficiency	Integration of design, construction, financing, operational, maintenance, and refurbishment responsibilities	x	x			x	3
Efficiency	Lower estimated lifecycle cost compared to PSC	x	x			x	3
Finance	Payment after commissioning		x		x	x	3
Quality	Incentive for availability	x	x	x			3
Efficiency	Improvement of construction quality for long-term use.	x	x				2
Finance	Design and fit-for-purpose risk on private sector	x	x				2
Finance	The potential for lower cost of assets				x	x	2
Finance	funding certainty for maintenance	x			x		2
Quality	High level of innovation	x	x				2
Quality	Transparency	x	x				2
Finance	Off-sheet balance treatment		x				1
Finance	Economically feasible projects		x				1
Finance	Impact on the economy (due to the PPP method)						0
Finance	Availability of funding for using the PSC budget on a different project						0
Finance	Urgency (due to funding availability)						0

[Source: Author's]

As stated earlier, the main problem of the PPP decision-making is the impossibility to quantify (in reliable terms) the benefits and problems of the PPP. To address this constraint, the concept of 'value' is introduced.

Perceived value is a personal (or agreed in a group) valuation of anything, based on moral and cognitive principles. Value is difficult to quantify but it is possible to rank it from a spectrum of alternatives. A judgement can be made on what delivers more or less value than another thing; i.e.

what is more valuable to the stakeholder. In order to make a structured logic thinking, the judgement is presented in a mathematical form, assuming that the components cannot be quantified.

To deliver a project using a PPP model, the following inequation must stand:

Equation 8-2: Value of PPP benefits and problems isolating PPP extra cost

$$v(\text{PPP extra cost}) < v(\text{PPP benefits}) - v(\text{PPP problems}) + v(\text{avoiding TP problems})$$

The value of the PPP extra cost should be less than the value of the PPP benefits, discounting the value of the PPP problems, and adding the value of the avoided TP problems. PPP benefits and avoided TP problem are presented separately because it clarifies the identification of both; however, they complement each other in most of the cases and they should not be identified more than once.

The qualitative Value for Money assessment should consider that, due to the specific nature of the project, the use of the PPP method can provide different benefits and problem. Moreover, depending on the project's nature, the valuation of these benefits and problems will be different in each case. Theoretically, an imaginary project can have no benefits except for one, which is so valuable due to the nature of the project, that it can outweigh all other problems and costs; and consequently, make a case for preferring that project as a PPP assuming the extra cost.

The valuation of each benefit and problem also depends on to which option is being compared. According to Heald (2003), the real decision-making is not between the PPP and the Public Sector Comparator (PSC) in terms of efficiency but in terms of funding availability. If the availability of public funding allows the PSC to be built; then, that is the option to compare with the PPP. If there is no available funding, the fall-back position needs to be considered instead of the PSC. The Fall-Back Position (FBP) is the option that should be done with the real available funding, which is probably a mitigation strategy, temporary improvement, or a time buying option by outsourcing the missing service. For example, the value of the PPP benefit in term of urgency (infrastructure available earlier) will be much higher if the PSC is not an option due to the funding availability, and moreover, if the FBP has a minimum scope.

The evaluation should suggest possible ways to judge if the value of the PPP (extra cost) is less than the value of the other side of the equation. Even though it is tempting to make a quantification process by putting scores and weights to every stakeholder and criteria; the principle of this evaluation is to make transparent reasoning for making a judgement, rather than quantifying and getting a final score.

8.6. Evaluation conclusion

The ex-post impact evaluation of PPP projects from the perspective of multiple stakeholders should conclude with clear statements about three aspects: about the project-system, about the procurement method, and about the conclusion objective. It is recommended to include a data summary of the evaluation to support and the conclusions and provide a clearer report (example in Table 8-3).

Table 8-3: Example of an evaluation summary applied to a health project

Project	
Health infrastructure PPP project	
Identified and included stakeholders	
Project Co (included)	Council
Builder	General Auditor
Operation and Management company	Medical Staff (included)
Hospital Operator (public) (included)	Patients (included)
Department of Health and Human Services (included)	Private Hospital (included)

Hospital PPP contract management (included)	Surrounding neighbours
Department of Treasury and Finance (included)	Health specialists and consultant
Workshop data collection	
1) Stakeholder's definition	5) Assessment of the evaluation criteria
2) Stakeholder's aim towards the project	6) Taxpayers' perspective (Value for Money analysis)
3) Key motivations and drivers of the stakeholder	7) The judgment of stakeholder principles
4) Evaluation criteria (critical and non-critical)	
Workshop analysis	
Critical criteria analysis	10 critical criteria were gathered, all assessed positively by their stakeholders
Negative criteria analysis	Risk allocation, contract design, and future-proofing are negative impacts. They are considered valid, but none of them is critical.
Missing criteria analysis	The contractor, the council, and the surrounding community perspectives were not reached. However, no support for evident negative impacts can be suggested.
PPP effect analysis	The negative impacts that are an effect of the PPP model are not considered decisive to conclude an overall negative impact of the PPP model
Taxpayers' perspective (Value for Money)	Gathering benefits and problem, the judgement can be 'overly' simplified by stating that an (almost certain) better maintenance benefit outweighs an (uncertain) negative higher capital return rates.
Stakeholder judgement analysis	All the reached stakeholders have a positive judgement towards the three principles. This is a strong argument to support a successfully working project-system

[Source: Author's]

The first part of the conclusion should be about the project-system. Here, the critical criteria, the possible deal-breakers, and the judgments of the stakeholders should be highlighted. The judgement principles of successful projects should be discussed and presented to accept/reject the successful functioning of the project-system.

The second part of the should be about the procurement method. A case should be made for judging if the PPP benefits outweigh the PPP problems or vice versa. This conclusion must be supported by the Value for Money analysis and the PPP effect analysis of the stakeholder impacts.

A third part of the conclusion should be about the objectives of the evaluation. Four general purposes for a PPP ex-post impact evaluation were presented: accountability, PPP legitimacy, industry learning, and public-sector learning. Each of the selected objectives should be discussed, highlighting how the results will influence the different aims. Industry and public-sector learning can be achievable by providing lessons learned that are linked to (1) critical criteria (2) negative impacts, (3) criteria in which the perception was evidently above the expectations, and (4) conflicts between stakeholders. Accountability and PPP legitimacy purposes are more difficult to achieve because the evaluation's sponsor needs to have sufficient power to officially engage the stakeholders (e.g. an academic exercise could not provide sufficient credibility) and be able to impact future decisions.

This procedure has described all the steps that are necessary to conduct an ex-post impact evaluation of PPP project from the perspective of multiple stakeholders. Examples of the results can be found in Chapter 7 of this thesis. Appendix B provides evaluation forms and templates for conducting the workshops.

8.7. Feedback from experts regarding the evaluation method

The results of the evaluation method, described as the results of the test cases, were presented to experts (a subgroup of the expert in stage A) for a final feedback process. This last step is to confirm the effectiveness and efficiency of the method, discussing its applicability and usefulness.

The results of test case 2 were presented to interviewee B (described in section 4.4.1) under a 1-hour meeting. According to interviewee B, the proposed method can be comparable to other existing methods for evaluating sustainable infrastructure development (not disclosed for confidentiality issues). However, the proposed method was recognised by the expert as more 'concrete' and straight-forward, showing clear steps and visual results. According to the expert, it is aligned with the aim of the public sector, which is to plan and conceptualise infrastructure following a comprehensive process with more public participation. In fact, a presentation of the method in greater detail to a bigger audience in the Chilean Public Infrastructure Ministry was suggested. Additionally, to link the method with an infrastructure planning tool was also discussed. The idea would be to work towards the best possible result, measured using the proposed ex-post impact evaluation method. Feedback about the process was also proposed; to include a clear way to define the boundaries of some stakeholders, how to consider beneficiaries in public infrastructure? Where to stop counting and enlarging the stakeholders?

The results of test case 2 were also presented to interviewee C, who said that the three judgement principles are correct. They are clearly conceptualised, providing a qualitative guideline for infrastructure project assessment. Another positive comment was that the method is operationalisable, by showing clear results and systematically addressing subjective information. The Value for Money process was discussed, emphasising that for the Chilean context, the comparison between public financing and PPP is artificial, as the possibility to generate mega infrastructure projects without a PPP model is rare and almost impossible (especially in transport projects; health infrastructure can be different). Weighting stakeholder by relevance was discussed, concluding that the method must address a way to understand the relative power between them. Public participation was mentioned as the current trend and that this approach addresses and (up to a certain point) solves this issue, generating accountability for what is expected by selecting criteria of interest and then judging it. The presentation (slides) of the results received some suggestions and mistake corrections. Interviewee C asked to be more consistent in the presentation of results, in order to follow a logical path towards the conclusion of the evaluation. According to him, the conclusion was presented isolated from the result analysis. Also, it is necessary to clarify two things: (1) differentiate the information regarding the general process from the specific test case that was evaluated, and (2) leave no ambiguity between positive and negative issues to avoid misinterpretation of results. The feedback was included in the test case 2 results and in the final evaluation method.

Within the Design Science approach, the goal of any artefact is to be effective and efficient (Hevner et al., 2004, p. 81). In this thesis, effectiveness is understood as the way the method solves the problem of evaluating PPP projects, including, for example, clear results, with high quality, reliable data, assertive analysis, and a reasonable and convincing final judgement. According to the author, the developed artefact is effective considering three things. First, it has a theoretical support that validates the way of thinking about PPP evaluation. Second, the testing of the proposed evaluation method employed real data and had an outcome that shows a real project evaluation. And third, the results were discussed with experts, validating the researcher's conclusion: the proposed method can be used to ex-post evaluate the impact of a PPP project, obtaining reasonably valuable results.

Efficiency is the second overall criteria to assess an artefact. For this research, this means that the method needs to be balanced between the employed resources and the obtained outcomes. The evaluation method should be easy to follow, with clear diagrams, without using excessive

stakeholder's time, able to be understood by the taxpayer, among others. According to the author, who was also the one employing the method in real projects, this attribute is also achieved. The proposed evaluation method is clear, easy to understand by the stakeholders who participated, it provides many tools to visualise the results, and (essentially) can provide valuable information with the resources that were employed.

The next chapter discusses the research study, including the design process and the main deliverable.

8.8. Summary

This chapter describes the final deliverable of the design process, an ex-post impact evaluation method for PPP projects. This evaluation procedure is based on the development of the conceptual framework 'Project Success Evaluation Pyramid Model' (Chapter 5).

The evaluation method has five main components: (1) evaluation setup, (2) stakeholder workshop preparation, (3) stakeholder workshop, (4) workshop analysis, and (5) evaluation conclusion.

First, the evaluation preparation consists of defining and declaring the timing, the budget, the evaluation team, the evaluation partner, the requester, and the use of the results. It also involves the stakeholder identification.

Second, to prepare the stakeholder workshop, five issues are described: (1) preliminary review of the project, (2) delivering preliminary information of the workshop to the stakeholder, (3) recording, (4) acceptance of the stakeholder, and (5) evaluation forms.

Third, the stakeholder workshop is the core component of the evaluation. It has the aim of gathering eight components from each stakeholder: (1) stakeholder's definition, (2) stakeholder's aim towards the project, (3) key motivations and drivers of the stakeholder, (4) evaluation criteria, (5) assessment of the evaluation criteria, (6) Support and documentation, (7) judgment of stakeholder principles, and (8) taxpayers' perspective over Value for Money.

Fourth, the analysis of the gathered information can be done employing the suggested approaches: consistency analysis, critical criteria analysis, theme analysis, negative criteria analysis, missing criteria and evaluation team analysis, conflicting criteria, PPP effect analysis, quantitative analysis, pro-project bias analysis, and taxpayers' perspective analysis.

Lastly, the conclusion of the evaluation should draw over three main topics: the project system (that includes the procurement method), the procurement method, and the evaluation objectives (accountability, PPP legitimacy, Industry learning, and public-sector learning).

Experts reviewed and gave final feedback to the evaluation method, which was presented with a test case evaluation report. They provided mostly positive comments, validating the employed theory and the way of addressing a PPP evaluation. This suggests that the artefact achieved the overall criteria of being efficient and effective to solve the problem that was intended for.

The next chapter will discuss the research project.

CHAPTER 9 DISCUSSION

9.0. Introduction

The adopted Design Science approach to develop an ex-post impact evaluation method for PPP projects was conducted and described throughout this thesis. Chapters 1, 2, 3 and 4 are introductory chapters describing the problem, the existing knowledge and the research method. Chapters 5, 6 and 7 are the main outcomes of the design process, showing how the artefact (evaluation method) was developed, tested, refined and then delivered as a procedure in Chapter 8.

This chapter discusses the validity and utility of the developed evaluation method, and the rigour of the design process. It argues on how the research contributes to the existing knowledge, and why certain decisions were taken throughout the design process.

The different aspects of the research are discussed under five sections: about the evaluation theory, about the PPP context, about the employed test cases, about the evaluation method, and about the design process.

9.1. About evaluation theory

This research studied the existing methods and approaches to evaluate, ex-post, a PPP project (literature review in Chapter 3). Chapter 5 described the conceptual framework 'Project Success Evaluation Pyramid Model', which is the theoretical development for evaluating multiple perspectives in a single project.

The developed evaluation method can be theoretically located between two big areas of knowledge, in public policy for its methodological aspect, and in construction management for its contextual aspect. The reviewed literature consisted on a set of different evaluation approaches categorised by discipline (or area) and by the level of stakeholder involvement. Here it was stated that most of the evaluation approaches did not include multiple stakeholders. This study has contributed to this aspect by addressing the multiplicity of actors.

To better 'locate' the proposed evaluation method within the existing evaluation theory, three distinctions are made. First, comparing it with the traditional survey method. Second, placing it between two types of impact evaluation. Finally, discussing how the proposed method deals with quantitative approaches.

The conceptual framework reshaped the evaluation concept towards a judgement role

Described by Owen and Alkin (2007) an evaluation is a technical process of enquiry, data collection, and data analysis that ends on a judgement of worth. Every aspect of the technical sphere needs to be synthesised on a decision; this is the aim of the evaluation, to decide and take action. This core concept is reshaped for this research, putting an extra step in the process. From where does the evaluation takes the information for making the judgement? Is there any restriction to the sources of information? When the project is conceptualised as a project-system and the Project Success Pyramid Model presents the stakeholder at the base of the pyramid, an implicit idea is stated. The evaluation starts by acknowledging the different perspectives that need to be considered. This is opposite to the conventional way of understanding an evaluation that can start with a very experienced team that might or might not take into account the existing perspectives that surround the evaluated project. From the public policy domain, this change in the understanding of the evaluation concept means more transparency and awareness of the existing social diversity. It is aligned with public participation schemes and with the need to democratise government processes. The proposed conceptual framework and subsequent evaluation procedure make very difficult for a practitioner to not consider

everyone's legitimate perspective. Even if the proposed method is not followed correctly, the core idea can be used to set a ground level. An evaluator should judge every perspective, rather than create a single perspective for themselves; now their role becomes to be judges of the project system, rather than expert viewers with unbalanced power.

The proposed evaluation method differs from a survey

The developed ex-post impact evaluation method is based on the principle that stakeholders select and decide what criteria are relevant to them, providing support for its selection and valuation. The evaluation is developed through workshops and the facilitator is a relevant actor. A possible option would have been to employ a survey method for addressing the problem, which is very common in the construction management discipline. The survey has certain advantages in terms of time-consumption and number of respondents, being possible to gather more information from large stakeholders such as users, personnel, investors, etc. However, a survey is an evaluation method that does not allow discussion, consensus, or deeper reasoning. The gathering of the stakeholder perspectives would have been superficial, limiting the subsequent judgement of the evaluators. Nevertheless, surveys could be very helpful for supporting a judgement within a stakeholder perspective (e.g. to confirm the user's perspective towards a project). It must be noticed that here the survey is understood as a methodological tool that can be employed for conducting many evaluation approaches (as the ones identified in the literature review).

The proposed method is between two types of impact evaluation: qualitative and evidence-based approach

Section 3.5 described that impact evaluations can be understood in two different ways. They can focus on evidence-based approaches that try to, quantitatively, demonstrate that the intervention X created a change in the Y indicator of interest (Copestake, 2014). They can also be understood as a perception-based approach that describes the impacts of a project/intervention usually based on beneficiaries' perception. The developed method for PPP projects is placed in the middle of the two understandings (closer to the perceptive approach), but also at a higher level. It is mostly perceptive because it relies on the stakeholder's perception, which can or cannot be supported by evidence (evidence is encouraged, but not mandatory). It is also at a higher level because it solves the question of 'what is the real Y indicator of interest that is being observed?'. As many views can have many indicators, this approach systematically addressed all the possibilities and try to solve existing conflicts. It differs from a positivistic view of defining the 'correct' indicator to measure the impact of a project.

Quantitative methods have a limited use within the evaluation method

The proposed evaluation method underused quantitative methods such as ratios, analytical hierarchy process, pairwise comparison, or fussy logic approaches. The reasons for adopting this limitation are three. First, de Wit (1988) concluded that any attempt to generate a quantitative result for Project Success would be only artificial. The PPP ex-post impact evaluation method is based on the conceptual framework 'Project Success Evaluation Pyramid Model', which is a theoretical approach to reach a judgement of success. De Wit's conclusion is the cornerstone of this study, generating an approach that employs 'principles' instead of scores, and follows a systematic approach of judging, instead of valuating (scoring). The second reason is that this study is under a structural-systemic epistemological approach (based on Aristotelian thinking) that employs qualitative methods for understanding the structure behind a phenomenon (in this case, the judgement of stakeholders). Conversely, a Cartesian-Humean approach that focuses on cause-effect is consistent with quantitative methods and, therefore, not appropriate in this case (Toomela, 2010). Even though statistical tests can be used to suggest a judgement to be right or wrong, the proposed procedure is intended to use reasoning and qualitative analysis in a structured and systematic way, mitigating the risk of presenting numbers that

can mislead the conclusions of a holistic analysis. Third, as one of the evaluation objectives is to show legitimacy, the reliability of quantitative methods can be most appropriate when quantifiable criteria are discussed. However, when qualitative information is an essential component of the phenomenon (political and ideological issues), the use of quantitative methods can reduce the legitimacy of the project and procurement method judgement; this is what has been suggested for the Value for Money estimation.

9.2. About the PPP context

This section discusses the issues related to how the study is aligned or contribute to the research and practice of PPP projects.

Identifying and understanding the PPP impacts should contribute to a better management of the PPP factors

The outcome of the research is an evaluation method that was designed for PPP projects. The mainstream discussion regarding PPPs (aim, effectiveness, ideology, etc.) was not part of the scope, but the way this research contributes to the overarching discussion needs to be addressed. Many studies focus on the factors that affect PPPs performance (eg. Debela, 2019; Hardcastle et al., 2005; Osei-Kyei & Chan, 2017), assuming the criteria and impacts are clear. For example, the PPP contract design is determinant for many of the impacts that the PPP project generates to the stakeholders. The use of a Design-Build-Maintain or a Design-Build-Operate can make a substantial difference in the impacts that the project will have during its lifecycle. Also, the payback option (return to the public sector) directly affects how much Value for Money is the taxpayer getting (different from the initial estimation). This research studied how to holistically assess each PPP by their outcomes and impacts to each stakeholder, which is a departing point for the discussion of how to deal, solve, improve, or keep those impacts. Here, the proposed method can be described as the tool to better analyse what is happening in the projects. It is aligned with Rodney Turner's words about the importance of studying the criteria before studying the factors (Ika, 2009). A better recognition and understanding of the impacts should lead to better analysis of the factors of successful PPP projects.

How the proposed evaluation compares the PPP option with the traditionally procured option

The trigger for adopting a PPP model is because it presumably provides better Value for Money to the taxpayer when compared to a traditionally procured project (P. Raisbeck, C. Duffield, & M. Xu, 2010). Then, a reasonable claim is to question how the proposed evaluation method can address this difference if it only evaluates PPP projects without evaluating their counterparts. Here, two insights from the research are important. First, the proposed evaluation method can also be applicable to Traditional Procurement projects, especially if they have a complex stakeholder structure comparable to a PPP. Second, research about how PPPs differ from traditionally procured projects is extended (Engel et al., 2014), and that is why this study fulfils the gap in knowledge by proving a common ground for evaluating the impact of a PPP project in an independent/isolated way.

The incentives to apply the proposed evaluation method are not clear

In Chapter 6 (refinement insight EF2), experts stated that there are no clear incentives to apply any kind of PPP ex-post evaluation. PPPs are megaprojects originally supported and led by the government and the private sector. After the operational stage or after a reasonable time of operation, not the government, nor the private sector have a real interest in evaluating the ongoing project because they risk facing accountability issues. However, there should be an effort from societal groups, independent organisations, universities, etc. to push the agenda towards better ex-post evaluation processes. This claim is already seen by the evaluation association, where practitioners have a real interest in evaluating more and better. Nevertheless, the proposed method contributes to this agenda by

providing a step-by-step procedure for evaluation, which is the first barrier when making an evaluation compulsory; not having a method. In the same way that the Victorian Government established that projects above AUD 50 million should be considered for a PPP option; projects with certain attributes should be always evaluated ex-post.

9.3. About the test cases

The research tested the developed evaluation method using three test cases: (1) the PPP social housing project in Chile; (2) the PPP transport project in Chile, and (3) a health PPP project in Australia. The three are intended to represent the diversity of existing PPP arrangements when testing the developed evaluation method. The primary aim of using the test cases is to develop the PPP evaluation method. The test cases are real projects in which the prototype evaluation method was applied for refining the procedure (evaluation method). However, the testing process delivered three project evaluations which are considered relevant collateral results of the study. These project evaluations are not essential to this research project, but they are considered a contribution to the contexts of the three test cases. Therefore, these project evaluation results are also discussed.

Limited use of the results of the test cases

The results and the data obtained in the testing process should be recognised as a collateral contribution of the study. This contribution has certain methodological limitations that need to be highlighted to avoid incorrect extrapolations. The testing process is an academic exercise and the results cannot enforce any resolution in the contexts of those project; i.e. stakeholder perspectives are not considered officials. Also, as the evaluation method was being developed, the testing was done with ‘underdeveloped versions’ of the method, meaning that the evaluation procedure presented in Chapter 8 was applied partially. The conclusions of every test case, which is an overall judgement of the project, must be taken as exploratory answers and as referential information about the projects. Nevertheless, the type of information that was gathered using the evaluation method is new and innovative, being valuable for lessons learned, strategic decision-making and for future research in those contexts (social housing, transport, and health infrastructure).

The test case selection could be biased towards projects with overall positive outcomes

The selection of the test cases was driven by the availability of stakeholders and information. For each case, key people were contacted for recommending certain projects and providing access (or guidance) to their stakeholders; this was essential to have access to the test cases. The recommendations on what projects to select were made mostly by PPP experts that were also part of the first research stage (Chapter 6). The recommendation and the willingness of stakeholders to participate in the evaluation suggests an unintended bias in the project selection. In general terms, all the recommended projects were considered good, successful or above de standard; no evidently bad or failed project was suggested. This possible bias could be either because failed PPP projects are not very common, or because stakeholders tend to avoid been exposed to a project that is not considered good. It is a risk to search for responsibility and accountability of past actions. Nevertheless, this bias is not considered a relevant problem for the research because the final aim is to test the evaluation method rather than to get official information and enforceable actions from the test case results. Moreover, the evaluation method is not limited to further improvement, and a ‘stress test’ will be faced if the method is applied to projects that are considered an evident failure or badly developed.

9.4. About the evaluation method

The deliverable of the thesis is a procedure to conduct a PPP evaluation by gathering a single perspective from each of the involved project stakeholders through a series of stakeholder workshops. It is the main contribution to practice. The procedure contains five parts: (1) evaluation preparation, (2) stakeholder workshop preparation, (3) stakeholder workshop, (4) workshop analysis, and (5) evaluation conclusion.

Within the evaluation procedure, there are four issues that need to be discussed. They were pointed out throughout the process and they help to clarify some parts of the evaluation method.

The PPP evaluation method is applicable to projects that are run by not-for-profit organisations

One of the assumptions for evaluating PPPs is that profit is a relevant component to analyse, but this research provided insights for evaluating PPP not-for-profit projects. The first testing of the evaluation method was on a social housing project which was managed and built by a not-for-profit organisation. Section 7.3.1 argues why this project is considered a PPP. This test case presents a lack of one key benefits of PPPs (according to the literature): the for-profit incentive within a whole-of-life approach (bundling) (Engel et al., 2014). In this project, the profit was not a factor, and even the project budget (government subsidy) could have been overreached using NGO funds. In this social housing project, profit, cost, and project complexity are not criteria for judging Project Success. This is a stress-test for the evaluation method. After the results of the three test cases presented in chapters 7, it is possible to suggest that there is no significant difference in the evaluation of PPP projects between a for-profit and a non-for-profit developer. Even though factors such as profit and risk are not present in the not-for-profit project, the evaluation method is able to assess the functioning of the project-system with all their stakeholders.

Use of existing research for guiding the evaluation criteria

Alternative ways to develop certain components of the evaluation method were discussed throughout the research process; one of them is the use of existing research to create criteria templates for stakeholders to judge on them. Existing research has developed factors and criteria for different PPP project contexts (e.g. Hardcastle et al., 2005). They use surveys and statistical tests to validate the significance of the identified factors or criteria. A possible alternative to the proposed evaluation method would be to use the existing research and create unique forms (templates) for each stakeholder to complete. This approach might give certain benefits such as consistency between evaluated projects (for example, the community will always use the same criteria, making it easier to compare different projects). However, it also creates a flaw that is against one of the theoretical principles discussed in this study; each project is different, and each stakeholder makes a specific judgement over a project even if they need to assess several projects. For example, a corporate board can judge two projects differently even if it is the same board and the projects are similar. Ideally, the board should use a comparable standard for judging both projects, but that is not something to be enforced with a method. Conversely, the freedom for choosing the evaluation criteria can bring other issues, such as a non-recognition of some variables that could be essential for future decision-making. Nevertheless, the adopted theory is aligned with the unrestrictive generation of a stakeholder perspective. That is the reason why a fixed set of criteria based on existing literature was not employed.

An implication of adopting a free selection of criteria is the difficulty for comparing two similar PPP projects (e.g. two PPP hospitals). In this case, public entities could standardise their criteria and apply them to all their projects, making possible the comparison and post-analysis. It must be highlighted

that this comparison is for technical purposes and learning, as the real comparison (following the proposed theory) is to know whether the two project-systems worked in comparable successful ways.

A positive/neutral/negative scale was preferred instead of a Likert scale for criteria valuation

Another issue to discuss within the evaluation method is the use of a 3-scale valuation format (positive/neutral/negative) instead of a 5-option Likert scale when assessing each of the identified criteria. The reason for adopting the former option was due to the underpinned principle of the evaluation, which is the transfer of the decision-making and the accountability to the stakeholder; i.e. stakeholders should take responsibility of their judgements and provide decision-making information to the evaluation. Employing a Likert scale that includes options such as very positive, not so positive, etc., would transfer the responsibility of decision-making to the evaluator. Decision-making processes cannot be probabilistic or in middle grounds; they involve discrete options such as accept or reject an investment, continue or cancel an intervention, contribute or not contribute to the goal, etc. Therefore, a positive/neutral/negative impact spectrum of valuation is more aligned with the evaluation's core structure.

The timing for conducting the evaluation could be affected by stakeholder availability

From the researcher's experience on testing the evaluation method, the timing of the evaluation is worth discussing. According to the procedure (chapter 8), the evaluator must decide and provide support for choosing a specific time for conducting the ex-post evaluation. However, the researcher's experience on three PPP projects (test cases) could suggest that if the evaluation is conducted too far into the operational stage, important tacit knowledge could be lost from the stakeholders that were part of the initial stage of the project. As described in the stakeholder selection procedure, tackling the whole project lifecycle can be a difficult task, and this argument (stakeholder availability) could be a reason to select an evaluation time as soon as the operational stage is considered stable.

9.5. About the design process

The research was developed by employing a Design Science approach. As explained throughout the thesis, it differs from a traditional descriptive approach because it aims to develop an artefact that can solve an existing problem in the real world. This artefact is developed and described in the most scientific and rigorous way, which makes it different from a pure design process. Currently, Design Science approach is being employed mostly in information system discipline for developing IT instantiations (software applications). This study applied the existing knowledge in IT to the context of PPP projects, which involves project management, construction management, and policy evaluation.

Chapters 5, 6, 7, and 8 described the design process. Following the general guidelines for Design Science research (summarised in Figure 4-2), this study (1) is based on existing knowledge, (2) was fully documented in terms of iterations, modifications and lessons learned, (3) concluded with a clear deliverable, (4) provides a solution to the existing identified problem, and (5) also generated new knowledge to the knowledge domain in which the artefact was developed.

The existing Design Science literature suggests many technics and tools to assess the designed artefacts, but mostly in the context of software engineering outcomes (instantiations), including indicators such as effectiveness, efficiency, compatibility, etc. (Prat, Comyn-Wattiau, & Akoka, 2014). Conversely, in this study, the artefact was assessed using experts' feedback over the results obtained through the evaluation method testing. The artefact is not in the domain of information technology,

therefore, the effectiveness and efficiency indicators are difficult to quantify and measure, so they were judged relative to existing similar artefacts (e.g. policy evaluation, PMS, or audit assessment).

Hevner and Chatterjee (2010) guidelines to support a well-developed Design Science process were presented in the methodology chapter (section 4.5). There, the validity of the study in terms of Design Science standards was supported by complying to the seven items proposed by the authors.

9.6. Implications outside the scope of the study

The research generated outcomes and ideas that were not considered part of the study, they are part of a broader project management and PPP knowledge domain, and they are discussed in this section.

Applicability of evaluation method outside its original purpose

The developed PPP ex-post impact evaluation has four components that could be employed outside the primary purpose of the method: the conceptual framework PSEPM, the stakeholder workshops, the incentive-based stakeholder map, and the qualitative Value for Money approach. The Project Success Evaluation Pyramid Model (Chapter 5) could be applied to any project in which the stakeholder structure is complex and with opposing views. The stakeholder workshop (sections 8.2 and 8.3) and the subsequent analysis is a straight forward methodology for gathering arguments in any social phenomenon, such as public policy, community organisation decisions, etc. The incentive-based stakeholder map (section 8.1.1) is a tool that provides clarity to the many actors that are involved in an endeavour, considering their position in a social system. Finally, the qualitative Value for Money approach (section 7.2.5 and 8.4) for addressing the taxpayers' perspective, embodies a principle that can be adopted for PPP decision-making, comparing PPP problems and benefits.

The evaluation method could be the goal of a broader Project Success Roadmap

The research outcome is an evaluation procedure that must be applied on a specific period during the lifecycle of a project; at the end of the project or advanced in the operational stage. This outcome could be a final component of a Project Success roadmap with the goal of achieving the best possible results in the PPP ex-post impact evaluation, which is to be a successfully working project-system. This roadmap for success should understand the roles of the stakeholders, the validity of their perceptions, and pursue early accountability of the success criteria for all stakeholders; among other steps and tasks. It can be complemented with value management processes and other areas of management including performance measurement systems and stakeholder management. Additionally, the roadmap for success could describe how or why certain expectations changed throughout the project lifecycle, which could, in turn, be judged in the ex-post evaluation. This point of view was mentioned during the expert interviews and during the stakeholder workshop regarding the relevance of the ex-ante evaluation (refinement insights EC3 and TP3). However, the roadmap should consider an aligned ex-ante evaluation for getting a clearer result.

Assuming that a successfully working project-system is the best goal for a project; then, the proposed ex-post evaluation method should be the selected approach for measuring this goal. Consequently, an ex-ante evaluation should be structured as a 'prequel' of the selected ex-post evaluation approach; both evaluations should be completely aligned. Currently, ex-ante evaluations include costs, scopes, timeframes, and expected impact, however, they do not address explicitly what is necessary to achieve a successfully working project-system. Using the proposed evaluation method, the ex-ante evaluation should have a holistic view of every stakeholder. In fact, one interviewed expert mentioned that projects were usually conceptualised without any user's perspective, but then the user has great importance when assessing that same project ex-post; which is logically inconsistent.

In practice, the possibility to gather stakeholders' criteria in an ex-ante position allows the ex-post evaluation to generate accountability. Moreover, the roadmap for success could include the participation or influence of every stakeholder over every criterion, declaring who might take responsibility for meeting each criterion. Most of the criteria might be a responsibility of the project team and project sponsor, but the responsibility-allocation process can improve every stakeholder's accountability. Finally, the nature of the project must also be considered when designing a roadmap, as 'not all projects are the same'. This generalisation problem is described by some authors in the project management field (Shenhar et al., 2001, p. 704).

Single project evaluations complement the general understanding of PPPs

The proposed evaluation approach provides a clear judgement for a single PPP project; thus, it contributes to the PPP discussion with information that is not commonly presented. The discussion regarding the use of PPPs is mostly based on PPPs average indicators. The usual claim is that PPPs are better or worse because 'in general' they have X or Y common attributes. Conversely, this approach tries to gather the existing benefits and problems for a single case, confirming if the discussed attributes are present in real operating projects. A recent example of this average-based view is the NAO report (National Audit Office, 2018), that is mostly against PPP developments due to the general behaviour of existing PPPs in the UK.

PPPs should provide better and clearer information to the public audience

Finally, the author recommends that PPPs should be communicated to the public audience with the clearest information possible and throughout the entire project lifecycle. This recommendation is aligned with what interviewees suggested in the first stage of the research (Section 6.2). For example, the available information of a PPP toll road project should include (1) the structure of the consortium, (2) the structure and governance of the public sector, (3) the roles in the PPP organisation (execution, monitoring, conciliation, planning, urban planning, beneficiaries, lobbyists, etc.), and (4) the financial information taking as a reference a private bank loan. The financial information should describe clear attributes of the PPP such as investment amount, contract duration, expected revenue, payback period, the worst-case scenario for the taxpayer, the best-case scenario for the taxpayer, available alternatives, etc. All the information should be presented in the plainest possible language for transparency and accountability purposes. In Australia, this information is possible to obtain, but a bigger effort could be done for making it accessible to the nonspecialised audience.

The discussion of the issues related to the validity, utility and rigor of the proposed evaluation method should give greater support to its development and clarify the possible issues that could be considered obstacles for its use in real world projects. The next chapter will present the conclusion of the study.

CHAPTER 10

CONCLUSION

10.0. Introduction

This final chapter presents the conclusions of the study and the research process. It restates the research problem in section 10.1. It provides a synthesis of the achieved research objectives in sections 10.2 and 10.3. The implications of the study in theory and practice are described in section 10.4. Finally, the study limitations are stated, and future research is explored in sections 10.5 and 10.6, respectively.

The difference between Design Science and behavioural science is relevant to mention, as it provides greater clarity for understanding the research conclusions. This research employs a Design Science approach, which is a theoretical framework for conducting a design process to build an artefact that can solve an existing problem in the real world. The process is rigorous and documented, and the outcome is effective and applicable. This approach differs from the traditional approach of behavioural sciences, which have the aim of describing nature with maximum accuracy.

10.1. Research problem

The nature of PPPs presents many challenges when attempting to evaluate them. One of the main difficulties is the multiple stakeholders that are interrelated within a single PPP project; all of them with legitimate power and interest. The public sector is the client, and usually, it involves more than one agency (treasury, health, councils, public infrastructure, auditor, etc.). The private sector is a network of companies and entities that create a body just for developing the PPP project usually called Single Purpose Vehicle (project co, facility management, contractor, credit agencies, law firms, advisors, etc.). The direct beneficiaries of the project are at the same time taxpayers, and because PPPs provide public services, the beneficiaries do not have a voluntary connection with the project, but a more right-based understanding (hospitals that provide public health, roads that provide connectivity, etc.). As megaprojects, the urban impacts and city-nation impacts are of concern of an extended society that is also a taxpayer and voter for the authority that leads the developments of PPP projects. The system of power and interest is complex, as it is described, for example, by Michel Foucault's theory of power, 'power is everywhere' (Daldal, 2014). Moreover, PPPs are usually employed for building mega-infrastructure projects, which have inherent multiple impacts at different levels (local, regional and national impact) and timeframes (short, medium and long-term impact).

What triggers PPP evaluation is the current questioning of existing PPP projects. This model of procurement is promoted by many international agencies (World Bank, Investment banks, etc), but is also being questioned by the media, the academia, and some public agencies. Their criticism relates to (1) the amount of profit that is conferred to the private entity, (2) the long-term implications for the public sector in term of service payments, (3) the quality of the service which can be driven by cost reduction goals; among others. Additionally, a complete rejection from international organisation of work unions (PSI) (Hall, 2015), transparency issues in the UK (National Audit Office, 2018), and the bankruptcy of the private companies that provide public services, among other cases, supports the notion that PPPs should be evaluated and judged carefully.

Assuming that PPPs should be assessed, a proper assessment tool or method should be identified. From the many options for this purpose, the ex-post evaluation was recognised as the approach that could tackle in the best way the existing PPP questioning. Moreover, according to the literature, the main concern regarding PPPs is their impact on society. Therefore, the research study is narrowed to conduct an ex-post impact evaluation of PPPs.

The recognition of multiple stakeholder views towards PPPs is the point of departure for this research. This attribute is found in many big projects but is much more complex and determinant in PPP projects. Therefore, this research starts by acknowledging that the multiple stakeholder perspectives is a key driver for any attempt to develop a PPP evaluation.

The research question and the objectives of this thesis, as presented in the introduction, are:

RQ: How can ex-post impact evaluation of Public-Private Partnership projects be conducted employing the perspective of its multiple stakeholders?

- *Objective 1: Develop a conceptual framework that informs a PPP evaluation*
- *Objective 2: Develop, test and refine a method that can be used to conduct a PPP ex-post impact evaluation based on the conceptual framework*

10.2. Objective 1: Develop a conceptual framework that informs a PPP evaluation

The literature review (Chapter 3) explored the existing frameworks for evaluating a PPP project. The addressed bodies of knowledge were: research on PPP evaluation and monitoring, management-based evaluation and monitoring approaches, program evaluation (public policy) approaches, and international agencies' approaches. From those bodies of knowledge, 30 approaches were gathered. Each of them was analysed in terms of their capability for addressing multiple stakeholder perspectives. i.e. which stakeholders were considered as a formal component of the evaluation. The review resulted in that only a few evaluation approaches include explicit multiple stakeholder perspectives. Here, four approaches were identified: Performance Prism, the PPP whole life assessment, project audit body of knowledge, and the Project Success concept. The Performance Prism is a Performance Measurement System; therefore, it does not address the judgement part of an evaluation; it is an important component for gathering KPIs but not for making a judgement out of that information. The PPP whole life assessment is based on the performance prism and it is more tailored to the PPP contexts. However, its main purpose is to monitor an ongoing project, without an explicit impact assessment. The project audit body of knowledge, not in the traditional way, answered the question 'was the right thing done?' Which differ from the original 'was the thing done right?'. It considers multiple stakeholders, however always informing the main view of the auditor. This approach has the power concentrated in the auditor, to the detriment of other stakeholders' views. The Project Success concept considered multiple stakeholder perspectives, but it lacks practical development; it is still in the theoretical and academic domain with rare applications in the real world.

From the reviewed approaches, the Project Success concept was selected as the basis for developing a conceptual framework that could evaluate the impact of a PPP project. Chapter 5 describes the development of the framework called 'Project Success Evaluation Pyramid Model' (PSEPM) (Figure 10-1). It is a model that operationalises the Project Success concept by presenting a pyramid that is built from bottom to top for reaching a theoretical Project Success judgement. The pyramid shows the different level of information gathering; the first level is the pool of success criteria that are used by the stakeholders to generate their own single success judgement in the second level. The multiple judgements are analysed against the third level of the pyramid: the stakeholder judgement principles. They are defined as the principles that allow a project-system to work successfully, assuming that the project is a system that contains stakeholders (people) and the endeavour (tangible deliverables and actions). The outcome of the framework is the process for reaching a judgment of Project Success, rather than the measurement of Project Success itself.

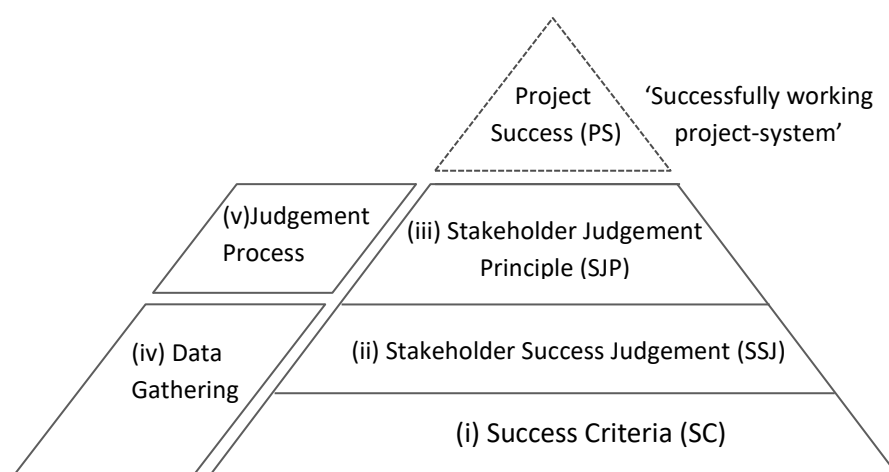


Figure 10-1: Project Success Evaluation Pyramid Model (PSEPM)

[Source: Author's]

The PSEPM was presented to PPP experts and practitioners from the private and public sector of Chile and Australia, receiving general approvals and recommendation on how to improve it. The presented version of the PSEPM has incorporated some changes to the original version presented to the experts.

The first objective was achieved using a systematic review of the possible alternatives for evaluating PPP projects and then developing a conceptual framework that can assess PPPs from multiple perspectives, which was discussed with experts in the field. The conceptual framework PSEPM is the deliverable of the first objective as indicated in the methodology chapter (Chapter 4), and it was the first iteration in the design process based on the Design Science approach.

10.3. Objective 2: Develop, test and refine a method that can be used to conduct a PPP ex-post impact evaluation based on the conceptual framework

The ex-post impact evaluation method for PPP projects employs the outcome of the first objective (the PSEPM) as a theoretical point of departure. The evaluation method was built through three design iterations of development, testing, and refinement. The design process included a first prototype version, a testing and refinement process using three PPP projects as test cases, a documentation and delivery process, and a final overall feedback from experts.

The prototype evaluation method was developed using as inputs the conceptual framework (PSEPM), the feedback from the experts regarding the PSEPM, and the prototyping exercises of the stakeholder workshops; all described in Chapter 6. The expert feedback includes 15 themes about PPP evaluation, and the prototyping exercises resulted in eight refinement insights. The prototype evaluation method is included in Appendix D.

The first testing on the social housing project gathered six stakeholders' perspectives, in total 47 success criteria were stated by the stakeholder, and eight refinement insights (lessons learned) were acknowledged from the process.

The second testing in a toll road project included 11 stakeholders' perspectives. In total 73 success criteria were gathered, and seven refinement insights were obtained. This test case provided two important additions to the evaluation method: a qualitative analysis of the Value for Money indicator and a stakeholder identification tool called incentive-based stakeholder map.

The third iteration tested the refined method in a hospital project in Australia. Eight perspectives were gathered with 53 success criteria in total. Six refinement insights (lessons learned) were developed for improving the final evaluation method.

The final procedure, considered the deliverable of the thesis, is described in Chapter 8 together with the final comments from the experts. The procedure has five main components: (1) evaluation setup, (2) stakeholder workshop preparation, (3) stakeholder workshop, (4) workshop analysis, and (5) evaluation conclusion.

The second objective was achieved by developing an ex-post impact evaluation method for PPP project based on the 'Project Success Evaluation Pyramid Model. This evaluation method was developed through three iteration processes of design, testing, and refinement. The outcome of this research objective is the evaluation procedure and documentation described in Chapter 8.

10.4. Implications in theory and practice

This section establishes the implications of the study in the existing theory and practice, highlighting the contributions to both aspects.

Contribution to the Project Success theory

This study explored the Project Success concept and challenged the existing discussion regarding the definition and application of it. It has been argued that, from the project management perspective, there are inconsistencies in the state of knowledge. The discussion around Project Success definition, including the concepts of success factors, success criteria, stakeholders' expectations and the notion of a perfect project, illustrates that the theory presents gaps and ambiguities (Chapter 5). At this point, the theoretical outcomes are three: a critical analysis of the Project Success understanding, a redefined Project Success definition, and a new Project Success framework. Even though it is difficult to find in the literature explicit definitions of Project Success, the discourse and the evolution of success can be synthesised as 'meeting stakeholders' expectations'. The refined definition 'meeting a substantial portion of stakeholders' expectation' clarifies what practice and academic discourse have noticed as the difficulty of the success concept: the different perspectives of the stakeholders and their unique judgement of what constitutes success for them. The proposed framework PSEPM is based on the understanding of stakeholders' expectations, which was interpreted as the common unit of analysis that resides in the reviewed existing success frameworks.

The relevance of the discussion around success is supported by the amount of research that has been dedicated to the topic over the last 50 years. There is a consistent effort from the research community to work towards an agreement in order to improve how projects are managed and assessed. Improving the precision of the theory impacts the industry, increasing the possibilities of assessing projects in a better way and potentially increasing the probability of success.

Stakeholder judgement principles developed using Alan Turing's test approach

Another contribution of the research is the theoretical development to judge successful projects. The Turing test approach was developed by the mathematician and philosopher Alan Turing to answer the question 'do machines think?' This research examined the Turing's approach and, on that basis, proposed a new perspective of the Project Success concept; applying the underlying reasoning of it to the question 'is this project successful?' The outcome and the contribution, in this case, is a set of three attributes that successful projects should have in common: willingness for repetition, no excessive negative impact, and no evident alternative. These principles are different from the existing frameworks related to time, cost, quality, safety, satisfaction, etc. The theoretical development also reframed the project concept as 'the implicit set of stakeholders to generate a temporal endeavour'.

This differs from the traditional project definition which is limited only to the endeavour, without taking into account the system created by the combination of stakeholders.

The first attribute of successful projects (through the lens of the stakeholders) is that projects should have a general willingness to be repeated. This means that, in general terms, the relevant stakeholders should have the disposition to hypothetically repeat the project if it was successful. The reason for this attribute is that, even when the expectations of the stakeholders are not met, the general judgement about the project can be best explained with a willingness for repetition; setting a threshold of conformity.

Second, successful projects should have a minimum requirement for judging a project as successful. The project should not generate an excessive negative outcome to one or more stakeholders, which means that none of them should suffer a large and evident negative impact over a 'fair-norm'. This is considered a 'minimum fairness' between stakeholders in any project, organisation, or society.

Finally, there is an exception that can be recognised when discussing successful projects. A project can have no willingness for repetition and may not have this minimum fairness explained; however, it might not be any evident better alternative to the ongoing project, which is a strong argument for judging that project as successful. This principle can support successful projects that have the purpose of minimising negative impact, rather than creating positive ones.

The significance of these principles can be explained with the understanding of descriptive and prescriptive knowledge. Descriptive knowledge tries to understand and describe a social phenomenon; in this case, how success behaves in the real world (through the identifications of factors and criteria). Conversely, prescriptive knowledge shapes future actions and change the behaviour of nature, creating knowledge that changes the way of thinking. This is similar to what was achieved by Porter's five forces model (Porter, 2008), which explained how strategy can be understood and managed. It is a mixing of behaviour (how things are being done) and design (how things can be done). Proposing what successful projects have or 'should' have in common not only describe projects, but also prescribe what they should pursue. This way of thinking promotes the search for other kinds of variables that can synthesise projects behaviour.

This approach can be applied to any project, from a simple and small to a large and complex one. However, its utility is more evident for projects that have a complex system of stakeholders, such as PPPs (which is currently being done by this study). Under this lens, it is possible to generate comparisons between projects and find trends that can complement the traditional iron triangle and satisfaction indicator.

Evaluating PPPs and discussing PPP policies in practice

Following the Design Science motto, the aim of this research is to provide a solution that can be applied in practice. The significance of the research should be assessed on whether the outcome is capable of solving that problem. PPP projects have been criticised and questioned, and a reasonable way to tackle this problem is by conducting a PPP ex-post impact evaluation. In the existing knowledge, there is no explicit method for this purpose, and therefore, this research developed an evaluation method that can be applied in practice, providing sufficient evidence about the rigour of the design process and about the results that can be obtained by applying it. The validity of the developed evaluation method is the way to prove the significance of the research.

Governmental agencies can employ the proposed evaluation method to assess existing projects which can be near closure, or at a stable operational stage. This method can provide a comprehensive analysis of a project and is intended for addressing the big spectrum of views of public projects. Its use can be a response to the increasing public interest on transparency and probity of public decisions.

Two specific concepts of the evaluation method must be highlighted: the incentive-based stakeholder map, and the qualitative Value for Money approach. The incentive-based stakeholder map describes in practice how to identify stakeholders in a project-system according to their roles. This map is an independent contribution to practice, as it can be applied to any kind of project, especially when they have complex stakeholder structures like PPPs. The utility of the map complements existing stakeholder management tools such as the interest/power matrix, or the rainbow diagram. The qualitative Value for Money approach is also an independent mechanism to assess Value for Money differently from the traditional quantitative approach that compares the 'whole of life net present value' between a PPP option and a Public Sector Comparator (PSC). The proposed approach presents a systematic reasoning process for addressing the benefits and problems of PPPs. It is a framework for making a case to support a PPP development beyond the financial approach, including the intangible and unquantifiable attributes that are already (informally) part of the decision-making in practice.

Currently, the discussion around PPPs is polarised. Governments and international organisations are facing divided opinions on whether PPPs should be fostered, transformed, or rejected. Moreover, the discussion is based on experiences that cannot usually be extrapolated, and conditions that cannot be changed or adapted rapidly. Megaprojects and PPPs are long-term endeavours with rigid structures and low flexibility in the middle-term. Additionally, governments have their own inertia and politics play an important role that is difficult to study and even more difficult to incorporate in any reasonably robust framework. As explained in the discussion chapter, this evaluation method provides a judgement for a single PPP project, which is rare. The PPP discussion is usually around average behaviour; i.e. PPPs are argued to be generally better or worse depending on the aspect, the definition of performance, the ideology, etc. This evaluation method contributes to the PPP discussion with new holistic and contextual specific evidence.

The proposed evaluation method should not be assessed only by its utility for providing answers, but also by the trend that is pushing towards transparency, probity, and rigour in the use of public resources. Even if the method is not followed accurately, the core concepts that support the method are also agents of change for policymaking and for complex project evaluation.

The long-term aim of the study is to foster the use of PPPs if evidence shows they are worth it. Ideally, the discussion should be based on clear and understandable principles such as the ones proposed here, but also considering everyone that deserves a stake in the discussion.

Contribution to Design Science Research (DSR) approach

The research also contributes to the use and development of the Design Science methodology in the project management and the Architecture, Engineering and Construction (AEC) disciplines. The seminal work of Design Science comes initially from Simon (1969), who presented the philosophical foundations of this kind of research. Then, Hevner et al. (2004) operationalised it with a strong focus on information system discipline, describing procedures, principles, and frameworks to conduct Design Science in practice. There are some authors that have employed DS in AEC, however, still, the body of knowledge is not robust. This study contributes to the DS application, presenting an artefact that is not an IT application, which is a common outcome for this research strategy. Chapter 4 described the design process in detail, following DS principles and providing evidence for its applicability.

Connection of the research with the industry

Finally, this research has a strong connection to the industry, mainly due to the Design Science aim of solving a real-world problem. Two of the panel members that provided supervision and advise for the development of this research are practitioners in the industry, outside the academia. This is

considered an advantage that complements the academic works, as both members provided pragmatic and ‘outcome-focused’ advice. One member is Director of PPP projects in a consulting firm, and the other is an advisor in private and governmental decision-making and value management. Both have overseen the PhD research project since the first year.

10.5. Study limitations

The limitations in terms of context, procedure, and implication are presented.

The stakeholder judgement principles should be applied as guidelines and not as rules

The first limitation is the applicability of the stakeholder judgment principles. The proposed principles (willingness for repetition, no excessive negative outcome, no evident alternative option) are presented as overarching concepts to guide the discussion that informs a success judgement. This means that they are general rules and they should be tailored for each project; many exceptions can be identified. For example, the perspective of the community towards an infrastructure that does not provide benefits at a local scale, but at a regional scale (power plant, prison, highway, etc.) would have most probably no willingness for repetition; their judgement would be probably negative, and their better alternative would be to build the project in any other location. This must be analysed differently, as the type of project gives no possibility of having a complete agreement of local communities (consider the expression ‘not in my backyard, NIMBY’). Another example is the financial result of a private entity that participates in a project. Lack of profit or bad business is a critical negative outcome for any company that enters a project. However, as explained in the analysis of Chapter 7, private entities are moved by the aim of getting profit with the intrinsic risk of losing it. If the negative impact in terms of profit cannot be assigned to the ‘project’s responsibility’ (e.g. other stakeholders’ failure or illegal practices), this impact is part of what any company could have received just for entering a business. The similar exception can be drawn for expropriations compensations. A landowner knows that the regulation that affects every citizen involves the possibility of being expropriated for the construction of public infrastructure. The lack of compensation can be considered an excessive negative outcome, but a compensation according to the current regulation should not be defined as a deal-breaker of the project, even if it is a negative impact for the landowner. These examples show the relevance of the ‘case to case’ analysis when using the stakeholder judgement principles.

The evaluation method was developed with contextual limitations (project type and country).

The design process and the outcome delivered is limited by the context in which it was developed and tested; this is considered more as a risk of application. As stated in the test case selection rationale, the PPP model is applied to a variety of project types, and in this study, only three cases were employed. Even though it was argued that the test cases represent the diversity in PPP application, they cannot fully represent the whole PPP spectrum. Theoretically, there should be project types that were not covered by the test cases in terms of their structure and operation, meaning that the evaluation method could not work as planned in every case. The same issues could happen for the country selection. The Chilean and the Australian contexts were selected for testing the evaluation method and it was argued that the use of these countries represented the existing diversity within the many countries that use the PPP model. However, it cannot be proved that the developed evaluation method will indeed work on other countries with different formal and informal rules.

The test cases contextual results cannot be extrapolated or considered official information

Another limitation is the possibility to extrapolate the results of the test cases (Chapter 7). The results obtained for each test case should not be interpreted as official, confirmatory, or conclusive. The results of each test case describe the level of information and the quality of the evaluation method at the moment of application. For example, executives in a public agency can present the general view, but they cannot be accountable for the public agency's *lei motif*. Former authorities that were considered for representing a stakeholder are not officially capable of representing the current organisation they use to be in charge, etc. Moreover, the nature of the study is for academic purposes, which means that the gathered information is not official, and it does not imply any responsibility of the participants. A formal evaluation employing this method that could be enforced contractually could present different results, as every stakeholder would be more interested in developing their perspective more thoroughly. Nevertheless, the test case results could be an input for decision-making purposes, acknowledging that the results are part of an academic exercise in which the rigour is in the design of the method and not the partial results of the testing stage.

10.6. Recommendations and future research

Through the development of the study, several recommendations can be proposed to different stakeholders for the better development of PPP projects and their evaluations.

Recommendations to the public sector

The public sector should be the main stakeholder for adopting the proposed evaluation method. Ideally, the evaluation should be conducted by an authority capable of gathering information from all the relevant stakeholders, as they need to be contacted and encouraged to participate.

From the expert's interviews, two issues are important for the public sector. First, the need for plain language information that presents the PPP with all their characteristics, benefits, risks, and problems. The public sector should pursue a transparency agenda, providing more and better information to the public, especially in complex projects such as PPPs. Second, the user must be considered *ex-ante* for planning purposes and *ex-post*. If the public sector follows the proposed evaluation method, it should include the users as a valid perspective. However, to achieve successful projects, the user perspective should be monitored from the beginning and not only at the end when a judgement is needed.

Recommendations to private companies (consortiums)

Evaluation methods such the one proposed in this thesis are helpful tools for private risk analysis that ensures economic sustainability. The private sector should have an active role in evaluating past projects, as the information can be incorporated in future endeavours. Projects have the problem of being unique every time, then, the risk is difficult to assess. More information available about users, government agencies, community groups, etc. reduces the risk profile of future projects and give insight about how problems were solved and how expectations were met. From the third test case, one of the factors that contribute more to meeting public authority's expectations in term of contract management was the company's proactive behaviour, which does not mean less profit, but the ability to tackle problems fast and effective.

Recommendations to key overseers such a treasury department

One of the main goals for a treasury department is to ensure that every project complies with the strategic government objectives (such as fiscal balance and risk exposure) and is delivered in the most effective and efficient way. Here is where the Value for Money approach proposed as part of the evaluation method can help the judgement of any project independently. To the existing cost and risk analysis of the Value for Money estimation, other non-quantifiable benefits and problems must be

addressed in an explicit way. This provides legitimacy to the estimations by accepting that the best Value for Money option for the taxpayer cannot simply be calculated using Net Present Cost (NPC) comparisons. Declaring a transparent way to deal with subjective issues should provide better support to any later audit.

Recommendations to society and taxpayers

Research on PPPs mostly agree that this procurement method must be applied correctly and to the correct situations; it is not a black and white answer whether PPPs should be promoted or avoided. Consequently, society should enforce that the correct procedures to assess this investment projects take place. To ask for information, audits, and plain language explanations should be a responsibility of the taxpayer. In fact, ways to improve the system and deliver better projects is also a responsibility of the taxpayer, recognised as a relevant stakeholder in a PPP project-system.

Finally, this study also opens a new branch for future research in two specific research areas: PPP ex-post evaluation (context), and Project Success (theory).

Further research on political and ideological issues

Evaluating PPP projects is essential for strategic decisions regarding investment agendas. This study presents an overarching model for assessing PPPs, however, it leaves areas that should be further researched. The political issues surrounding PPPs were considered in the study by addressing consistency between aim and evaluation criteria (section 8.3), and by recommending transparency and plain language for the taxpayers understanding. Nevertheless, deeper research on this topic can improve the evaluation process. Engel et al. (2014) describe the general economic and political problems of PPPs, but a robust framework to analyse and assess those problems and issues for specific political contexts is considered a gap in knowledge. The ideological bias is also a topic to further discuss, specifically, how it can affect the evaluation of a PPP. In cases like Chile, the ideological component has influenced the PPP agenda, but it has only been analysed from the media; research about political ideology and PPPs would contribute to improving PPP evaluation in the future, identifying biases and supporting judgement with better evidence.

Further research on Project Success theory and practice

To further advance the discourse in the theory of Project Success, the 'success' concept still requires further research. This study has analysed the concept, has redefined it and has proposed a framework to assess it (PSEPM). However, a deeper understanding of Project Success can contribute to a greater applicability of the concept. The theoretical development and the assumptions presented in this study could be further researched employing empirical data and quantitative methods. Surveys to confirm, for example, how the proposed judgement principles correlate with other project attributes is a next step in the research agenda. The mathematical forms of the proposed success framework could also drive more longitudinal studies in which the expectations for different project stakeholders could be tracked throughout the lifecycle of a project.

Finally, the proposed evaluation method also opens the possibility to develop a roadmap for Project Success. This could be considered the next step for improving PPP projects by ensuring they become successfully working project-systems. The roadmap should include actions and tools to manage projects in order to achieve the best possible results when evaluated ex-post with the proposed evaluation method. Using the context of PPPs, this thesis contributed to clarify what is a successful project and how it can be assessed. Consequently, now it should be easier to reach something that is clearer to identify; a successful project.

REFERENCES

- ABC News. (2015). East West Link: Taxpayers hit with \$339 million bill as Government strikes deal to scrap East West Link. Retrieved from <http://www.abc.net.au/news/2015-04-15/victorian-government-to-pay-339-million-east-west-link-contracts/6393536>
- Abedrapo, E. (2016). Concesiones de Obras Públicas en Chile 20 años. *PPP Americas 2016*. Santiago, Chile: Banco Interamericano de Desarrollo.
- Allard, P., & Nazar, M. (2019). *Costanera Norte - Relatos de la Primera Autopista Urbana de Chile*. Universidad del Desarrollo. Chile.
- Atkinson, R. (1999). Project management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. *International Journal of Project Management*, 17, 337-342. doi:10.1016/S0263-7863(98)00069-6
- Australian Government. (2016). *National Public Private Partnership Policy Framework*. Canberra, Australia: Australian Government.
- Baccarini, D. (1999). The Logical Framework Method for defining project success. *Project Management Journal*, 30(4), 25. doi:10.1177/875697289903000405
- Badewi, A. (2016). *Investigating benefits realisation process for enterprise resource planning systems*. (PhD), Cranfield University, UK.
- Balanced Scorecard Institute. (2016). Balance Scorecard Basics. Retrieved from <http://balancedscorecard.org/Resources/About-the-Balanced-Scorecard>
- Barclay, C., & Osei-Bryson, K.-M. (2010). Project performance development framework: An approach for developing performance criteria & measures for information systems (IS) projects. *International Journal of Production Economics*, 124(1), 272-292.
- Barton, R., Aibinu, A. A., & Oliveros-Romero, J. (2019). The Value for Money Concept in Investment Evaluation: Deconstructing its Meaning for Better Decision Making. *Project Management Journal*, 50(2), 210-225.
- Bauld, S., & McGuinness, K. (2006). *Value for money*. Paper presented at the Summit.
- Bing, L., Akintoye, A., Edwards, P. J., & Hardcastle, C. (2005). The allocation of risk in PPP/PFI construction projects in the UK. *International Journal of Project Management*, 23(1), 25-35.
- Bowell, T., & Kemp, G. (2014). *Critical Thinking: A Concise Guide* (4th ed.): Taylor and Francis.
- Boyer, E. J., & Scheller, D. S. (2018). An Examination of State-Level Public-Private Partnership Adoption: Analyzing Economic, Political, and Demand-Related Determinants of PPPs. *Public Works Management & Policy*, 23(1), 5-33. doi:10.1177/1087724x17729097
- Bronte-Stewart, M. m. b.-s. u. a. u. (2015). Beyond the Iron Triangle: Evaluating Aspects of Success and Failure using a Project Status Model. *Computing & Information Systems*, 19(2), 21-37.
- Burger, P., & Hawkesworth, I. (2011). How to attain value for money: Comparing PPP and traditional infrastructure public procurement. *OECD Journal on Budgeting*, 2011(1), 91-146.
- Butler, B., & Cook, H. (2014). Private Royal Children's hospital redevelopment project used taxpayer cash. *The Sydney Morning Herald*. Retrieved from [www.smh.com.au](http://www.smh.com.au/business/private-royal-childrens-hospital-redevelopment-project-used-taxpayer-cash-20140107-30ffz.html) website: <http://www.smh.com.au/business/private-royal-childrens-hospital-redevelopment-project-used-taxpayer-cash-20140107-30ffz.html>
- Carlsson, S. (2006). *Design Science research in Information Systems: A Critical Realist perspective*. Paper presented at the Proceedings of the 17th Australasian Conference on Information Systems, Adelaide, Australia. <https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsswe&AN=edsswe.oai.lup.lub.lu.se.1386253&site=eds-live&scope=site>
- Chavez, M. (2018, 19/11/2018). Hospitales concesionados han costado hasta 30% menos que los financiados por el Estado. *El Mercurio*. Retrieved from <http://www.economiaynegocios.cl/noticias/noticias.asp?id=522832>
- Copestake, J. (2014). Credible impact evaluation in complex contexts: Confirmatory and exploratory approaches. *Evaluation*, 20(4), 412-427. doi:10.1177/1356389014550559

- Craven, D. (2014). Alan Turing's legacy is even bigger than we realise. *The Conversation*. Retrieved from <https://theconversation.com/alan-turings-legacy-is-even-bigger-than-we-realise-34735>
- Daldal, A. (2014). Power and ideology in Michel Foucault and Antonio Gramsci: A comparative analysis. *Review of History and Political Science*, 2(2), 149-167.
- Davis, K. (2013). Different stakeholder groups and their perceptions of project success. *International Journal of Project Management*, 32(2), 189-201. doi:10.1016/j.ijproman.2013.02.006
- Davis, K. (2016). A method to measure success dimensions relating to individual stakeholder groups. *International Journal of Project Management*, 34(3), 480-493. doi:10.1016/j.ijproman.2015.12.009
- Davis, K. (2018). Reconciling the views of project success: A Multiple stakeholder model. *Project Management Journal*, 49(5), 38-47. doi:10.1177/8756972818786663
- de Wit, A. (1988). Measurement of project success. *International Journal of Project Management*, 6(3), 164-170. doi:10.1016/0263-7863(88)90043-9
- Debela, G. Y. (2019). Critical success factors (CSFs) of public-private partnership (PPP) road projects in Ethiopia. *International Journal of Construction Management*, 1-12. doi:10.1080/15623599.2019.1634667
- Department of Treasury and Finance. (2018, 2018). Gateway key decision points, guidance and templates. Retrieved from <https://www.dtf.vic.gov.au/gateway-review-process/gateway-key-decision-points-guidance-and-templates>
- Dewulf, G., Blanken, A., & Bult-Spiering, M. (2012). *Strategic Issues in Public-Private Partnerships* (2 ed.): Wiley.
- Doty, A. (2008). An examination of the value of the Victorian Government's investment logic map as a tool for front-end evaluation of investment proposals. *Evaluation Journal of Australasia*, 8(1), 26-39. doi:10.1177/1035719x0800800105
- Engel, E., Fischer, R., & Galetovic, A. (1997). Cómo licitar una concesión vial urbana. *Estudios Públicos*(67), 1-38.
- Engel, E., Fischer, R. D., & Galetovic, A. (2014). *The economics of public-private partnerships: A basic guide*: Cambridge University Press.
- Eskerod, P., Huemann, M., & Savage, G. (2015). Project Stakeholder Management--Past and Present. *Project Management Journal*, 6-14. doi:10.1002/pmj.21555
- European foundation for quality management. (2015). The EFQM model excellence model. Retrieved from <http://www.efqm.org/the-efqm-excellence-model>
- European PPP Expertise Centre. (2011a). Guide to guidance: How to prepare, procure and deliver PPP projects. In: European Investment Bank.
- European PPP Expertise Centre. (2011b). The non-financial benefits of PPPs: an overview of concepts and methodology. In: European Investment Bank.
- Fitzgerald, L., Brignall, S., Silvestro, R., Voss, C., & Robert, J. (1991). *Performance measurement in service businesses*: Chartered Institute of Management Accountants London.
- Frost, P. (2008). *The New Royal Women's Hospital - a public private partnership*. Melbourne, Australia: Victorian Government.
- Galles, G. (2015). *Self interest is not selfishness*. Retrieved from Alabama, USA: <https://mises.org/library/self-interest-not-selfishness>
- Gao, L., & Zhao, Z.-Y. (2020). The Evolutionary Game of Stakeholders' Coordination Mechanism of New Energy Power Construction PPP Project: A China Case. *Sustainability*, 12(3), 1045.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2016). *Impact evaluation in practice*: The World Bank.
- Gordon, J. (2016). Victoria to pay even more for East West Link cancellation. *The Age*. Retrieved from <http://www.theage.com.au/victoria/victoria-to-pay-even-more-for-east-west-link-cancellation-20160922-grm6bs.html>
- Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, 37(2), 337-355.

- Grimsey, D., & Lewis, M. K. (2005). Are Public Private Partnerships value for money?: Evaluating alternative approaches and comparing academic and practitioner views. *Accounting Forum*, 29(4), 345-378. doi:<http://dx.doi.org/10.1016/j.accfor.2005.01.001>
- Guasch, J. L. (2004). *Granting and renegotiating infrastructure concessions: Doing it right*: World Bank Publications.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, 2(163-194), 105.
- Hall, D. (2015). *Why public-private partnerships don't work*. Retrieved from University of Greenwich, UK:
- Hardcastle, C., Edwards, P. J., Akintoye, A., & Li, B. (2005). Critical success factors for PPP/PFI projects in the U.K. construction industry. *Construction Management and Economics*.
- Heald, D. (2003). Value for money tests and accounting treatment in PFI schemes. *Accounting, Auditing & Accountability Journal*, 16(3), 342-371.
- Hellowell, M., & Pollock, A. M. (2010). Do PPPs in Social Infrastructure Enhance the Public Interest? Evidence from England's National Health Service. *Australian Journal of Public Administration*, 69, S23-S34. doi:10.1111/j.1467-8500.2009.00658.x
- Hevner, A. R., & Chatterjee, S. (2010). Design Science Research Frameworks. *Design Research in Information Systems*, 23.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75-105.
- Hill, G. (2009). *A framework for Valuing the Quality of Customer information*. (Doctor of Philosophy), University of Melbourne,, Melbourne, Australia. Retrieved from <http://ghill.customer.netspace.net.au/docs/summary.html>
- Hinde, D. (2012). *PRINCE2 Study Guide*. UK: Wiley.
- HM Treasury. (2003). *The green book appraisal and evaluation in central government: Treasury guidance*: Norwich (United Kingdom): TSO.
- Hoddy, E. T. (2019). Critical realism in empirical research: employing techniques from grounded theory methodology. *International Journal of Social Research Methodology*, 22(1), 111-124. doi:10.1080/13645579.2018.1503400
- Ika, L. A. (2009). Project success as a topic in project management journals. *Project Management Journal*, 40(4), 6-19.
- Irwin, T. (2007). *Government guarantees: Allocating and valuing risk in privately financed infrastructure projects*: The World Bank.
- Jackson, E. T., & Kassam, Y. (1998). *Knowledge shared: Participatory evaluation in development cooperation*: IDRC.
- Järvinen, P. (2007). Action research is similar to design science. *Quality & Quantity*, 41(1), 37-54.
- Jepsen, A. L., & Eskerod, P. (2009). Stakeholder analysis in projects: Challenges in using current guidelines in the real world. *International Journal of Project Management*, 27(4), 335-343. doi:<http://dx.doi.org/10.1016/j.ijproman.2008.04.002>
- Jingfeng, Y., Yajun Zeng, A., Skibniewski, M. J., & Qiming, L. I. (2009). Selection of performance objectives and key performance indicators in public-private partnership projects to achieve value for money. *Construction Management & Economics*, 27(3), 253-270. doi:10.1080/01446190902748705
- Johnson, G., Scholes, K., & Whittington, R. (2008). *Exploring corporate strategy: text & cases*: Pearson Education.
- Kaplan, R. S., & Norton, D. P. (2007). Using the Balanced Scorecard as a Strategic Management System. *Harvard Business Review*, 85(7/8), 150-161.
- Khasnabis, S., Dhingra, S. L., Mishra, S., & Safi, C. (2010). Mechanisms for Transportation Infrastructure Investment in Developing Countries. 136(1), 94-103. doi:doi:10.1061/(ASCE)0733-9488(2010)136:1(94)

- KMPG. (2015). *Public private partnerships: Emerging global trends and the implication for future infrastructure development in Australia*. Retrieved from
- Koops, L., Coman, L., Bosch-Rekvelde, M., Hertogh, M., & Bakker, H. (2015). Public Perspectives on Project Success – Influenced by National Culture? *Procedia - Social and Behavioral Sciences*, 194, 115-124. doi:<http://dx.doi.org/10.1016/j.sbspro.2015.06.126>
- Langbein, L. I., & Felbinger, C. L. (2006). *Public program evaluation: A statistical guide*: Armonk, N.Y. : M.E. Sharpe, c2006.
- Lim, C. S., & Mohamed, M. Z. (1999). Criteria of project success: an exploratory re-examination. *International Journal of Project Management*, 17, 243-248. doi:10.1016/S0263-7863(98)00040-4
- Liu, A., & Walker, A. (1998). Evaluation of project outcomes. *Construction management and economics*, 16(2), 209-219. doi:10.1080/014461998372493
- Liu, J., Love, P. E., Davis, P. R., Smith, J., & Regan, M. (2013). *Performance measurement framework in PPP projects*. Paper presented at the P3Book Proceedings of International Conference on PPP Body of Knowledge in Preston, UK, University of Central Lancashire, Lancashire.
- Liu, J., Love, P. E. D., Davis, P. R., Smith, J., & Regan, M. (2015). Conceptual framework for the performance measurement of public-private partnerships. *Journal of Infrastructure Systems*, 21(1). doi:10.1061/(ASCE)IS.1943-555X.0000210
- Liu, J., Love, P. E. D., Smith, J., Regan, M., & Palaneeswaran, E. (2015). Review of performance measurement: implications for public–private partnerships. *Built Environment Project and Asset Management*, 5(1), 35-51. doi:doi:10.1108/BEPAM-12-2013-0070
- Love, P. E., Skitmore, M., & Earl, G. (1998). Selecting a suitable procurement method for a building project. *Construction Management & Economics*, 16(2), 221-233.
- Lund-Thomsen, P. (2007). *Assessing the impact of public-private partnerships in the Global South*. Retrieved from Switzerland:
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251-266. doi:[http://dx.doi.org/10.1016/0167-9236\(94\)00041-2](http://dx.doi.org/10.1016/0167-9236(94)00041-2)
- McElroy, B., & Mills, C. (2000). Managing stakeholders. *Gower handbook of project management*, 757-775.
- McKee, M., Edwards, N., & Atun, R. (2006). Public-private partnerships for hospitals. *Bulletin of the World Health Organization*, 84(11), 890-896.
- McLeod, L., Doolin, B., & MacDonell, S. G. (2012). A perspective-based understanding of project success. *Project Management Journal*, 43(5), 68-86.
- Milton-Barker, A. (2016). Mitsuku chatbot wins Loebner Prize for most humanlike A.I., yet again. Retrieved from <https://venturebeat.com/2016/10/15/mitsuku-chatbot-wins-most-humanlike-a-i-loebner-prize-again/>
- Ministerio de Desarrollo Social. (2017). Evaluacion social ex post. Retrieved from <http://sni.ministeriodesarrollosocial.gob.cl/evaluacion-iniciativas-de-inversion/evaluacion-ex-post/>
- Ministerio de Obras Publicas. (2016). Coordinacion de concesiones de obras publicas. Retrieved from http://www.concesiones.cl/quienes_somos/Paginas/default.aspx
- MINVU. (2013). Ministerio de Vivienda y Urbanismo. Retrieved from <http://www.minvu.cl>
- Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of management review*, 22(4), 853-886.
- Mladenovic, G., Vajdic, N., Wündsche, B., & Temeljotov-Salaj, A. (2013). Use of key performance indicators for PPP transport projects to meet stakeholders' performance objectives. *Built Environment Project & Asset Management*, 3(2), 228.
- Müller, R., & Turner, J. R. (2007). The influence of project managers on project success criteria and project success by type of project. *European management journal*, 25(4), 298-309.

- Müller, R. A., & Jugdev, K. A. (2006). A retrospective look at our evolving understanding of project success. *IEEE Engineering Management Review*, 110.
- Nalewaik, A., & Mills, A. (2015). Project Performance Audit: Enhanced Protocols for Triple Bottom Line Results. *Procedia - Social and Behavioral Sciences*, 194, 134-145. doi:<http://dx.doi.org/10.1016/j.sbspro.2015.06.185>
- Nalewaik, A., & Mills, A. (2017). *Project performance review: Capturing the value of audit, oversight, and compliance for project success*: London : Routledge, 2017.
- National Audit Office. (2013). *Savings from operational PFI contracts*. UK: National Audit Office.
- National Audit Office. (2018). *PFI and PF2*. UK: National Audit Office.
- Neely, A., Adams, C., & Crowe, P. (2001). The performance prism in practice. *Measuring Business Excellence*, 5(2), 6-13.
- Neely, A., Gregory, M., & Platts, K. (1995). Performance measurement system design: a literature review and research agenda, 80.
- Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M., & Kennerley, M. (2000). Performance measurement system design: developing and testing a process-based approach. *International Journal of Operations & Production Management*, 20(9/10), 1119-1145.
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. In (Vol. 16).
- OECD. (2016). OECD recommendations on principles for public governance of public private partnership. Retrieved from <http://www.oecd.org/governance/oecd-recommendation-public-privatepartnerships.htm>
- Oliveros-Romero, J., & Aibinu, A. A. (2019). Ex post impact evaluation of PPP projects: an exploratory research. *Built Environment Project and Asset Management*, 9(2), 315-330.
- Oliveros R., J., & Aibinu, A. A. (2018). *Ex-post impact evaluation for PPP projects using multiple stakeholder perspectives*. Paper presented at the European Evaluation Society (EES) biennial Conference 2018, Thessaloniki, Greece.
- Oliveros R., J., & Gonzalez-Mathiesen, C. (2017). *Real estate social housing market employing a PPP approach: the Chilean case*. Paper presented at the State of Australian Cities Conference (SOAC 2017), Adelaide, Australia.
- Olsen, R. P. (1971). Can project management be defined? *Project Management Quarterly*.
- Osei-Kyei, R., & Chan, A. P. C. (2017). Implementing public-private partnership (PPP) policy for public construction projects in Ghana: critical success factors and policy implications. *International Journal of Construction Management*, 17(2), 113-123. doi:10.1080/15623599.2016.1207865
- Osterwalder, A. (2004). The business model ontology: a proposition in a design science approach, l'Université de Lausanne, 2004. On-line: http://www.hec.unil.ch/aosterwa/PhD/Osterwalder_PhD_BM_Ontology.pdf.
- Owen, J. M., & Alkin, M. C. (2007). *Program evaluation: Forms and approaches* (3rd ed.): New York : Guilford Press, c2007.
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, 64(1), 12-40.
- Peffer, K. E. N., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of management information systems*, 24(3), 45-77.
- Pinto, J. K., & Slevin, D. P. (1988). Project success: definitions and measurement techniques. *Project Management Journal*, 19(1), 68-72.
- Pollock, A. (2012). How PFI is crippling the NHS. *The Guardian*. Retrieved from <http://www.theguardian.com/commentisfree/2012/jun/29/pfi-crippling-nhs>
- Porter, M. E. (2008). The five competitive forces that shape strategy. *Harvard Business Review*, 86(1), 78-93.
- Prat, N., Comyn-Wattiau, I., & Akoka, J. (2014). *Artifact Evaluation in Information Systems Design-Science Research-a Holistic View*. Paper presented at the PACIS.

- Project Management Institute. (2013). *A guide to the project management body of knowledge (PMBOK® guide)* (Fifth edition ed.). Newtown Square, Pennsylvania: Project Management Institute.
- Public-Private-Partnership in Infrastructure Resource Center. (2016). Government Objectives: Benefits and Risks of PPPs. Retrieved from <http://ppp.worldbank.org/public-private-partnership/overview/ppp-objectives>
- Quade, E. S., & Carter, G. M. (1989). *Analysis for public decisions* (G. M. Carter. Ed. 3rd ed.): New York : North-Holland, c1989.
- Raisbeck, P., Duffield, C., & Xu, M. (2010). Comparative performance of PPPs and traditional procurement in Australia. *Construction Management and Economics*, 28(4), 345-359. doi:10.1080/01446190903582731
- Raisbeck, P., Duffield, C., & Xu, M. (2010). Comparative performance of PPPs and traditional procurement in Australia. *Construction management and economics*, 28(4), 345-359. doi:10.1080/01446190903582731
- Rapoport, R. N. (1970). Three Dilemmas in Action Research:With Special Reference to the Tavistock Experience. *Human Relations*, 23(6), 499-513. doi:10.1177/001872677002300601
- Rawls, J. (1999). A Theory of Justice, revised edition. Cambridge, MA: Harvard University Press.
- Renger, R. (2015). System evaluation theory (SET): A practical framework for evaluators to meet the challenges of system evaluation. *Evaluation Journal of Australasia*, 15(4), 16.
- Ringer, F. K. (2004). *Max Weber: an intellectual biography*: University of Chicago Press.
- Rockart, J. F. (1980). The changing role of the information systems executive : a critical success factors perspective. In.
- Rogers, D. (2013). Funding the future after the demise of PPPs. *The Conversation*. Retrieved from theconversation.com website: <http://theconversation.com/funding-the-future-after-the-demise-of-ppps-18869>
- Saaty, R. W. (1987). The analytic hierarchy process—what it is and how it is used. *Mathematical Modelling*, 9(3), 161-176. doi:[https://doi.org/10.1016/0270-0255\(87\)90473-8](https://doi.org/10.1016/0270-0255(87)90473-8)
- Saint. (2016). Legal Systems of the World. Retrieved from <http://saint-claire.org/wp-content/uploads/2016/01/Legal-Systems-of-the-World.pdf>
- Sandoval, G., & Villegas, D. (2014). Hospitales concesionados: Minsal plantea que modelo sube 77,9 costo de construcción. *La Tercera*. Retrieved from <http://www.latercera.com/noticia/hospitales-concesionados-minsal-plantea-que-modelo-sub-779-costo-de-construccion/>
- Sepulveda, R. (2016). Las brutales ganancias de las rutas concesionadas: se podrían hacer más de 90 carreteras. Retrieved from <http://www.biobiochile.cl/especial/noticias/reportajes/reportajes-economia/2016/10/07/las-brutales-ganancias-de-las-rutas-concesionadas-se-podrian-hacer-mas-de-90-carreteras.shtml>
- Shenhar, A. J., Dvir, D., Levy, O., & Maltz, A. C. (2001). Project success: A multidimensional strategic concept. *Long Range Planning*, 34(6), 699-725. doi:10.1016/S0024-6301(01)00097-8
- Simon, H. A. (1969). The sciences of the artificial. Cambridge, MA.
- Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.): MIT press.
- State Government of Victoria. (2008). *Investment lifecycle guidelines, post-implementation review*. Melbourne, Australia: The secretary department of Treasury and Finance.
- State Government of Victoria. (2016). Public Private Partnership projects. Retrieved from <http://www.dtf.vic.gov.au/Infrastructure-Delivery/Public-private-partnerships/Projects>
- State Government of Victoria. (2017). *Partnerships Victoria, Excellence in public private partnerships*. Melbourne, Australia: State of Victoria
- Stevenson, A. (2011). *Oxford dictionary of English. [electronic resource]* (3rd ed. ed.): Oxford University Press.

- Tarazona Vento, A. (2017). Mega-project meltdown: Post-politics, neoliberal urban regeneration and Valencia's fiscal crisis. *54*(1), 68-84. doi:10.1177/0042098015625025
- The Economist Intelligence Unit. (2014). *Evaluando el entorno para las asociaciones publico-privadas en America Latina*. Retrieved from www.fomin.org
- Toomela, A. (2010). Quantitative methods in psychology: inevitable and useless. *Frontiers in psychology*, *1*, 29.
- Toor, S.-u.-R., & Ogunlana, S. O. (2010). Beyond the 'iron triangle': Stakeholder perception of key performance indicators (KPIs) for large-scale public sector development projects. *International Journal of Project Management*, *28*(3), 228-236. doi:<http://dx.doi.org/10.1016/j.ijproman.2009.05.005>
- Turner, J. R. (2009). *The handbook of project-based management: leading strategic change in organizations* (3rd ed.). New York McGraw-Hill.
- Turner, J. R., & Zolin, R. (2012). Forecasting success on large projects: Developing reliable scales to predict multiple perspectives by multiple stakeholders over multiple time frames. *Project Management Journal*, *43*(5), 87-99.
- Valenzuela, N. (2009). Corrupción en las EGIS: ¿Se derrumba el modelo chileno de Vivienda Social? . *Plataforma Urbana*. Retrieved from [plataformaurbana.cl](http://www.plataformaurbana.cl) website: <http://www.plataformaurbana.cl/archive/2009/07/15/corrupcion-en-las-egis-%C2%BFse-derrumba-el-modelo-chileno-de-vivienda-social/>
- van der Loop, H., Haaijer, R., & Willigers, J. (2016). New Findings in the Netherlands about Induced Demand and the Benefits of New Road Infrastructure. *Transportation Research Procedia*, *13*, 72-80. doi:<https://doi.org/10.1016/j.trpro.2016.05.008>
- Vassallo, J. M. (2006). Traffic risk mitigation in highway concession projects: the experience of Chile. *Journal of Transport Economics and Policy (JTEP)*, *40*(3), 359-381.
- Vedoya, S., & Fernández, O. (2018, 27/10/2018). Justicia estudia concesionar nuevas cárceles. *La Tercera*. Retrieved from <https://www.latercera.com/nacional/noticia/justicia-estudia-concesionar-nuevas-carceles/377973/>
- Venable, J., Pries-Heje, J., & Baskerville, R. (2012). A comprehensive framework for evaluation in design science research. In K. Peffers, M. Rothenberger, & B. Kuechler (Eds.), *Design Science Research in Information Systems. Advances in Theory and Practice: 7th International Conference, DESRIST 2012, Las Vegas, NV, USA, May 14-15, 2012. Proceedings* (pp. 423-438). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Victorian Auditor-General's Office. (2017). *Results of 2016–17 Audits: Public Hospitals*. Retrieved from Melbourne:
- Waite, M. (2012). *Paperback Oxford English dictionary*. Oxford, UK: Oxford University Press.
- Wateridge, J. (1998). How can IS/IT projects be measured for success? *International Journal of Project Management*, *16*(1), 59-63.
- Westerveld, E. (2003). The Project Excellence Model®: linking success criteria and critical success factors. *International Journal of Project Management*, *21*(6), 411-418. doi:[http://dx.doi.org/10.1016/S0263-7863\(02\)00112-6](http://dx.doi.org/10.1016/S0263-7863(02)00112-6)
- White, H. (2010). A Contribution to Current Debates in Impact Evaluation. *Evaluation*, *16*(2), 153-164. doi:10.1177/1356389010361562
- Williams, T. (2016). Identifying Success Factors in Construction Projects: A Case Study. *Project Management Journal*, *47*(1), 97-112. doi:10.1002/pmj.21558
- World Bank. (2012). *Strategic environmental assessment in the World Bank*. Retrieved from Washington DC:
- World Bank. (2015). *H1 2015 global PPI update*. Retrieved from
- World Bank. (2016). Beneficiary assessment. Retrieved from <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTSOCIALDEVELOPMENT/EXTPCEN/G/0,,contentMDK:20507670~pagePK:148956~piPK:216618~theSitePK:410306,00.html>
- World Bank. (2019). *H1 2019 Private participation in infrastructure (PPI)*. Retrieved from

- World Bank, Asian Development Bank, & Inter-American Development Bank. (2014). *Public private partnerships - Reference guide 2.0*. Retrieved from Washington, US:
- World Bank national accounts data, a. O. N. A. d. f. (2020). GDP per capita (current US\$) - Australia, Chile. Retrieved 20/04/2020, from The world Bank <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=AU-CL>
- World Economic Forum. (2013). *Strategic infrastructure steps to prepare and accelerate public-private partnerships*. Retrieved from
- Xiong, W., Yuan, J.-F., Li, Q., & Skibniewski, M. J. (2015). Performance objective-based dynamic adjustment model to balance the stakeholders' satisfaction in PPP projects. *Journal of Civil Engineering & Management*, 21(5), 539-547. doi:10.3846/13923730.2014.895409
- Yang, R. J., & Shen, G. Q. P. (2015). Framework for stakeholder management in construction projects. *Journal of Management in Engineering*(4), 4014064. doi:10.1061/(ASCE)ME.1943-5479.0000285
- Yin, R. K. (2013). Validity and generalization in future case study evaluations. *Evaluation*, 19(3), 321-332. doi:10.1177/1356389013497081
- Zimring, C. M., & Reizenstein, J. E. (1980). Post-occupancy evaluation: An overview. *Environment and Behavior*, 12(4), 429-450.

APPENDIX A: ETHICS APPROVAL, CONSENT FORM AND PLAIN LANGUAGE DECLARATIONS

09 December 2016

Dr A.A. Aibinu

Architecture, Building and Planning
The University of Melbourne

Dear Dr Aibinu

I am pleased to advise that the Architecture, Building and Planning Human Ethics Advisory Group has approved the following Minimal Risk Project.

Project title: **Evaluation tool to assess the impact of social infrastructure PPP projects on stakeholders**
Researchers: **J Oliveros Romero, Dr A A Aibinu**
Ethics ID: **1648165**

The Project has been approved for the period: **06-Dec-2016 to 31-Dec-2017**.

It is your responsibility to ensure that all people associated with the Project are made aware of what has actually been approved.

Research projects are normally approved to 31 December of the year of approval. Projects may be renewed yearly for up to a total of five years upon receipt of a satisfactory annual report. If a project is to continue beyond five years a new application will normally need to be submitted.

Please note that the following conditions apply to your approval. Failure to abide by these conditions may result in suspension or discontinuation of approval and/or disciplinary action.

(a) **Limit of Approval:** Approval is limited strictly to the research as submitted in your Project application.

(b) **Amendments to Project:** Any subsequent variations or modifications you might wish to make to the Project must be notified formally to the Human Ethics Advisory Group for further consideration and approval before the revised Project can commence. If the Human Ethics Advisory Group considers that the proposed amendments are significant, you may be required to submit a new application for approval of the revised Project.

(c) **Incidents or adverse effects:** Researchers must report immediately to the Advisory Group and the relevant Sub-Committee anything which might affect the ethical acceptance of the protocol including adverse effects on participants or unforeseen events that might affect continued ethical acceptability of the Project. Failure to do so may result in suspension or cancellation of approval.

(d) **Monitoring:** All projects are subject to monitoring at any time by the Human Research Ethics Committee.

(e) **Annual Report:** Please be aware that the Human Research Ethics Committee requires that researchers submit an annual report on each of their projects at the end of the year, or at the conclusion of a project if it continues for less than this time. Failure to submit an annual report will mean that ethics approval will lapse.

(f) **Auditing:** All projects may be subject to audit by members of the Sub-Committee.

Please quote the ethics registration number and the name of the Project in any future correspondence.

On behalf of the Ethics Committee I wish you well in your research.

Yours sincerely



Dr Robert Crawford
Acting Chair
Architecture, Building and Planning Human Ethics Advisory Group



msd
Melbourne
School of Design

FACULTY OF
ARCHITECTURE,
BUILDING AND
PLANNING
www.msd.unimelb.edu.au

Consent Form (experts)

Melbourne School of Design, faculty of Architecture, Building and Planning



Project: Evaluation tool to assess the impact of social infrastructure PPP projects on stakeholders

Primary Researcher: Ajibade Aibinu

Additional Researchers: Jose Oliveros – Student Researcher

Name of Participant: _____

1. I consent to participate in this project, the details of which have been explained to me, and I have been provided with a written plain language statement to keep.
2. I consent to / my name being used in any publications arising from this research / be referred to by a pseudonym. (~~strikethrough one option~~)
3. I understand that the purpose of this research is to develop an evaluation tool to assess the impacts that a social infrastructure PPP project has on the different stakeholders.
4. I understand that my participation in this project is for research purposes only.
5. I acknowledge that the possible effects of participating in this research project have been explained to my satisfaction.
6. In this project I will be required to give my opinion regarding the applicability and the context of the proposed impact evaluation for PPP projects
7. I understand that my interviews may be audio-taped.
8. I understand that my participation is voluntary and that I am free to withdraw from this project anytime without explanation or prejudice and to withdraw any unprocessed data that I have provided.
9. I understand that the data from this research will be stored at the University of Melbourne and will be destroyed after 5 years.
10. I have been informed that the confidentiality of the information I provide will be safeguarded subject to any legal requirements; my data will be password protected and accessible only by the named researchers.
11. I understand that given the small number of participants involved in the study, it may not be possible to guarantee my anonymity.
12. I understand that after I sign and return this consent form, it will be retained by the researcher.

Participant Signature: _____ **Date:** _____

Consent Form (stakeholders)

Melbourne School of Design, faculty of Architecture, Building and Planning



Project: Evaluation tool to assess the impact of social infrastructure PPP projects on stakeholders

Primary Researcher: Ajibade Aibinu

Additional Researchers: Jose Oliveros – Student Researcher

Name of Participant: _____

1. I consent to participate in this project, the details of which have been explained to me, and I have been provided with a written plain language statement to keep.
2. I consent to / my name being used in any publications arising from this research / be referred to by a pseudonym. *(strikethrough one option)*
3. I understand that the purpose of this research is to develop an evaluation tool to assess the impacts that a social infrastructure PPP project has on the different stakeholders.
4. I understand that my participation in this project is for research purposes only.
5. I acknowledge that the possible effects of participating in this research project have been explained to my satisfaction.
6. In this research I will be required to give my opinion in terms of expectations and perceptions to determine the impact that the project has on me as a relevant stakeholder representative.
7. I understand that my interviews may be audio-taped.
8. I understand that my participation is voluntary and that I am free to withdraw from this project anytime without explanation or prejudice and to withdraw any unprocessed data that I have provided.
9. I understand that the data from this research will be stored at the University of Melbourne and will be destroyed after 5 years.
10. I have been informed that the confidentiality of the information I provide will be safeguarded subject to any legal requirements; my data will be password protected and accessible only by the named researchers.
11. I understand that given the small number of participants involved in the study, it may not be possible to guarantee my anonymity.
12. I understand that after I sign and return this consent form, it will be retained by the researcher.

Participant Signature: _____ **Date:** _____

HREC Number: 1648165.1 Project Start Date: 26/11/2015 Version: v1

PLAIN LANGUAGE STATEMENT (Experts)

First Interview: "PPP Assessment Tool" Evaluation

Dr Ajibade Aibinu (Supervisor)

Faculty of Architecture, Building and Planning
University of Melbourne
Ph: +61 3 8344 6811
Email: aaibinu@unimelb.edu.au

Mr. Jose Oliveros (PhD Student)

Faculty of Architecture, Building and Planning
University of Melbourne
Ph: +61 4 5251 4129
Email: joseo@student.unimelb.edu.au



Project: "Evaluation tool to assess the impact of Social Infrastructure PPP on stakeholders"

Introduction

You are invited to participate in the above research project, which is being conducted by Dr. Ajibade Aibinu (**Supervisor**) and Mr. Jose Oliveros (**PhD Student**) of the Faculty of Architecture, Building and Planning at the University of Melbourne. The aim of the study is to develop an evaluation tool to assess the impacts that a social infrastructure PPP project has on the different stakeholders.

We are interested in talking to you about a new way of evaluating a Public-Private Partnership (PPP) Project presented by the researcher.

What will I be asked to do?

As an expert in the field, if you agree to participate, you would be asked to contribute in the following ways:

You will be presented with a newly developed way (tool) of assessing a PPP after it has been implemented. You will be provided with an explanation of the core concepts that supports the tool. After the presentation, you will be asked to provide feedback regarding the tool highlighting the validity of the tool concept, the usefulness, the applicability in practice, as well as what should be included or removed and how it could be improved. The total time of interaction will be no more than one hour.

With your permission, this interview would be digitally recorded so that we can make an accurate record of what you say. When the recording has been transcribed, you would be provided with a copy of the transcript, so that you can verify that the information is correct and/or request deletions.

How will my confidentiality be protected?

This project has been approved by the Human Research Ethics Committee and we intend to protect your anonymity and the confidentiality of your responses to the fullest possible extent, within the limits of the law. Your name and contact details will be kept in a separate, password-protected computer file from any data that you supply. This will only be able to be linked to your responses by the researchers, for example, in order to know where we should send your interview transcript for checking. In the final report, if you agree, you will be referred to by a pseudonym. We will remove any references to personal information that might allow someone to guess your identity; however, you should note that as the number of people we seek to interview is very small, it is possible that someone may still be able to identify you. The data will be kept securely in the Faculty of Architecture, Building and Planning for five years from the date of publication, and may be destroyed after this time. It is also possible that the data you provide may be used by the researchers for related research in the future.

Date: 11 October 2016 / HREC Number: 1648165 Version 1 /

How will I receive feedback?

Once the thesis arising from this research has been completed, a brief summary of the findings will be available to you on application at the Faculty of Architecture, Building and Planning. It is also possible that the results will be presented at academic conferences and in published papers.

Do I have to take part?

Participation in this study is voluntary. Should you wish to withdraw at any stage, or to withdraw any unprocessed data you have supplied, you are free to do so without prejudice. Your decision to participate or not, or to withdraw, will not affect any relationship you may have with this department or University and will not affect any services you may receive now or in the future.

Where can I get further information?

Please contact the researchers if you have any questions or if you would like more information about the project please do not hesitate to contact the researchers at the contact details above.

If you have any concerns about the conduct of the project which you do not wish to discuss with the research team please contact the Manager, Human Research Ethics, Office for Research Ethics and Integrity, The University of Melbourne, on ph: 8344 2073, or email address: humanethics-complaints@unimelb.edu.au.

How do I agree to participate?

If you would like to participate, please indicate that you have read and understood this information by signing the consent form.

APPENDIX B: EVALUATION FORMS TEMPLATES

Ex - post Evaluation tool						FORM A - Stakeholder Perspective									
Project						Stakeholder									
Description of the project						Purpose of the stakeholder towards the project									
Success Criteria	Relevance	Rank	Expectation	Supporting Documentation	Actual Value	Supporting Documentation	Reason for the difference exp/perc.	Supporting Documentation	Impact	Lessons Learned					
Willingness for repetitions						No excessive negative Outcome (fairness)					No evident alternative option Supporting/Challenging Criteria				
Supporting/Challenging Criteria															

Ex-post evaluation FORM B: EXAMPLES OF MOTIVES, CRITERIA AND SUPPORTING DOCUMENTATION

Stakeholder	Motives/Mission	Type	Success Criteria	Supporting Documentation
Private Sector	Existence	Financial	Profit	Project Objectives
	Profit		Net Present Value	Contract
	Prestige		revenue	Business Case
	relationship with society		whole life cost	Investment Logic Map
Public Authority	Learning Curve		Capital cost	Work Breakdown Structure
	Compliance to a higher Authority		Payback	Stakeholder Agreements
	transparency		Risk analysis and reponse	Stakeholder Negotiations
	Political Revenue		Time	Push Communication issues
	Consistency with strategy		Business Model	Audits and review reports
	Efficiency		contract Modification	Formative evaluation reports
User	Achievement of Particular Objectives	Management	economic growth	Satisfaction surveys
	End user/client satisfaction		Productivity ratios	internal estimation
	internal improvement		maintenance	executives estimation
	fulfil a specific need		complains	minimum logical requirement
Community/ Society	Quality in product/service		Rate of absences	historical statistics
	good experience		HHRR effectiveness	
	Positive social Impacts	Process	Accident rates	
	Positive economic impacts		health safety, and environmental impact	
	Positive environmental impacts		Conflict resolution	
	transparency		participation	
	fairness		Timely completion of project	
			objective achievement	
		Satisfaction/ Perception	Adequate management	
			Compliance	
			safety	
			intervention during construction	
		social	compensations	
			Client/user satisfaction	
			relationship with a stakeholder	
			employee satisfaction	
			Connectivity	
			Urban impact	
			social impact	
			Public policy purpose	

Ex-post evaluation

FORM C: TAXPAYER PERSPECTIVE

Qualitative Value for Money Assessment

PPP possible Benefit	PPP possible Problems	possible TP avoided problems
Efficiency	Excessive Renegotiations	Deficient project selection
Operational Efficiency due to bundling	Loss of public budget future flexibility (due to future service fee transfers)	Low quality maintenance
On time	High (out of the norm) private profits	Inefficient service-fee price adjustment
On budget	Rigidity in service provision	Regulation problems and corruption
Integration of design, construction, financing, operational, maintenance, and refurbishment responsibilities	Risk of non-well defined functional specifications	Incentive for developing new infrastructure to the detriment of investing in infrastructure maintenance
Improvement of construction quality for long-term use.	Higher tendering costs	Limited risk appetite due to the political impact
Lower estimated lifecycle cost compared to PSC	need for additional skills in negotiation, tendering and management	
	Stakeholder rejections due to private involvement (ideological reasons).	
	Higher return rates over capital investment	
Finance/Economy		
Better risk allocation		
Impact on the economy (due to PPP method)		
Funding availability		
Off-sheet balance treatment		
Economically feasible projects		
Design and fit-to-purpose risk on private sector		
The potential for lower cost of assets		
Availability of funding for using the PSC budget on a different project		
Urgency (due to funding availability)		
Payment after commissioning		
Quality		
Incentive for high-quality maintenance		
High level of innovation		
Clear performance standards		
Transparency		
Incentive for availability		
Better residual infrastructure after transfer to public		

for taxpayers, the following should stand...

$$\begin{array}{c} \text{Value of} \\ \text{PPP extra} \\ \text{cost} \end{array} < \begin{array}{c} \text{Value of} \\ \text{PPP} \\ \text{benefits} \end{array} - \begin{array}{c} \text{Value of} \\ \text{PPP} \\ \text{problems} \end{array} + \begin{array}{c} \text{Value of} \\ \text{Avoiding TP} \\ \text{problems} \end{array}$$

APPENDIX C: PRESENTATION EMPLOYED IN INTERVIEWS

msd
Melbourne School of Design

A new ex-post Impact Evaluation tool for assessing social infrastructure PPP projects

José Oliveros Romero, PhD student

Supervisor
Dr. Ajibade Aibinu

December 2016

msd
Melbourne School of Design

Research Summary

Research Problem The lack of a method to evaluate the impact of a PPP project, specifically a infrastructure PPP project.

Research Outcome Designed ex-post evaluation tool for social Infrastructure PPP projects

Research Aim (1) Improve future PPP developments and (2) provide arguments for the discussion about the use of PPPs

msd
Melbourne School of Design

Underlying Assumptions

- The complexity of PPPs is due to (1) the mix of stakeholders interests and power, (2) the multiple perceptions of the same outcomes, (3) the time frame, and (4) the political and ideological frame (public interest)
- The fulfilment of all the possible stakeholder expectations is a theoretical perfect result of a PPP
- Every stakeholder evaluates a PPP formally or informally; it seems better to make it formal
- Project impacts are different from PPP impacts

msd
Melbourne School of Design

Existing Pattern in Evaluation Methods

msd
Melbourne School of Design

New Approach for PPP Evaluation

Conceptual Framework

msd
Melbourne School of Design

Evaluation example

Stakeholder

	Private	Public	User	Society
Criteria	profit	users	Price	CO2 emissions
Expectation	10M	2K	Free	200/MM
Perception	8M	4K	Paid	800/MM
Reason	Fix service Price	Quality of service	Public means free	No standard adopted
Repetition (why)	Yes	Yes	No	Yes
Fairness (why)	Yes	Yes	Yes	No
Best option (why)	Yes	Yes	No	No

Impact Evaluation tool flow chart

GAP Grid

PPP criteria

MCDM tool

- AHP
- Interval AHP or
- Cloud Theory

Stakeholder judgements

Judgement

Repetition Project - PPP

Fairness Project - PPP

Best Option Project - PPP

Principle X Project - PPP

Impact Evaluation tool flow chart

Risk's Significance		What impacts	What impacts
Reason for the assessment	Assessment		
	Positive result		
	Failure result		
Objectives of the assessment	Assessment Results	What will we do it	How is going to be done
	Start		
	Source		
	Monitors		
	Validity		
	Assignment of gaps		
	Assessors' Interest		
	Assignment Principles	Significance	
	a		
	b		
	c		
	d		
	e		

	Price	Costs	Results
End of contract			
Contract Renewal			
Start of Service. Contract after PPP Others			

Category	Stakeholder	Description	Positive impacts from the project	Negative impacts from the project	Bias/Interests	Included in the assessment (Y/N)	Why / Why not
Private							
Public							
End User							
Society							

Selected Stakeholders Unselected Stakeholders

Why the assessment is not better including this stakeholder? Why the assessment is not better excluding this stakeholder?

	Motive 1	Motive 2	Motive 1	Motive 2	Motive 1	Motive 2
Stakeholder 1	Motive 1					
	Motive 2					
Stakeholder 2	Motive 1					
	Motive 2					
Stakeholder 3	Motive 1					
	Motive 2					

[illegible]

LAP PHASES			
Gap Code	00-1	Area	N/A
Stakeholder	Private Industry	Source	Contract
Category	Transportation	Medium	Profit
Criteria	Total cost		
Project Stage	Construction		
Breakdown unit	Infrastructure (not service)		
Fill in by the stakeholder			
Qualitative		Quantitative	
Metric	Total cost per project	Reason of the gap	
Units	millions USD	How comes that deal arises from bidding	
Best Possible Value	5 MM	Implication of the gap	
Most Possible Value	15 MM	Why is it a disadvantage and modifications	
Expected Value	8 MM	Why benefits from the gap	
Perceived Value	12 MM	Why losses from the gap	
Perceived score of achievement	75/100	The gap in the government by having a cheaper bidding process	
Relevance of the gap		The contribution to time and cost	
Ranking of the gap		First	
Fill in by the evaluator			
Cost-benefit ratio		Cost-benefit ratio	
Cost-benefit between a gap and its reasons		Effect of the gap	
Conflicts with other gaps		Level of influence	
Modifications		Reasons	
Comments		Comments	

Stakeholders	Contractor	Stakeholder_2	Stakeholder_3	Stakeholder_4	Stakeholder_5
Regulation Support	Yes	Answer	Answer	Answer	Answer
	Gap 1.5%	Gap 1.5%	Gap 1.5%	Gap 1.5%	Gap 1.5%
	Gap 2.5%	Gap 2.5%	Gap 2.5%	Gap 2.5%	Gap 2.5%
	Profit 20%	Gap 3.5%	Gap 3.5%	Gap 3.5%	Gap 3.5%
	Gap 4.5%	Gap 4.5%	Gap 4.5%	Gap 4.5%	Gap 4.5%
Fairness Support	Answer	Answer	Answer	Answer	Answer
	Gap 1.5%	Gap 1.5%	Gap 1.5%	Gap 1.5%	Gap 1.5%
	Gap 2.5%	Gap 2.5%	Gap 2.5%	Gap 2.5%	Gap 2.5%
	Gap 3.5%	Gap 3.5%	Gap 3.5%	Gap 3.5%	Gap 3.5%
	Gap 4.5%	Gap 4.5%	Gap 4.5%	Gap 4.5%	Gap 4.5%
Best option Support	Answer	Answer	Answer	Answer	Answer
	Gap 1.5%	Gap 1.5%	Gap 1.5%	Gap 1.5%	Gap 1.5%
	Gap 2.5%	Gap 2.5%	Gap 2.5%	Gap 2.5%	Gap 2.5%
	Gap 3.5%	Gap 3.5%	Gap 3.5%	Gap 3.5%	Gap 3.5%
	Gap 4.5%	Gap 4.5%	Gap 4.5%	Gap 4.5%	Gap 4.5%
Other Support	Answer	Answer	Answer	Answer	Answer
	Gap 1.5%	Gap 1.5%	Gap 1.5%	Gap 1.5%	Gap 1.5%
	Gap 2.5%	Gap 2.5%	Gap 2.5%	Gap 2.5%	Gap 2.5%
	Gap 3.5%	Gap 3.5%	Gap 3.5%	Gap 3.5%	Gap 3.5%
	Gap 4.5%	Gap 4.5%	Gap 4.5%	Gap 4.5%	Gap 4.5%

[illegible]

Criteria	Profit	Price	Gap 3	Gap 4
Profit	1	2	2	1
Price	0.5	1	3	2
Gap 3	0.5	0.333333	1	2
Gap 4	1	0.5	0.5	1
sum	3	3.333333	6.5	6

Pairwise Comparison

Criteria	Profit	Price	Gap 3	weight
Profit	0.333333	0.521739	0.307692	0.332558
Price	0.166667	0.260870	0.461538	0.305662
Gap 3	0.166667	0.080695	0.153846	0.185201
Gap 4	0.333333	0.320495	0.769231	0.768839

$$Total\ Impact\ Score = \sum_1^i GAP_i WEIGHT_i$$

$$PPP\ Impact\ Score = \sum_{i=1}^I PPPGAP_i WEIGHT_i$$

$$PPP\ influence = \frac{PPP\ impact\ score}{total\ impact\ score}$$

PPP influence
score

Gap Grid

Stakeholders
judgements

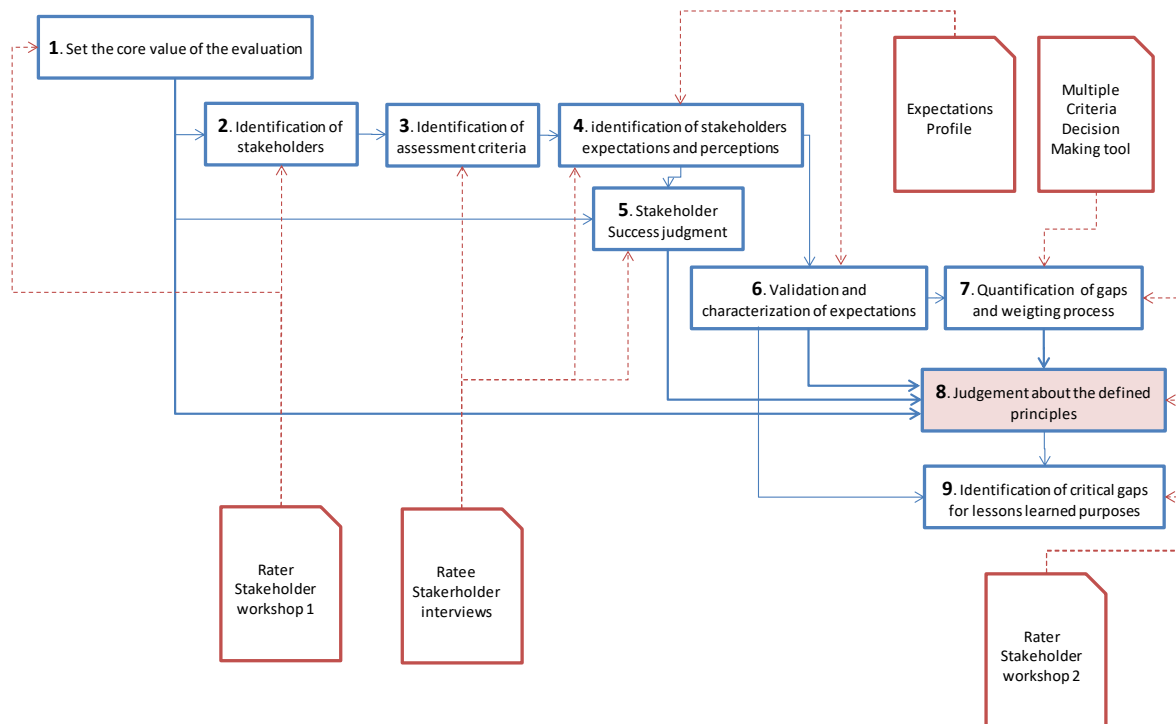
PPP criteria

Judgment Principles	Principal	Comment	POP	Comment
Repetition	Statement		Statement	
Reason 1				
Reason 2				
Reason 3				
Fairness	Statement		Statement	
Reason 1				
Reason 2				
Reason 3				
Recid option	Statement		Statement	
Reason 1				
Reason 2				
Reason 3				
Other	Statement		Statement	
Reason 1				
Reason 2				
Reason 3				

Principles	Critical gap	Comment/support	Project	FRF	Future recommendation
Abjection					
Fairness					
Best Option					
Other					

- **Conceptual Development:** is there any problem in the theoretical development for evaluating the impact of PPP projects?
- **Evaluation Tool:** Can the tool be improved? How?
- **Evaluation tool Applicability:** How to make the tool more applicable to real projects
- **Tool Limitations:** What are the limitations of the tool?
- **Incentives:** What are the incentives to apply something like the proposed tool?
- **Results:** What can be done with the results obtained with the tool?

APPENDIX D: PROTOTYPE EVALUATION METHOD PROCEDURE



Step 1: set the requirement for the assessment process

Establish (1) the objective of the assessment, (2) the consequences of the assessments regarding the possible results, and (3) the final user of the evaluation report.

- Why the assessment should be performed?
 - What are the objectives of the assessment?
 - Who benefits from the performing of the assessment?
- What are the implications of a successful result and to whom?
- What are the implications of a failure result and to whom?
- Who will use the results of the assessment and how?

Define the time for the evaluation application. It should be coherent with the objectives of the evaluation

Define the rater Stakeholder who will be responsible of assessing the perceptions of all the other stakeholders involved in the project

Define the Stakeholder Judgement Principles for the final outcome. Establish the principles that the rater stakeholder will use at the end of the evaluation.

This step should be performed by the rater stakeholder or the entity that can be triggering the evaluation.

Reason for the assessment	Who impacts		what impacts
	Assessment		
	Success result		
	Failure result		
Objectives of the assessment	Assessment Results		
	Who will use it		How is going to be used
	Gaps		
	Scores		
	Motives		
	validity		
	judgment of gaps		
	lessons learned		

Figure: Prototype of the rationale section in the tool interface

Source: Author's

	Pros	Cons	Result
End of contract			
Concession Renewal			
Start of Service Contract after PPP			
Others			

Figure: Possible milestones from assessment application

Source: Author's

Step 2: identification of the stakeholders

Select and evaluate which stakeholders will be part of the evaluation. Assess possible conflict of interest, positive and negative impacts from the project, and any possible bias (ideological, political, religious, etc.).

Category	Stakeholder	Description	Positive impacts from the project	Negative Impacts from the project	Bias/Interests	Included in the assessment(Y/N)	Why / Why not
Private							
Public							
End User							
Society							

Figure: Stakeholder identification and selection

Source: Author's

For the selected stakeholders, test the validity of the selection by applying:

- Marginal value analysis: Compare the value of taking out the last selected (less relevant) stakeholder and the value of including the most relevant non-selected stakeholder.
- Cross-motive Analysis: Compare the motives of each stakeholder against all the other motives to check any conflict between the stakeholders.

This step is performed by the rater stakeholder.

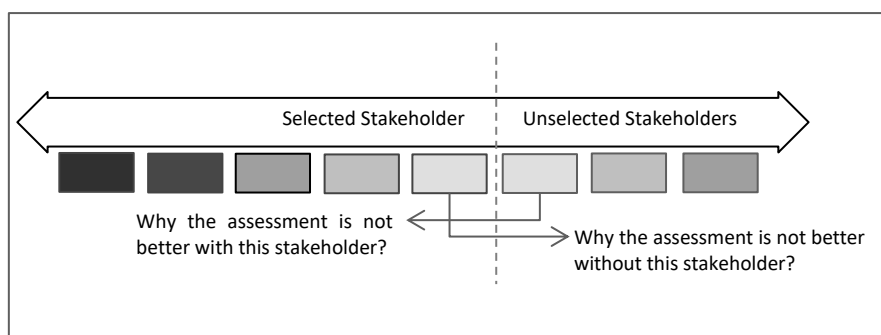


Figure: Marginal Value Analysis

Source: Author's

		Stakeholder 1		Stakeholder 2		Stakeholder 3	
		Motive 1	Motive 2	Motive 1	Motive 2	Motive 1	Motive 2
Stakeholder 1	Motive 1						
	Motive 2						
Stakeholder 2	Motive 1						
	Motive 2						
Stakeholder 3	Motive 1						
	Motive 2						

Figure: Cross Motive Analysis' table

Source: Author's

Step 3: identification of assessment criteria

Select the assessment criteria by combining (1) the motives (interest) of the stakeholders, and (2) the sources of criteria. The motives of the stakeholders are separated in a generic way as private goals, public goal, end user goals, and society goals. The sources of criteria are the original PPP contract, stakeholder issues and disputes, media discussion, etc.

Each selected criteria should be linked with at least one stakeholder and the total list of criteria should be refined with each stakeholder.

This step is performed by the third party or a rater stakeholder and reviewed by the other stakeholders in step 4.

Stakeholder Goals/Motives		Sources of Criteria
Public Sector Public Welfare End User Approval Positive Value for Money Positive Political Impact Compliance No unforeseen negative impacts On budget +	Private Entity Profit Reputation Environmental compliance Long Term Alliance for future PPP Controlled Risks Safety on site No unforeseen negative impacts Transparent Public Process Positive Learning Curve +	Project Objectives Contract Business Case Investment Logic Map Work Breakdown Structure Stakeholder Agreements Stakeholder Negotiations Push Communication issues Audits and review reports Formative evaluation reports Satisfaction surveys
	End User Quality Product/service On time Long Term usability +	
	Society Positive Social Impact Positive Economic Impact Positive Environmental Impact Not excessive private profit No unforeseen negative impacts Transparent process Positive economic/environmental trade off Positive social/environmental trade off +	

Figure: Preliminary stakeholders' goals/motives and sources of criteria

Source: Author's

Stakeholder Goals/Motives	Success Criteria	Sources of Criteria
Private Entity		
Profit	NPV	Contract
Society	Gains and Losses	Business Case
No Excessive private Profit		Financial report

Figure: Example of the establishment of assessment criteria

Source: Author's

Step 4: collection of stakeholder expectations and perceptions

Each stakeholder should develop an expectations portfolio. This consists in addressing all the gaps of expectations from a single stakeholder. Each gap of expectation will be addressed in a gap profile that contains:

- Quantitative information: Metrics, scales, targets, expected value, perceived value, relative relevance of the gap.
- Qualitative information: Reason for the gap, implication of the gap, who benefits from the gap, who losses from the gap.

GAP PROFILE	
Gap Code	Area
Stakeholder	Source
Category	Motive
Criteria	
Project Stage	
breakdown unit	
Quantitative	Qualitative
Metric	Reason of the gap
Units	
Best Possible Value	Implication of the gap
Worst Possible Value	
	Who benefits from the gap
Expected Value	
Perceived Value	Who loses from the gap
Perceived Score of achievement /100	
Relevance of the gap	
Ranking of the gap	

Figure: Prototype of gap profile

Source: Author's

Step 5: stakeholders' success judgement

Each Stakeholder should answer the threshold questions selected in step 1. The partial impression of each answer will be part of the input that the rater stakeholder needs, in order to make the final judgement. The bias of the rater stakeholder is partially controlled by the transparency of the process in which all the different perspectives need to be acknowledged and then validated. This is more robust compare to an evaluation that only needs to acknowledge the arguments that support a decision.

Step 6: validation and characterisation of expectations

The expectations of all the stakeholders addressed in the gap profile should allow a feasible version of the project. This validation process needs to be performed by the rater stakeholder.

The analysis consists in comparing each expectations of each stakeholder against all the other expectations. Each comparison should be performed in terms of what was expected, how it was expected, when it was expected.

			Stakeholder 1						Stakeholder 2						Stakeholder 3					
			Exp 1			Exp2			Exp 1			Exp2			Exp 1			Exp2		
			What	How	When	What	How	When	What	How	When	What	How	When	What	How	When	What	How	When
Stakeholder 1	Exp 1	What																		
		How																		
		When																		
Stakeholder 2	Exp 2	What																		
		How																		
		When																		
Stakeholder 3	Exp 1	What																		
		How																		
		When																		
Stakeholder 3	Exp 2	What																		
		How																		
		When																		

Figure: Prototype of cross-gap map
Source: Author's

Any conflict, between the expectations, that makes the project theoretically unfeasible needs to be solved and corrected by the rater stakeholder. The conflicts should be solved using three steps.

- 1) Decide which expectation is more relevant to the project (by ranking the stakeholders)
- 2) Solve the conflict by reducing or eliminating the less important expectation
- 3) Provide recorded and public reasons for the decision that was taken

A separate analysis needs to be done, to make a subset of criteria that is directly related to the decision of adopting the PPP strategy. Criteria that are dependent of the PPP strategy or were a cause of the PPP should be part of the subset.

Step 7: quantification and characterisation of expectations

To inform the judgement of the rater stakeholder, a quantitative overall project success score is necessary to develop using a Multiple Criteria Decision Making (MCDM) tool. The process could use a pair wise comparison to obtain the relative weights of each gap. Then a regular AHP equation can obtain the final project success score. More comprehensive approaches such as interval AHP or cloud theory can also be employed, depending on the complexity of the project.

Two score will be obtained, one for the Project success and another for the PPP success. The latter consist on weighting just the criteria of the PPP subset to get a quantitative overall score.

	More important than								equal	Less important than								
Gap 1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Gap 1
Gap 2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Gap 2
Gap 3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Gap 3
Gap 4	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Gap 4

Criteria	Gap 1	Gap 2	Gap 3	Gap 4
Gap 1				
Gap 2				
Gap 3				
Gap 4				

Figure: Pair wise comparison for the gaps in expectations
Source: Author's

Step 8: judgement about the defined principle

The final judgement is performed by the rater stakeholder. This judgement is based on four inputs: (1) The stakeholder partial judgement of the principles, (2) the gap grid containing all the quantitative and qualitative information about the gaps in expectations, and (3) the quantitative output obtained

from the MCDM tool, and (4) the identification of the criteria that are affected by the PPP procurement strategy. With the four pieces of information the rater stakeholder will need to provide a judgement for each stakeholder judgement principle.

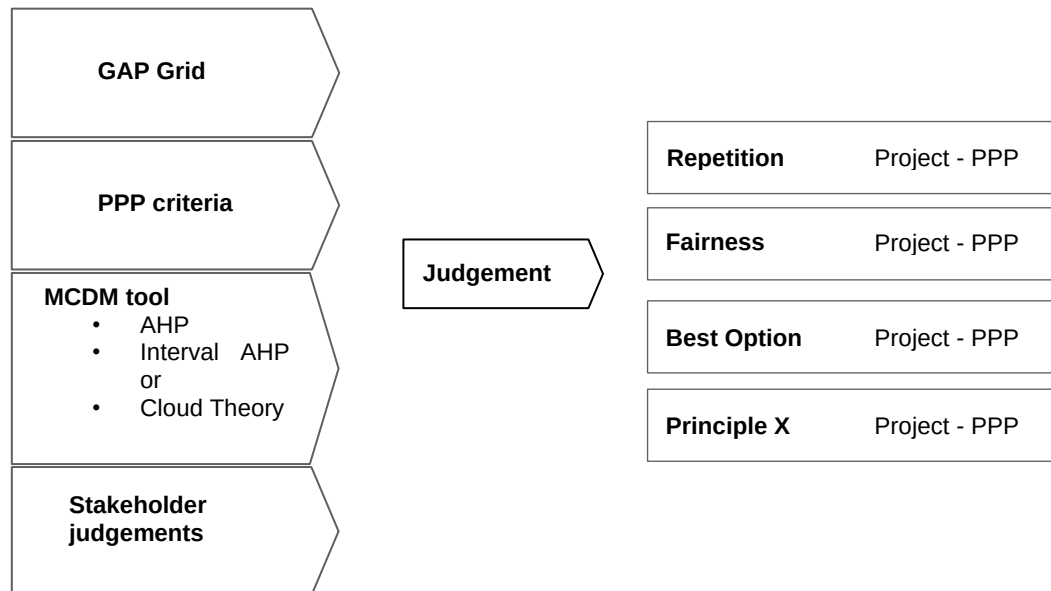


Figure: inputs and outputs of the final judgement
Source: Author's

Step 9: identification of critical gaps for lesson learned purpose

From the gap profile, the most critical gaps (relevant to inform successfulness) should be addressed to show explicitly the lessons learned. Also, it needs to be provided to whom these gaps are important.