

**School of Management & Marketing**

**An integrated procurement assessment model for the  
measurement of development effectiveness and value for money  
in international development**

Bryan Cook

18964039

This thesis is presented for the Degree of  
Doctorate of Philosophy  
of  
Curtin University

August 2022

## **DECLARATION**

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

**Human Ethics** (For projects involving human participants/tissue, etc....) The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2014. The proposed research study received human research ethics approval from the Curtin University Human Research Ethics Committee (EC00262), Approval Number # HRE2018-0649.

Signature:

Date: 25 August 2022

## **DEDICATIONS**

This thesis is dedicated to my family: Michelle, Ellie and Michael Cook who have always supported me during this period of study.

## ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisors, Dr. Elizabeth Jackson and Dr. Paul Alexander, who guided me throughout this research.

Thank you to Dr. Jenny Lalor who provided guidance and support on the statistical analysis.

*“A little knowledge is a dangerous thing”*. Often considered to be a misquote from a verse by Pope (1716) for the quote: *“A little learning is a dangerous thing”*(Pope, 1716)

Implies that a small amount of knowledge, or learning, can lead people to believe they know more than they actually do. The quote is included in this research to support the notion of continuous pursuit of knowledge for transparency, truth and clarity in research.

## **ABSTRACT**

International development organisations disburse over US\$100 billion of Official Development Assistance globally per annum; it is estimated that 50% of this is procured in the form of goods and services in development projects. This substantial investment in development is controversial and has experienced enduring examination and debate over the effectiveness and impact of aid disbursement. This research examines the relationship between project success and procurement performance to identify predictors that contribute to Development Effectiveness (DE) and Value for Money (VFM) in international development. The literature review examines existing procurement assessment models and current project evaluation methodology used to measure procurement and project performance by leading international development organisations. Data were collected in two key phases of the research. The first was an evaluation of procurement process maturity through feedback from ten international development and three private sector organisations by an Integrated Procurement Maturity Model (IPMM). This was followed by a statistical analysis of 1,920 project evaluation reports from the same development organisations to examine comparative project performance results and impact of projects on DE and VFM. Results showed that, despite the similarity between the procurement and project management processes, there is a very weak statistical correlation between the procurement performance and the project evaluation results. The findings indicate that the existing project evaluation criteria, used by the selected international development organisations, have inconsistencies and do not provide a reliable measurement of DE and VFM. Following the analysis of the project evaluation criteria and the development of the IPMM model, several recommendations for improvement are presented. The research contributes to theory on the measurement of DE and VFM in international development projects.

## TABLE OF CONTENTS

DECLARATION .....	I
DEDICATIONS.....	II
ACKNOWLEDGEMENTS .....	III
ABSTRACT .....	IV
TABLES .....	VII
FIGURES .....	IX
ABBREVIATIONS .....	X
1 INTRODUCTION .....	1
1.1 A brief description of international development.....	1
1.2 Overview of the Research.....	6
2 OVERVIEW OF INTERNATIONAL DEVELOPMENT ORGANISATIONS.....	8
2.1 Introduction .....	8
2.2 Multilateral Development Banks.....	11
2.3 United Nations agencies .....	13
2.4 Bilateral development organisations .....	15
2.5 Summary.....	17
3 LITERATURE REVIEW .....	18
3.1 Introduction .....	18
3.2 The role of procurement in international development .....	22
3.3 Defining Development Effectiveness in international development .....	28
3.4 Defining Value for Money in international development.....	32
3.5 Private sector definitions and approaches to project effectiveness.....	38
3.6 Criticism of development effectiveness; Where does the money go?.....	43
3.7 Corruption and waste.....	44
3.8 Commitments towards improvements in Development Effectiveness .....	53
3.9 Measurement of development effectiveness.....	59
3.10 Evaluation of development performance.....	62
3.11 Measurement of procurement performance.....	68
3.12 Procurement maturity assessment models .....	72
3.13 Preliminary theoretical framework from the literature .....	83
3.14 Research Questions .....	86
3.15 Research Aims and Objectives .....	87
3.16 Conclusion.....	88
4 RESEARCH DESIGN .....	90
4.1 Introduction .....	90
4.2 Methodological steps in the research.....	93
4.2.1 Research philosophy.....	93
4.2.2 Research approach.....	96
4.2.3 Research methods .....	97
4.2.4 Time horizons .....	101
4.3 Data sources and descriptions.....	101
4.3.1 Selection of the international development organisations .....	103
4.3.2 Selection of the procurement expert panel .....	107
4.3.3 Secondary data collection .....	110
4.3.4 Primary data collection .....	115
4.3.5 IPMM questionnaire design.....	124
4.3.6 VFM factors and weighted IPMM questionnaire .....	129
4.4 Administration of the questionnaire .....	133
4.5 Data analysis.....	135
4.5.1 Normalisation of the different ratings scales .....	137
4.6 Statistical testing and analysis .....	140
4.7 Research validity and integrity .....	146

4.8	Presentation of data .....	147
4.9	Approach to data analysis in this research.....	147
5	RESULTS .....	150
5.1	Procurement maturity results.....	150
5.1.1	Introduction .....	150
5.1.2	Integrated procurement maturity model questionnaire results .....	151
5.1.3	Statistical distribution for IPMM results .....	156
5.1.4	Weighted IPMM questionnaire results .....	160
5.1.5	Statistical distribution for the weighted IPMM results.....	165
5.1.6	Comparison between original and weighted IPMM scores .....	168
5.1.7	Comparison of the three types of international organisations .....	170
5.1.8	Comparison of IPMM results; private sector and development organisations ....	172
5.1.9	Comparison of procurement groups in development organisations .....	174
5.1.10	Analysis of the four procurement phases.....	177
5.1.11	Analysis of the 12 procurement criteria.....	179
5.2	Project Evaluation Results.....	181
5.2.1	Introduction .....	182
5.2.2	Overall project evaluation results by the total mean score .....	186
5.2.3	Analysis of project evaluation by type of organisation .....	192
5.2.4	Analysis of project evaluation results by region.....	195
5.2.5	Statistical distribution of project evaluation results by region .....	198
5.2.6	Analysis of project evaluation results by sector .....	201
5.2.7	Statistical distribution of project evaluation results by sector .....	203
5.2.8	Project evaluation ratings by region and sector.....	206
5.2.9	Project evaluation criteria ratings .....	209
5.2.10	Cost estimation and actual costs .....	213
5.3	Correlation of the project evaluation mean results .....	223
5.4	Regression analysis of the project evaluation mean results.....	227
5.5	Summary.....	230
6	DISCUSSION .....	233
6.1	Introduction .....	233
6.2	Procurement capacity and maturity assessment .....	237
6.3	Project evaluation performance assessment methodology .....	244
6.4	Relationship between procurement maturity and project performance .....	254
6.5	Measurement of development effectiveness and value for money .....	257
6.6	Summary.....	261
7	CONCLUSION.....	264
7.1	Introduction .....	264
7.2	Concluding remarks.....	267
7.3	Limitations of the research .....	274
7.4	Contributions to theory.....	277
7.5	Contributions to practice .....	280
7.6	Recommendation for further research .....	282
7.7	Summary of the chapter.....	285
	REFERENCES .....	289
	APPENDIX 1. IPMM QUESTIONNAIRE.....	347
	APPENDIX 2. QUESTIONS FOR PROCUREMENT EXPERT PANEL ON VFM .....	362
	APPENDIX 3. COMPARISON OF TRANSFORMATION EQUATIONS FOR SCALE NORMALISATION ...	365

## TABLES

TABLE 1. MULTILATERAL DEVELOPMENT BANKS.....	12
TABLE 2. SUB-REGIONAL MULTILATERAL DEVELOPMENT BANKS .....	13
TABLE 3. UNITED NATIONS AGENCIES .....	14
TABLE 4. UNITED NATIONS SPECIALISED AGENCIES.....	15
TABLE 5. UNITED NATIONS OTHER ENTITIES AND RELATED ORGANISATIONS.....	15
TABLE 6. PROCUREMENT EXPENDITURE OF MULTILATERAL DEVELOPMENT BANKS .....	25
TABLE 7. OPERATIONAL PROCUREMENT EXPENDITURE OF UN AGENCIES .....	26
TABLE 8. SUPPLIERS TO UNITED NATIONS AGENCIES BY COUNTRY AND REGION .....	27
TABLE 9. EVALUATION REPORTS WORLD BANK AND ASIAN DEVELOPMENT BANK .....	65
TABLE 10. INDEPENDENT EVALUATION GROUP RATING SCALE.....	65
TABLE 11. PROCUREMENT ASSESSMENT TOOLS USED IN DEVELOPMENT ORGANISATIONS	77
TABLE 12. ARTICLES ON PROCUREMENT MATURITY ASSESSMENT MODELS .....	78
TABLE 13. COMMON PROCUREMENT MATURITY ASSESSMENT MODELS .....	79
TABLE 14. COMPARISON OF RESEARCH PHILOSOPHIES IN MANAGEMENT RESEARCH .....	95
TABLE 15. COMPARISON OF DEDUCTIVE, INDUCTIVE AND ABDUCTIVE APPROACHES .....	97
TABLE 16. JUSTIFICATION OF COMBINING QUALITATIVE AND QUANTITATIVE RESEARCH ..	99
TABLE 17. PARAMETERS FOR THE SELECTION OF THE DEVELOPMENT ORGANISATIONS..	105
TABLE 18. EXPERT PROCUREMENT PANEL PARTICIPANTS.....	110
TABLE 19. DELPHI SURVEY DESIGN TYPES .....	117
TABLE 20. REFERENCES FOR KEY PROCUREMENT MATURITY CRITERIA .....	127
TABLE 21. VALUE FOR MONEY FACTORS .....	131
TABLE 22. RATINGS SCALES (SIX-POINT AND FOUR-POINT SCALES) .....	138
TABLE 23. PROJECT EVALUATION NORMALISED RATINGS .....	139
TABLE 24. SUMMARY OF RELATIONSHIPS, DIFFERENCES AND TRENDS BY DATA TYPE.....	143
TABLE 25. OVERALL IPMM QUESTIONNAIRE TOTAL SCORES.....	153
TABLE 26. STATISTICAL DISTRIBUTION DATA FOR IPMM RESULTS.....	157
TABLE 27. OVERALL WEIGHTED IPMM RESULTS.....	162
TABLE 28. STATISTICAL DISTRIBUTION DATA FOR WEIGHTED IPMM RESULTS .....	166
TABLE 29. GAP ANALYSIS BETWEEN ORIGINAL AND WEIGHTED IPMM RESULTS .....	168
TABLE 30. FREQUENCY DISTRIBUTION FOR IPMM RESULTS BY ORGANISATION TYPE.....	171
TABLE 31. IPMM RESULTS FOR PRIVATE SECTOR (OIL AND GAS) ORGANISATIONS .....	173
TABLE 32. FREQUENCY DISTRIBUTION FOR IPMM RESULTS BY PROCUREMENT GROUP ..	175
TABLE 33. FREQUENCY DISTRIBUTION FOR IPMM RESULTS BY PROCUREMENT PHASE...	177
TABLE 34. FREQUENCY DISTRIBUTION IPMM RESULTS BY PROCUREMENT CRITERIA .....	180
TABLE 35. TOTAL PERCENTAGE DISTRIBUTION OF PROJECT EVALUATION RATINGS.....	184
TABLE 36. PROJECT EVALUATION RATINGS TOTAL MEAN RESULTS .....	187
TABLE 37. FREQUENCY DISTRIBUTION DATA FOR PROJECT EVALUATION RESULTS.....	188
TABLE 38. STATISTICAL DISTRIBUTION PROJECT EVALUATION BY ORGANISATION TYPE.	193
TABLE 39. BONFERRONI STATISTICAL TEST PROJECT EVALUATION BY ORGANISATION ..	194
TABLE 40. TOTAL MEAN RESULTS FOR PROJECT EVALUATION RATINGS BY REGION .....	196
TABLE 41. PROJECT EVALUATION RESULTS AND CORRUPTION INDEX BY REGION.....	198
TABLE 42. CORRELATION OF PROJECT EVALUATION AND CORRUPTION BY REGION.....	200
TABLE 43. TOTAL MEAN RESULTS FOR PROJECT EVALUATION RATINGS BY SECTOR .....	202
TABLE 44. STATISTICAL DISTRIBUTION OF PROJECT EVALUATION RESULTS BY SECTOR ..	204



TABLE 45. PROJECT EVALUATION RATINGS BY REGION AND SECTOR .....	206
TABLE 46. PROJECT EVALUATION CRITERIA RATINGS.....	211
TABLE 47. PERCENTAGE VARIATION OF ACTUAL COSTS FROM COST ESTIMATE .....	214
TABLE 48. STATISTICAL DISTRIBUTION FOR ACTUAL COSTS FROM COST ESTIMATE.....	217
TABLE 49. CAUSES OF COST OVERRUN AND UNDERRUN IN ADB PROJECTS.....	221
TABLE 50. SUMMARY OF PROCUREMENT AND PROJECT EVALUATION RESULTS .....	224
TABLE 51. CORRELATION OF PROJECT EVALUATION RESULTS AND COST DATA.....	225
TABLE 52. REGRESSION ANALYSIS FOR OVERALL PROJECT EVALUATION AND CRITERIA .	228

## FIGURES

FIGURE 1. OFFICIAL DONOR ASSISTANCE EXPENDITURE BY MAJOR DONORS IN 2018.....	16
FIGURE 2. HIERARCHY OF PERFORMANCE MEASUREMENT DEVELOPMENT EFFECTIVENESS	21
FIGURE 3. VALUE FOR MONEY DEFINED BY THE ASIAN DEVELOPMENT BANK .....	35
FIGURE 4. PROCUREMENT MATURITY LEVELS .....	76
FIGURE 5. PROCUREMENT IPMM PROCESS DIAGRAM .....	82
FIGURE 6. THEORETICAL FRAMEWORK SUGGESTED BY THE LITERATURE.....	84
FIGURE 7. FLOW DIAGRAM OF THE RESEARCH PROCESS .....	93
FIGURE 8. THE RESEARCH ONION .....	94
FIGURE 9. DEPARTMENT FOR INTERNATIONAL DEVELOPMENT RATING SCALE .....	138
FIGURE 10. ANALYSIS OF RESULTS .....	148
FIGURE 11. DATA ANALYSIS .....	149
FIGURE 12. OVERALL IPMM RESULTS .....	154
FIGURE 13. OVERALL WEIGHTED IPMM RESULTS.....	163
FIGURE 14. TOTAL PERCENTAGE DISTRIBUTION OF PROJECT EVALUATION RATINGS .....	185
FIGURE 15. PROJECT EVALUATION RATINGS TOTAL MEAN RESULTS .....	189
FIGURE 16. TOTAL MEAN PROJECT EVALUATION RESULTS BY REGION .....	196
FIGURE 17. PROJECT EVALUATION MEAN RESULTS BY SECTOR .....	203
FIGURE 18. TOTAL PROJECT EVALUATION RATINGS BY REGION AND SECTOR .....	206
FIGURE 19. TOTAL PROJECT EVALUATION CRITERIA RATINGS .....	212
FIGURE 20. PERCENTAGE COST VARIATION BY THRESHOLD.....	214

## ABBREVIATIONS

ADB	Asian Development Bank
AFDB	African Development Bank
AIIB	Asia Infrastructure Investment Bank
ARDE	Annual Review of Development Effectiveness
APA	Alternative Procurement Arrangements
APQC	American Productivity and Quality Center
BOAD	West African Development Bank
BSTDB	Black Sea Trade and Development Bank
BSC	Balanced Scorecards
CAF	Development Bank of Latin America
CDB	Caribbean Development Bank
CABEI	Central American Bank for Economic Integration
CTBTO	The Preparatory Commission for Comprehensive Nuclear Test Ban Treaty Organisation
CSR	Corporate Social Responsibility
CPI	Corruption Perception Index
CIPS	Chartered Institute of Procurement and Supply
CPE	Country Programme Evaluation
CCPE	Cluster Country Programme Evaluation
CLRR	Validation of Completion Learning Reviews
CAPE	Country Assistance Programme Evaluations
CPSFR	Country Partnership Strategy Final Review
CMMI	Capability Maturity Model Integration
CEB	Corporate Executive Board
DE	Development Effectiveness
DEC	Development Experience Clearinghouse, USAID
DFAT	Department of Foreign Affairs and Trade, Australia
DFID	Department for International Development, United Kingdom
ECOSOC	The Economic and Social Council
EBRD	European Bank for Reconstruction and Development
EIB	European Investment Bank
EADB	East African Development Bank
ETDB	Economic Co-operation Trade and Development Bank
EDB	Eurasian Development Bank
EvD	Evaluation Department
EIRR	Economic International Rate of Return
ECG	Evaluation Co-operation Group
EIPM	European Institute of Purchasing Management
EFQM	European Foundation for Quality Management
EAP	East Asia Pacific
ECA	Europe Central Asia
FAO	Food and Agriculture Organisation
FCV	Fragility, Conflict and Violence
FEED	Front-End Engineering Design
FID	Financial Investment Decision
GNI	Gross National Income

GATT	General Agreement for Tariffs and Trade
GPA	Government Procurement Agreement
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
HIC	High Income Countries
HIPC	Heavily Indebted Poor Countries
HSE	Health Safety and Environment
IA	Implementing Agency
IADB	Inter-American Development Bank
IAEA	International Atomic Energy Association
IAEG	Inter-Agency and Expert Group
IBRD	International Bank for Reconstruction and Development
ICAI	International Commission for Aid Impact
ICAO	International Civil Aviation Authority
ICR	Implementation Completion Results Report
ICRR	Implementation Completion Results Report Review
ICSID	International Centre for Settlement and Investment Disputes
IDA	International Development Association
IEG	Independent Evaluation Group
IFAD	International Fund for Agriculture and Development
IFC	International Finance Corporation
IFI	International Financial Institution
ILO	International Labor Organisation
IMF	International Monetary Fund
IMO	International Maritime Organisation
INT	Institutional Integrity Department
IOM	International Organisation for Migration
IP	Implementation Partner
IPF	Investment Project Financing
ISDB	Islamic Development Bank
ITO	International Telecommunication Union
JICA	Japan International Co-operation Agency
JUI	Joint Inspections Unit of the United Nations System
KPMG	Klynveld Peat Marwick Goerdeler
LBD	Local Business Development
LCR	Latin and Central America
LIC	Low Income Countries
LNG	Liquid Natural Gas
MAPS	Method for Assessing Procurement Systems
MDB	Multilateral Development Bank
MDG	Millennium Development Goals
MfDR	Managing for Development Results
MDRI	Multilateral Debt Relief Initiative
MIC	Middle Income Countries
MIGA	Multilateral Investment Guarantee Agency
MNA	Middle East Region
NDB	New Development Bank
NGO	Non-Governmental Organisation

ODA	Official Development Assistance
ODI	Open Data Institute
OECD	Organisation for Economic Co-operation and Development
OKF	Open Knowledge Foundation
PCR	Project Completion Report
PEAT	Project Efficiency Assessment Tool
PMCC	Pearson's Product Moment Correlation Coefficient
PMA	Procurement Maturity Assessment
PPAR	Project Performance Assessment Report
PPER	Project Performance Evaluation Report
PVR	Project Valuation Report
ROAR	Results Oriented Annual Report
ROI	Return on Investment
ROSMA	Return on Supply Management Assets
SAPE	Sector Assistance Programme Evaluation
SCM	Supply Chain Management
SCOR	Supply Chain Operations Reference
SEI	Software Engineering Institute
SME	Small and Medium Enterprise
TQM	Total Quality Management
UN	United Nations
UNAIDS	United Nations
UNCTAD	United Nations Conference of Trade and Development
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
UNEP	United Nations Environmental Programme
UNESCO	United Nations Educational Scientific and Cultural Organisation
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
UNISDR	United Nations Office for Disaster Reduction
UNODC	United Nations Office of Drugs and Crime
UNOPS	United Nations Office for Project Services
UNPD	United Nations Procurement Division
UNPF	United Nations Population Fund
UNRWA	United Nations Relief and Works Agency for Palestinian Refugees
UN Women	United Nations Women
UPU	Universal Postal Union
USAID	United States Agency for International Development
VFM	Value for Money
WBG	World Bank Group
WFP	World Food Programme
WHO	World Health Programme
WIPO	World Intellectual Property Organisation
WMO	World Meteorological Organisation
WTA	World Tourism Agency
WTO	World Trade Organisation

## **1 INTRODUCTION**

International development is a broad concept and classical definitions combine the approach of a long-term process of structural and societal transformation and short to medium term outcomes and targets (Scholte & Söderbaum, 2017; A. Sumner & Tribe, 2008). There has been a gradual expansion of international development boundaries towards a global development definition that includes global challenges and co-operation to address emerging issues such as sustainable development and climate change (Scholte & Söderbaum, 2017). This chapter provides an introduction to the research and how it was conducted.

### **1.1 A BRIEF DESCRIPTION OF INTERNATIONAL DEVELOPMENT**

International development is comprised of a pluralistic group of different multi-national, national, and non-profit organisations. The definition and understanding of the term Development is often ambiguous and sometimes misunderstood (De Haan, 2009). This is partly because there are multiple forms of development aid such as loans, grants, projects, and technical assistance. There have been different definitions and classifications of recipient groups in the past including; Developed Countries, Developing Countries, the Third World and variations of these terms (Sumner, 2010; World Bank Group, 2020). There are many development organisations of different structures and functions. These include Multilateral Development Banks (MDBs) such as the World Bank Group (WBG) and the Asian Development Bank (ADB), United Nations agencies, bilateral or national development organisations and Non-Governmental Organisation (NGOs). Historically, each of these organisations has been created and established in different ways by the formation of charters, agreements and declarations that outline the organisation membership, governance, structure, mission and goals (United States Institute of Peace, 2018).

Definitions for development also change and these new definitions then become commonly adopted across organisations. An example was the World Bank Group's (WBG) dis-adoption of the term Developing Countries in 2016 and now uses the term Low-Income Countries (LICs), defined as Gross National Income (GNI) of below US\$1,085. Lower Middle-Income Countries (MICs) are defined with GNI of between US\$1,086 and US\$4,255 and Upper-Middle-Income Countries of GNI between US\$4,256 and US\$13,205. High Income Countries with GNI above US\$13,205. For example, MICs represent over 70% of the world population and 60% of the WBG loans are disbursed to these countries (World Bank Group, 2022). The United Nations Statistics Division (UNSD) however, refers to Developed Countries, Developing Countries, Developing Economies in Transition and Least Developed Countries (LDCs) (UNDP, 2022). These terms are routinely used in the Annual Statistical Report on United Nations Procurement (ASR) (United Nations Office for Project Services, 2022). For the purposes of discussion, with the exception of the ASR references in Section 3.2, the World Bank Group's terms (HIC, MIC and LIC ) are the terms used in this research.

Development has taken place at varying rates in different parts of the world, and history has recorded differences in relative economic, political, and military progress for hundreds if not thousands of years. As a result, international development and distribution of wealth can be emotive and controversial subjects and several studies on the effectiveness of development aid, have found the results to be inconclusive (Burnside & Dollar, 2000; Doucouliagos & Paldam, 2009; Hansen & Tarp, 2000).

Major influences and effects on the distribution of global development policy and funding over the last two hundred years include the industrial revolution (Lucas, 2002), the growth of the petrochemical industry (Hassan, Meyer, & Kot, 2019), innovation and technological advances (Allen, 2017) and geopolitical events such as the two World

Wars (Francis, 2020). In more recent times, the advent of globalisation the application of technology and automation in banking has enabled rapid and widespread economic growth although, as Blair (2010) argued, not without introducing new risks (Blair, 2010). Despite many beneficiaries of such advancements, global wealth inequality has persisted and remains a considerable concern where influence, wealth and power is perceived to be disproportionately concentrated (Piketty, 2018). Nevertheless, the record of international development organisations in the creation of global growth is accompanied by prevailing and persistent questions regarding the effectiveness of the disbursement of development funds (Doucouliagos & Paldam, 2009; Easterly, 2003).

Precise measurement and evaluation is required to understand the influence that international development may have had on global growth and the impact and results of development activity on a country, sector, or region (Kaplan, 2009). The measurement and evaluation of development activities is performed by dedicated departments in the major international organisations according to specific parameters based on the OECD-DAC evaluation methodology (Asian Development Bank, 2019; IEG, 2013; OECD, 2019). Such evaluations remain the leading barometer for success of the development aid industry; however, there are questions about the criteria that currently define success and how well they measure how funds are ultimately disbursed (Chianca, 2008).

There are also differences in the definition, application, and the level of measurement of the evaluation criteria between different organisations. Evaluations are targeted on theme, region, country and sector levels and at a discrete project and implementation level and sub-level. Such distinction and level of evaluation is often described as the macro and micro paradox of evaluation measurement (Howes, Otor, & Rogers, 2011; Mosley, 1986). There are fundamental challenges in establishing the legitimate impact of development, and it is not always clear or demonstrable that more aid leads to more



growth. Roodman (2007) noted that more growth can lead to less aid as donors perceive that less development aid is required (Roodman, 2007). Further, microeconomic data evaluation of aid-financed projects shows that the majority of projects are successful while analysis of macroeconomic data are frequently less positive (Howes et al., 2011).

The effectiveness and performance of development impact, and its results are the subject of increasing focus and attention (Khagram, Thomas, Lucero, & Mathes, 2009; OECD-DAC, 2019). Pressure on the global economy and a changing political climate has led to calls for more accountability, evidence of cost performance and value for money. The issue was so significant that global-scale forums, like the Paris Declaration on Aid Effectiveness, 2005 and the Accra Agenda for Action, 2008, resolved to improve Development Effectiveness (DE) and set the path towards the development of the concept of Value for Money (VFM) (Love et al., 2018; Wood et al., 2011).

The quantities of funding dedicated to development are overwhelming. The international development industry spent an unprecedented US\$161.2 billion of global donors' and taxpayers' money in 2020, boosted by COVID-19, towards multiple development purposes with numerous financial and non-financial modes of delivery (OECD, 2021). It is estimated that over half of Official Development Assistance (ODA) is disbursed as goods and services through procurement (Ellmers, 2011; OECD, 2017; 2021). The international development institutions such as the MDBs and UN agencies are responsible for the delivery and impact of their development objectives and to ensure that every dollar is maximised for the purpose directed. However, there are recurring questions on the effectiveness, efficiency and value for money of the funds put towards international development goals (Easterly & Pfitze, 2008). A considerable part of the development industry is concerned with the matter of securing funds and advocacy towards goals, outcomes and projects across multiple countries and sectors. The critical

responsibility for the implementation and delivery of programs and projects is held directly by the international organisations. These organisations utilise a combination of Implementation Agencies (IAs), Non-Government Organisations (NGOs), suppliers and consultants (Reinsberg & Westerwinter, 2021; World Bank Group IEG, 2016). The measurement and evaluation of the performance of international development for MDBs, United Nations agencies is managed and reported by dedicated evaluation departments and groups (Asian Development Bank, 2019; World Bank Group IEG, 2016).

International development organisations have several evaluation reports designed to measure Development Effectiveness (DE) from different perspectives. While they claim to evaluate cost and efficiency, they often lack detailed criteria on impact effectiveness and how and where funds are disbursed (Clements, Chianca, & Sasaki, 2008; Khagram et al., 2009). Criticism of current development evaluation models concerns the lack of clarity between the measurement indicators at the outcome level of evaluation and the output and project levels and inconsistency in cost analysis and reporting (Denizer, Kaufmann, & Kraay, 2013). Also, the rating method used in the evaluation of development results is often subjective, difficult to correlate and the predominance of self-evaluation in projects lacks impartiality (Clements et al., 2008; Denizer et al., 2013).

International development organisations disburse funds through different methods, and this is mainly dependent on their charter, mission and the type of development aid and the delivery mechanism (Asian Development Bank, 2022; International Monetary Fund, 2021; United Nations, 2019). Funds are disbursed in the form of loans, grants, technical assistance and projects of multiple types, scale and purposes (Brech & Potrafke, 2014). Development is channelled by region, country and sector, including governance, finance, health, education, energy, transportation, industry and water. Other development objectives include disaster preparedness, humanitarian relief, sustainability, climate

change, gender equality and migration (Khang & Moe, 2008). Questions on how effective development organisations are at achieving their objectives and in disbursing and accounting for donor and taxpayer's funds persist and have led to criticism and calls for improvement in effectiveness (Doucouliagos & Paldam, 2011; Maren, 2009).

This research reviews development funds disbursed in the form of projects managed through the procurement and project management processes, which represent the two mechanisms of project delivery. Procurement and project management performance is assessed from primary and secondary data from a procurement questionnaire and project evaluation reports of international development organisations. The research analyses the relationship between the procurement and project management processes, the project evaluation criteria and cost factors. The results are used to examine the contribution and influence of the procurement and project management processes towards DE and VFM.

## **1.2 OVERVIEW OF THE RESEARCH**

The research seeks to understand the performance of the procurement and project management processes in development projects, being the critical mechanisms for effective project delivery. The research examines the relationship between the procurement and project management processes and their relative contribution towards project performance and outcomes. The results are analysed to explore whether the procurement and project management processes and criteria contribute towards Development Effectiveness (DE) and Value for Money (VFM) in development projects.

The first phase of the research was to assess and understand the performance of procurement in leading international development organisations and whether existing Procurement Maturity Assessment (PMA) models could be applied towards this purpose. After the literature review and analysis of existing PMAs, a new Integrated Procurement

Maturity Model (IPMM) was developed for the assessment of procurement maturity and capacity of international development organisations. The IPMM assesses and compares the procurement performance of ten selected international development organisations and three private sector organisations. The main objective of the IPMM is to assess the procurement maturity and performance of the procurement process and explore the relationship and contribution of procurement to project performance, DE and VFM. The second phase of the research was to provide review and analysis of the project evaluation process from a sample of ( $n = 1,920$ ) project evaluation reports for the ten international development organisations collected between 2015 and 2017. The review and analysis included comparison of the overall project evaluation of the international development organisations by region and sector with further review of the individual project evaluation criteria. The research explored the roles and relationship of both the procurement and project management processes in international development projects and examined their combined contribution towards project performance, outputs, outcomes, DE and VFM. In addition, the research conducted analysis of cost performance and Critical Success Factors (CSFs), such as cost, schedule and quality to seek new knowledge on the measurement of development performance in projects.

Chapter 2 introduces the main international development organisations, the MDBs, UN agencies and bilateral development organisations, and provides an overview of their operating environment. Chapter 3 provides a detailed but not exhaustive literature review, Chapter 4 explains the methodology used in the research and Chapter 5 describes the research results and findings. Research discussion is provided in Chapter 6 and Chapter 7 outlines conclusions, limitations and recommendations for further research.

## **2 OVERVIEW OF INTERNATIONAL DEVELOPMENT ORGANISATIONS**

This chapter is dedicated to outlining the three key types of organisations that are the subject of this research: Multinational Development Banks (MDBs), Agencies of the United Nations and bilateral development organisations. To understand the scale and method of disbursement of development funds it is necessary to consider the nature and type of the multiple organisations involved in international development activities.

### **2.1 INTRODUCTION**

The MDBs, UN agencies and bilateral development organisations are all development organisations but are comprised of a numerous and complex group of different entities that often work in overlapping geographies, sectors, and ministries. The following description and summary provides some perspective and background to the subject of this research. The different types of international development organisations are often referred to as bilateral and multilateral and each have different operational objectives and use different types of financial disbursement methods (Gulrajani, 2016). These include loans, grants, programmes, projects, and other mechanisms. International development organisations have struggled to demonstrate and convince critics that the results and impact of development funds disbursement is effective (Ika, Diallo, & Thuillier, 2012). Projects are a prevalent mode of development implementation and the focus of this research. The Procurement process and function plays an important and integrated role when projects include goods and services and suppliers and contractors are required to implement projects (de Araújo, Alencar, & de Miranda Mota, 2017).

The International Monetary Fund (IMF) and World Bank Group (WBG) were created in July 1944 at Bretton Woods in New Hampshire, USA, as a result of an international conference of 44 nations. The membership of both organisations stands at 189-member nations each. The two organisations were designed to form the basis of a renewed and

firm international economic and financial charter after the turmoil of the first part of the twentieth century (United Nations, 2019).

Shortly after, in October 1945, as a replacement for the League of Nations, the United Nations (UN), was formed as another group of international organisations to prevent a repeat of similar global conflict. At its founding, the UN had 51 member countries and currently has 193 member countries and purposes, principles and organisation outlined in the charter (United Nations, 2019). The United Nations has six principal bodies: The General Assembly; the Security Council; the Economic and Social Council; the Trusteeship Council; the International Court of Justice and the Secretariat. These principal bodies are not well understood and confused in general discussion regarding the United Nations. The General Assembly is the only body in which all the members are represented and is responsible for supervisory, financial and elective functions related to the UN charter. The Security Council has primary responsibility for the maintenance of internal peace and security and has five permanent members and ten non-permanent members. The Security Council reviews complaints or disputes from any country and initially explores the possibility of peaceful resolution. The Security Council can authorise peacekeeping forces to be deployed to ensure stability pending negotiations. The Security Council may also call upon UN members to apply diplomatic or economic sanctions, and the UN charter allows the possibility of military action to be taken against a country (s) where no alternative is considered possible (United Nations, 2019).

It is the Economic and Social Council (ECOSOC) that is responsible for the economic, social, humanitarian and cultural areas of the UN and specialised agencies included in this research. Many global development programmes and projects concern human rights, narcotics, population, social development, environment, statistics, gender equality and diversity, science, and technology. Programmes are often comprised of several

component projects (Papadaki et al., 2014). The Secretariat acts as the chief administrative office of the United Nations and an important political function acting as chief spokesperson and high-level negotiator (United States Institute of Peace, 2018). The United Nations organisations have adapted and changed significantly since establishment following global development, the changing environment and pressures to respond to emerging challenges, disasters and emergencies (United Nations, 2019). The UN, specialised agencies and bilateral development organisations represent the largest source of disbursement of development funds for LIC, MIC and borrowing countries (United States Institute of Peace, 2018).

In terms of development effectiveness, the Independent Evaluation Group (IEG) of the World Bank Group claimed that 39% of World Bank projects are rated below satisfactory (Chauvet, Collier, & Duponchel, 2010; World Bank, 2020). Concerns with project management include aspects such as poor project design, poor stakeholder management, delays, cost overruns and coordination between stakeholders (Ahsan & Gunawan, 2010). In 2020 over US\$161.2 billion per annum is estimated as disbursed on international development including loans, grants and over 50% through procurement and contracts (Ellmers, 2011; OECD, 2021). The Paris Declaration on aid effectiveness noted donors are often behind in commitments due to lack of policy structures, lack of compliance, poor alignment and disconnects between corporate strategies and aid agendas (Wood et al., 2011). The deficiencies in the realisation of DE remain unspecified, and there is a growing recognition of the requirement to improve evaluation methods for the measurement of how funds are disbursed (Easterly & Pfitze, 2008; Gulrajani, 2014).

## 2.2 MULTILATERAL DEVELOPMENT BANKS

Multilateral development banks (MDBs), also called International Finance Institutions (IFIs), provide financing and professional advice for development in borrowing countries. MDBs have large memberships including HIC, LIC and MIC borrower countries. MDBs finance projects in the form of loans at market rates, loans below market rates, grants and technical assistance (World Bank Group, 2020).

The World Bank Group (WBG), for example, provides two main types of Investment Project Financing (IPF) loans names as International Bank for Reconstruction and Development (IBRD) and International Development Association (IDA) loans. These loans provide support, often as projects, towards a wide variety of purposes, chiefly used in the infrastructure, human development, agriculture, and public administration sectors. IBRD loans are provided to MICs at competitive market rates made possible by the AAA ratings of the organisation and between 0.5% and 1.0% depending on the repayment term of under 12 years' and up to 18 years' average with a maximum maturity of 30 years. IBRD loans are mainly used for loans, guarantees, risk management, knowledge and technical advisory services (Sumner, 2010; World Bank Group, 2020).

IDA provides loans on concessional terms to 75 poorest member countries, 39 of which are in Africa with zero or very low interest charges and repayments are scheduled over a term of 25 to 40 years with a 5 to 10 year grace period. IDA also provides significant levels of debt relief through the Heavily Indebted Poor Countries (HIPC) initiative and the Multilateral Debt Relief Initiative (MDRI). For example, since 1960, IDA has provided US\$345 billion for investments in 113 countries (World Bank Group, 2020). MDBs also often provide trust funds and grants in fragile and crisis-affected situations where the ability to lend is difficult and to provide immediate assistance to natural



disasters and emergencies and provide financing, direct investment and guarantees through MIGA and IFC (World Bank Group, 2020).

The most prominent MDBs are as shown in Table 1.

TABLE 1. MULTILATERAL DEVELOPMENT BANKS

The World Bank Group (WBG) - itself comprised of five institutions; the International Bank for Resettlement and Development (IBRD); International Centre for Settlement and Investment Disputes (ICSID); International Development Association (IDA); International Finance Corporation (IFC) and the Multilateral Investment Guarantee Agency (MIGA)
International Fund for Agriculture and Development (IFAD)
European Investment Bank (EIB)
Asian Development Bank (ADB)
African Development Bank (AFDB)
Inter-American Development Bank (IADB)
Islamic Development Bank (ISDB)
European Bank for Reconstruction and Development (EBRD)
Asian Infrastructure Investment Bank (AIIB)
Development Bank of Latin America (CAF)

In addition, there are several sub-regional multilateral development banks, which have only member countries and lend to borrowing countries from international capital markets shown in Table 2. The loans are competitive partly due to shared responsibility for repayment from the member countries, which are global in the case of WBG and IFAD, or regional in the case of ADB, EBRD, AFDB and others.

TABLE 2. SUB-REGIONAL MULTILATERAL DEVELOPMENT BANKS

Caribbean Development Bank (CDB)
Central American Bank for Economic Integrations (CABEI)
East African Development Bank (EADB)
West African Development Bank (BOAD)
Black Sea Trade and Development Bank (BSTDB)
Economic Co-operation Organisation Trade and Development Bank (ETDB)
Eurasian Development Bank (EDB)
New Development Bank (NDB) formerly BRICS Development Bank

The World Bank Group has operations globally in multiple sectors including governance, health, education, transportation, infrastructure and agriculture. In 2016, WBG disbursed over US\$50 billion in the form of loans, grants and financial instruments to borrowing countries in over 186 countries (World Bank Group, 2017).

### 2.3 UNITED NATIONS AGENCIES

In addition to Multilateral Development Banks, there are the United Nations Agencies. The UN system is a complex group of over 34 organisations made up of the United Nations and many affiliated programmes, funds, and specialised agencies, all with their own membership, leadership, and budget. The programmes and funds receive finance through voluntary contributions from member countries, and some take public donations and contributions (Global Policy Forum, 2013; United Nations, 2019). The Specialised Agencies are independent international organisations funded by both voluntary and assessed contributions. Some of these organisations are representatives of the UN system, frequently appearing in publications and press coverage, some are MDBs, and others are much less well known. The specialised United Nations Agencies are integrated into the UN system under the UN Charter Article 57, and include the IMF and WBG (United Nations, 2018). The specialised UN agencies have different mandates from advocacy, disbursement of

financial instruments and project delivery and management. The smaller UN organisations are referred to as related UN agencies and have a relationship with the UN defined by agreements established under Article 57 and 63 of the United Nations Charter (United Nations, 2019). The programmes and funds are comprised of the following 12 prominent organisations; as specified in Table 3.

TABLE 3. UNITED NATIONS AGENCIES

United Nations Development Fund (UNDP)
United Nations Children’s Fund (UNICEF)
United Nations High Commissioner for Refugees (UNHCR)
World Food Programme (WFP)
United Nations Office of Drugs and Crime (UNODC)
United Nations Population Fund (UNPF)
United Nations Conference of Trade and Development (UNCTAD)
United Nations Environmental Programme (UNEP)
United Nations Relief and Works Agency for Palestinian Refugees (UNRWA)
United Nations Women (UN Women)
United Nations Human Settlements Programme (UN-Habitat)
International Trade Centre (UN/WTO)

In addition to the programmes and the UN specialised agencies are independent organisations, which work with the United Nations. They have formed a relationship with the UN through individually negotiated agreements. Some of these predate the First World War and are associated with the League of Nations, others have been created in the wake of the Second World War. Another group, such as UN Women, has been established to meet emerging diversity and advocacy requirements (Charlesworth & Chinkin, 2013). The UN specialised agencies are shown in Table 4.

TABLE 4. UNITED NATIONS SPECIALISED AGENCIES

The World Bank Group (WBG)
The International Monetary Fund (IMF)
World Health Organisation (WHO)
United Nations Educational, Scientific and Cultural Organisation (UNESCO)
International Labor Organisation (ILO)
Food and Agriculture Organisation (FAO)
International Fund for Agriculture and Development (IFAD)
International Maritime Organisation (IMO)
World Meteorological Organisation (WMO)
World Intellectual Property Organisation (WIPO)
International Civil Aviation Organisation (ICAO)
International Telecommunication Union (ITU)
United Nations Industrial Development Organisation (UNIDO)
Universal Postal Union (UPU)
World Tourism Agency (WTA)

There are other groups listed other entities and related organisation in Table 5.

TABLE 5. UNITED NATIONS OTHER ENTITIES AND RELATED ORGANISATIONS

The United Nations Programme on HIV/AIDS (UNAIDS)
United Nations Office for Disaster Reduction (UNISDR)
United Nations Office for Project Services (UNOPS)
International Atomic Energy Association (IAEA)
International Organisation for Migration (IOM)
World Trade Organisation (WTO)
Commission for the Comprehensive Nuclear Test Ban Treaty Organisation (CTBTO)
The Organisation for Prohibition of Chemical Weapons (OPCW)

#### 2.4 BILATERAL DEVELOPMENT ORGANISATIONS

As well as Multilateral Development Banks (MDBs) and the United Nations Agencies, Bilateral development organisations are main feature of the global development organisations. Bilateral development organisations include national governments and the EU and direct their Official Development Assistance (ODA) to

support development projects in many sectors including health, governance, sustainable development, education, transportation, infrastructure and agriculture. They also work closely with MDBs, UN agencies and Non-Government Organisations (NGOs) in joint projects or by channelling and disbursing funds as part of development agreements (Mitchell, 2014). The largest bilateral development donors in the order of spend value for 2018 are represented in as Figure 1 follows:

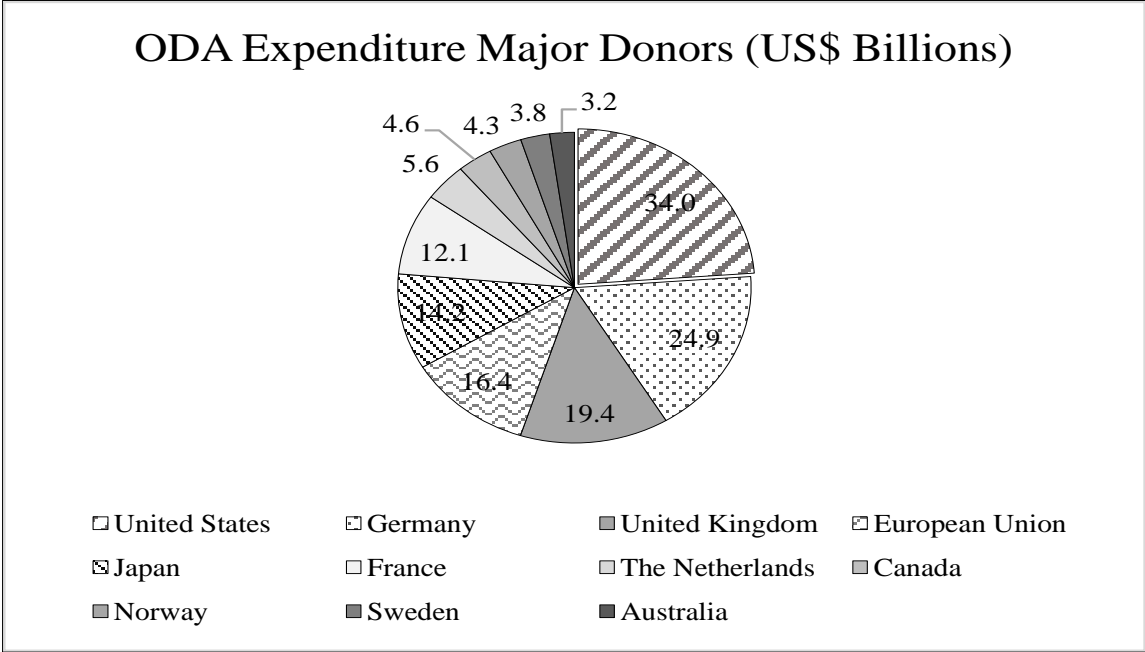


FIGURE 1. OFFICIAL DONOR ASSISTANCE EXPENDITURE BY MAJOR DONORS IN 2018 (OECD, 2020).

The bilateral development organisation such as Department for International Development (DFID), Department of Foreign Affairs and Trade (DFAT), United States Agency for International Development (USAID) and Japan International Co-operation Agency (JICA) have a combination of financial disbursement mechanisms that include grants, programmes, technical assistance and projects. These organisations also collaborate with both MDBs and UN agencies for joint development initiatives and disbursement of funds on their behalf (Gulrajani, 2016).

## 2.5 SUMMARY

The large number of development organisations with multiple, often competing missions and objectives ensures a broad global coverage for many sectors of development. At the same time, there is significant overlap both in geography and sector focus combined with challenges of coordination between organisations. Overall, development organisations are loosely classified as Multilateral Development Banks (MDBs), UN agencies and bilateral organisations. This classification framework will be used in subsequent parts of this research for data analysis. The complexity and over-lapping nature of many of these organisations contribute to some of the criticism of DE, VFM and delivery of development results (Bourguignon & Sundberg, 2007). Now that a comprehensive overview of the industry context has been provided in Chapter 2 and Chapter 3 of this document will overview the theory behind procurement and project process performance in the development industry.

### **3 LITERATURE REVIEW**

This chapter is dedicated to a critical analysis of the literature surrounding procurement and project management performance and discussion of the definition and assessment of Development Effectiveness (DE) and Value for Money (VFM). The chapter reviews the literature and current knowledge on the performance of development projects, commitments to improving measurement and delivery, and the effects of corruption and waste. The section closes with the presentation of a theoretical framework for the research, the research questions and research aims and objectives before moving on to the chapter dedicated to research design.

#### **3.1 INTRODUCTION**

The theory of procurement and its impact on development effectiveness and value for money in projects provides the underlying basis of this research. Over half of Official Development Assistance (ODA) is disbursed in the form of goods and services through procurement (Ellmers, 2011). The main assumption is that robust procurement performance should greatly influence the improved performance of DE and VFM (Biscaye et al., 2017; Mensah, 2016; Rimkūnienė, 2013). Development evaluations are an embedded part of the procurement and project management process and are performed by the evaluation department of respective international development organisations. For example, these include the Independent Evaluation Group (IEG) of WBG and the Independent Evaluation Department (IED) of ADB, which follow the Organisation for Economic Co-operation and Development's (OECD) methodology. The OECD methodology evaluates the criteria: relevance; coherence; effectiveness; efficiency; impact and sustainability as defined in this Chapter (Ahsan & Gunawan, 2010; Asian Development Bank, 2019; OECD, 2019; World Bank Group, 2012).

The measurement of procurement performance is most frequently assessed using Procurement Maturity Assessment models (PMAs) which have developed from earlier project management maturity models; there are several leading PMAs utilised in both the public and private sector. As with many business functions, there are metrics for measuring internal procurement process performance and the PMAs enable additional perspectives from benchmarking (Brandmeier & Rupp, 2010). In international development, the OECD has developed the Methodology for Assessing of Procurement Systems (MAPS) tool (OECD, 2010), which forms the basis of PMAs used by many MDBs, UN and bilateral development agencies for procurement assessments. The MAPS model is an assessment tool that is chiefly concerned with the regulatory environment, governance and compliance of the procurement process and less focussed on process effectiveness and efficiency measures. Public and private sector PMA models have evolved to incorporate criteria that assess best practices of the procurement process, organisation and technology including effectiveness and efficiency (Safari et al., 2021). This research reviews existing PMA models used in the public and private sectors and in development organisations and examines how they capture and measure criteria such as effectiveness, efficiency and VFM. The focus on procurement is due to the evidence that approximately 50% of ODA is disbursed and procured in the form of goods, equipment services and works (Ellmers, 2011). For example, in ADB, from January 1966 to December 2015, 63% of loan project funds globally were expended on the procurement of goods, services and works (Ahsan & Paul, 2018; Asian Development Bank, 2018). Secondly, both the procurement and project management processes include performance indicators that enable the measurement of Development Effectiveness (DE) and Value for Money (VFM) in projects (Lindstrom, 2014).



There are existing measures of development performance and DE performed by the evaluation departments of international development organisations, and these mostly follow the OECD methodology of evaluation with some variation between different organisations (OECD-DAC, 2012).

The issue of DE has arisen from a growing perception that there are challenges and weaknesses regarding the effectiveness of ODA and the disbursement of development funds (de Montclos, 2012; Easterly & Pfitze, 2008). There are various terms and descriptions used to describe and measure development results and effectiveness including effectiveness, outcomes, outputs and spending activity illustrated in Figure 2 (Biscaye et al., 2017, p. 1429). Figure 2 illustrates the four levels of the hierarchy of performance measurement for DE and suggests that each level can be measured independently. As the definition of DE remains elusive, there can be a disconnect between the macro and the micro-levels of development in practice (Denizer et al., 2013; Khagram et al., 2009). This research undertakes analysis of the level of development activity that includes procurement and project management and examines the connectivity between the micro and macro-levels of measurement of delivery of DE.

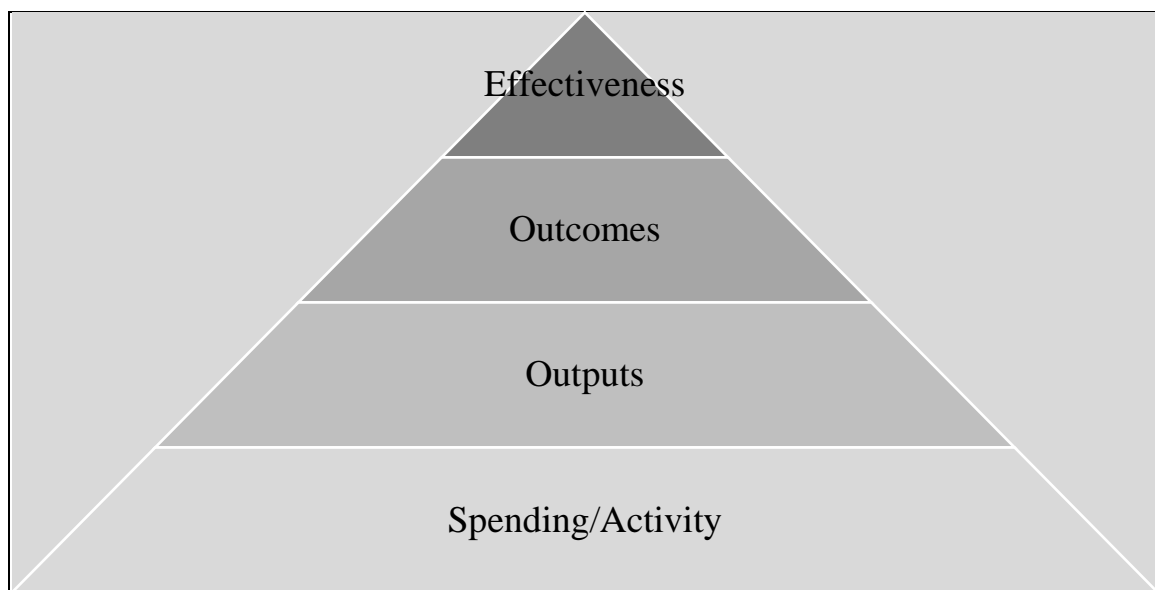


FIGURE 2. HIERARCHY OF PERFORMANCE MEASUREMENT DEVELOPMENT EFFECTIVENESS (Biscaye, LaFayette, & Martin, 2015, p. 2; Biscaye et al., 2017, p. 1428)

Several initiatives have been developed to pursue both the measurement and improvement of DE, including Management for Development Results (MfDR) reports and regular development evaluation reports (McLiesh & Arizti, 2006). Many agencies produce annual reviews of their development results such as the WBG's Annual Review of Development Effectiveness (ARDE), the UN's Results-Oriented Annual report (ROAR) and DFID's Development Effectiveness Report. The DFID report acknowledges the difficulty in reporting development goals and demonstrating outcomes. All of the reports from development organisations indicate that they have challenges connecting the different levels of measurement with reliable evidence of results (White, 2005).

The measurement of development effectiveness has been inconsistent. This issue was addressed at the United Nations Conference on Financing for Development Report, Monterrey, Mexico, 2002 to the Preparatory Committee. The report contained 87 recommendations to remedy the problem of inconsistency in the measurement of development effectiveness. Ironically, the report did not include plans for the implementation of the recommendations (Picciotto, 2002; United Nations, 2002).

However, as a result of the conference, public sector organisations have sought ways to increase accountability with increased transparency, social responsibility, triple bottom line reporting and other forms of performance measurement. The levels of accountability are reflected by development organisations within corporate scorecards defined as impacts, outcomes, reach, outputs and inputs from highest to the lowest level (Lloyd, Poate, & Villanger, 2014; Picciotto, 2002, 2018). Another outcome of the conference was the development of OECD-based standard methods for results measurement established across donor and development organisations (Biscaye et al., 2015; Biscaye et al., 2017). The conference was followed by the Paris Declaration for Aid Effectiveness 2005, Accra Agenda for Action 2008, High-Level Forum for Aid Effectiveness in Busan, 2011 and development of the Sustainable Development Goals (SDGs) (Eyben, 2013). As discussed, the procurement process contributes substantially towards the development goals, programmes and projects and potentially provides a window to answer Easterly and Pfutze (2008) question of “where does the money go?” (Ahsan & Paul, 2018; Easterly & Pfutze, 2008). The role of procurement is further explored in the next section.

### **3.2 THE ROLE OF PROCUREMENT IN INTERNATIONAL DEVELOPMENT**

The procurement function has not been given the same recognition, understanding, allocation of resources and awareness in LICs and MICs as in HICs despite the efforts of major international organisations and this may be deliberate or due to lack of awareness (Mensah, 2016). In LICs and MDCs, procurement activities are transitioning from a clerical, transactional, non-strategic approach towards a strategic, value-adding process with the ability to influence organisational decision making. The reforms include steps towards changes in procurement regulations, methods and processes. There remain many challenges with the measurement and effective benchmarking of procurement

performance despite the use of several procurement assessment models (Harland, Telgen, Knight, Callender, & Thai, 2007).

Approximately 50% of the ODA development aid of the total US\$131.5 billion in 2015, US\$142.6 billion in 2016 and US\$146.6 billion in 2017, rising to US\$161.2 billion in 2020, was procured in the form of goods and services (OECD, 2017; 2021). This significant spend substantiates and supports the argument of the importance of the procurement process and function for implementation and delivery of international development (Ellmers, 2011; OECD, 2017; OECD Inter Agency Task Force on Financing for Development, 2017). Despite this, the function is often poorly understood, under-resourced and incorporates different procurement processes, methods and priorities in respective development organisations (Tassabehji & Moorhouse, 2008). However, due to a combination of factors that include economic pressure, political change, and specific challenges, such as natural disasters and emergencies, procurement started to receive increased focus, interest and attention (Mena, Christopher, & van Hoek, 2014). The World Trade Organisation (WTO) encourages public procurement to be efficient and to increase transparency and accountability (Hoekman, 1998). Competition is one of the key principles of public procurement, and it has long been understood that competition among suppliers can result in savings and efficiencies without compromising the quality of goods and services (Domberger, Hall, & Li, 1995; Knight et al., 2012).

In the late 1970s under the General Agreement for Tariffs and Trade (GATT), the Government Procurement Agreement (GPA) was negotiated by several high-income countries. The GPA seeks to ensure the extension of the principles of non-discrimination, national treatment and transparency to the competitive procurement process of government entities (Mavroidis, 2005). The issue of non-discrimination would seem to

be unambiguous in increasing competition and mirror the effect of the marketplace. However, organisations increasingly encourage contracts with specific groups such as indigenous minorities, female-owned business, gender equality, marginalised communities and domestic business in LICs and MICs (World Bank Group, 2013). The preference for specific attributes of bidders can be achieved through allocating additional evaluation points for those preferred criteria, although this may result in increased costs, depending on the desired attribute. For example, preferred criteria may include attributes such as specialist technical, Health, Safety and Environment (HSE) and Corporate Social Responsibility (CSR), which are widely used to select contractors in the natural resources industries (Ernst & Young, 2014). Many organisations support the case for competitive bidding, and Transparency International (2018) notes that non-competitive procurement procedures may increase procurement costs by as much as 30%.

Transparency and disclosure of procurement information are also regarded as effective ways to improve the visibility of procurement and decrease fraud and corruption (Rohwer, 2009). However, there are objections against transparency, which include suggestions that it may have the effect of decreasing efficiency, increasing complaints, time for publishing bids and extend process cycle times (Halachmi & Greiling, 2013).

To provide perspective for visibility of the annual estimated procurement expenditure under management by the MDBs, the approximate procurement volume and value between 2015 and 2016 is shown in Table 6.

TABLE 6. PROCUREMENT EXPENDITURE OF MULTILATERAL DEVELOPMENT BANKS

Multilateral Development Bank (MDB)	2015 (US\$)	2016 (US\$)
World Bank Group (WBG)	12.5bn	14.3bn
Asian Development Bank (ADB)	6.4bn	6.5bn
African Development Bank (AfDB)	2.1bn	1.7bn
European Bank for Reconstruction & Development (EBRD)	1.9bn	2.1bn
Inter-American Development Bank (IADB)	4-5bn	4-5bn
Islamic Development Bank (IsDB)	4.9bn	5.77bn

(Africa Development Bank, 2018; Asian Development Bank, 2020; European Bank for Reconstruction and Development, 2017; Islamic Development Bank (ISDB), 2016; Roberto Aiello, 2016; World Bank Group, 2017).

The largest volume and value of procurement expenditure is represented by operational procurement for goods, services and works and funds are disbursed in the form of loans and grants to borrowing countries which manage the procurement processes and procedures (Asian Development Bank, 2022; World Bank Group, 2020). The MDBs' operational procurement function is concerned with oversight, governance and reporting of procurement performed by borrowing country and Implementing Agencies (IAs). MDBs make a distinction between the corporate procurement function, which handles internal institutional procurement, and corporate procurement functions (OECD, 2015).

The specialised agencies, with the largest procurement expenditure, include the United Nations Development Programme (UNDP) disburses grants and manages multiple sector areas with an emphasis on project management and delivery of goods, works and services and has an annual procurement expenditure of US\$2.76 billion in 2015 and US\$1.70 billion in 2016 (United Nations Office for Project Services, 2017). UNDP manages sectors involved in the environment, health, governance, poverty reduction and crisis prevention and recovery and has some overlaps with other UN agencies. The United Nations Procurement Division (UNPD) with the largest procurement expenditure manages procurement, logistics and support to peacekeeping operations and annual expenditure of US\$3.1 million in 2015 and US\$3.32 million in 2016. The United Nations

Children's Fund (UNICEF) manages significant project procurement of a value of US\$2.4 million in 2015 and has a leading advocacy role in support of its activities.

The World Health Organisation (WHO) has a proportionately high element of goods delivery and procurement of medical and supporting goods, equipment, and services of US\$881 million in 2015 and US\$757 million in 2016. The United Nations agencies disbursed US\$18 billion in 2016 and US\$22.3 billion in 2020. This includes goods, services, works and consultancy (United Nations Office for Project Services, 2020).

TABLE 7. OPERATIONAL PROCUREMENT EXPENDITURE OF UN AGENCIES

UN Agency	Goods (US\$ million)	Services (US\$ million)	Total (US\$ million)	Total (%)
	2016	2016	2016	2016
UNICEF	2,630,480	854,713	3,485,193	19.67
UNPD	1,301,515	1,931,699	3,233,215	18.25
WFP	1,605,301	1,356,602	2,961,904	16.72
UNDP	576,278	1,121,188	1,697,466	9.58
UNHCR	562,628	617,133	1,179,761	6.66
UNOPS	321,859	578,320	900,179	5.08
WHO	187,833	569,107	756,940	4.27
PAHO	677,512	112,162	789,674	4.45
IOM	202,537	319,952	522,489	2.95
FAO	158,019	158,985	317,004	1.78
Others	484,962	1,383,745	1,868,707	10.55
<b>Total</b>	<b>8,708,924</b>	<b>9,003,607</b>	<b>17,712,531</b>	<b>100</b>

([ref]: United Nations Office for Project Services, 2017, p. 125)

Table 7 shows that US\$17,712,531 was disbursed by UN agencies in 2016 in the form of goods and services awarded to suppliers throughout the world and therefore benefits the recipient countries as well as their local suppliers and contractors.

In 2016, the most considerable amount of procurement (38.50%) was channelled through suppliers and contractors from Developed Countries shown in Table 8. There are several initiatives to promote the use of country suppliers in Developing countries and Least Developed Countries (LDC) to improve supplier capacity and the financial benefits derived from the use of local business and infrastructure (United Nations Office for

Project Services, 2020). Organisations such as the Multilateral Development Banks (MDBs) often require or encourage LDCs to use established MDB and UN procurement procedures and bidding processes and these procedures are familiar in many African countries (Quinot & Arrowsmith, 2013).

TABLE 8. SUPPLIERS TO UNITED NATIONS AGENCIES BY COUNTRY AND REGION

<b>Supplier Country and Region</b>	<b>Goods</b>	<b>Services</b>	<b>Total</b>
	<b>2016</b>	<b>2016</b>	<b>2016</b>
<b>Developed countries</b> (US\$ thousand)	3,580,832	3,238,499	6,819,331
Percentage (%)	20.22	18.28	38.50
<b>Developing countries &amp; economies in transition</b> (US\$ thousand)	4,092,696	3,856,018	7,943,714
Percentage (%)	23.11	21.77	44.88
<b>Africa</b> (US\$ thousand)	375,054.1	744,182	1,119,236
Percentage (%)	2.12	4.20	6.32
<b>Arab Countries</b> (US\$ thousand)	56,004	107,056	163,060
Percentage (%)	0.32	0.60	0.92
<b>Asia</b> (US\$ thousand)	3,195,393	1,970,650	5,166,042
Percentage (%)	18.04	11.13	29.17
<b>Europe</b> (US\$ thousand)	156,533	447,837	604,370
Percentage (%)	0.88	2.53	3.41
<b>Latin America and the Caribbean</b> (US\$ thousand)	305,292	524,726	830,019
Percentage (%)	1.72	2.96	4.69
<b>Northern America</b> (US\$ thousand)	0.00	38,542	38,453
Percentage (%)	0.00	0.22	0.22
<b>Oceania</b> (US\$ thousand)	4,420	23,114	27,535
Percentage (%)	0.02	0.13	0.16
<b>Least Developed Countries (LDC)</b> (US\$ thousand)	832,845	1,567,120	2,400,443
Percentage (%)	4.70	8.85	13.55
<b>Unspecified countries</b> (US\$ thousand)	202,551	341,493	544,044
Percentage (%)	1.14	1.93	3.07
<b>Grand Total</b> (US\$ thousand)	<b>8,708,924</b>	<b>9,003,607</b>	<b>17,712,531</b>

([ref]: United Nations Office for Project Services, 2017, p. 125)



The largest global bilateral agency is USAID with over US\$19 billion expended in 2016 and US\$19.3 billion in 2017 in terms of commitments to international development including procurement, project, program, training and contribution to international organisation activities (USAID, 2020). The Japan International Co-operation Agency (JICA) disbursed over US\$7 billion of total bilateral ODA in 2016, which includes loans and grants with a major proportion of procurement activity (JICA, 2017). The Department for International Development (DFID) awarded US\$1.4 billion of procurement and contracts in goods, services, works and consultancies in 2016 (ICAI, 2018). Similarly, the Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH (GIZ) disbursed over US\$1.5 billion and approximately US\$1.48 billion on procurement and contracts in 2016 and 2017, respectively (GIZ, 2016). Table 6 and Table 7 indicate the value of goods and services that are channelled through the procurement process of the MDBs and UN Agencies is assessed at over US\$70 billion between 2016 and 2017.

### **3.3 DEFINING DEVELOPMENT EFFECTIVENESS IN INTERNATIONAL DEVELOPMENT**

International development organisations are interested and increasingly expected to understand and demonstrate whether the funds that they provide to recipient countries are used efficiently to achieve the desired outcomes (Biscaye et al., 2017). The Millennium Development Goals and the later Sustainable Development Goals outline key development targets and results by the international community and donor countries and taxpayers, civil society organisations and individual citizens increasingly demand to know how funds are spent (Biscaye et al., 2017; Sachs, 2012; Stijns, 2012). In order to improve Development Effectiveness (DE), Kaplan's (2001) premise of only managing what can be accurately measured is critical to consider (Kaplan, 2001). The multiple types of financial instruments, such as loans, grants, programs and projects, that

international development organisations include in their charters and mission, make the consistency and definition and measurement of DE a complex challenge (White, 2005).

Development evaluations have different layers of measurement and development is viewed from the outcomes, or macro, perspectives which are an essential viewpoint to the ultimate intention of the development objectives (Chianca, 2008). However, as discussed in Section 3.1, the definition and measurement of outcomes are different from those of outputs (Biscaye et al., 2017; World Bank Group IEG, 2016); Figure 1 suggests that outcomes measure a broader (macro) level of development effectiveness than outputs. The criteria of relevance, effectiveness, efficiency and sustainability used in the evaluation are applied inconsistently to macro and micro levels. The lack of consistency can make it difficult to understand and interpret what is being measured and defined project success and DE at the project level (Chianca, 2008; De Wit, 1988).

Country evaluations targeted at the health and education sectors, for example, primarily approach effectiveness from the macro country outcome perspective and seek to measure the activities at the country level (Asian Development Bank, 2019). Many of the development results, when defined and measured at country or macro-level, are influenced by the activities of multiple external factors, economics, market trends and the activities of other development organisations. The macro definition for measures of success, such as the eradication of a disease on a national basis, for example, represents an outcome-based type of measurement indicator. At the project micro-level for the health sector, for example, there are specific indicators for results, such as the delivery of medicines, hospital construction and medical services rendered. The identification and distinction between macro and micro levels of measurement is essential to in order to distinguish what is being measured (Howes et al., 2011). Secondly, while DE can be expressed at both the macro and micro-level, the micro-level is critical for understanding

DE and VFM for project metrics in the hierarchy performance measurement in Figure 2.

Measurement of results is expressed differently depending on the audience and needs including external stakeholders, internal institutional performance measurement purposes, sector performance measurement and at the level of discrete project delivery. The levels are illustrated as a hierarchy of performance measurement, as shown in Figure 2, which include effectiveness, outcomes, outputs, and direct project, spend and cost management levels and indicators as discussed in Section 3.1 (Biscaye et al., 2015).

The United Nations Conference on Financing for Development in Monterrey, Mexico in March 2002 reached agreement on the basic elements of global partnership towards development effectiveness. The agreement reflected new paradigms concerning results orientation, domestic ownership and improved policies, partnerships between governments, private sector and civil society and a long term holistic approach to development (Picciotto, 2002, 2018). The new approach sets very ambitious goals at the top levels of the hierarchy of performance measurement and demonstrates a shift from project level programs to country-level evaluation as the unit of account (Picciotto, 2018). The United Nations Conference in Monterrey, 2002 led to increased efforts to measure DE, cost benefit analysis towards development effectiveness (Stern, 2004). Efforts to define and measure results have remained mixed in both their methods and success and some of the challenge is due to inconsistency in the measurement and clarity of the macro and micro measurement indicators. The MDBs and UN agencies measure Development Effectiveness (DE) using evaluation criteria based on the OECD Development Assistance Committee (DAC) methods. The development organisations in this research use the OECD criteria to measure development with some variations in rating scales. The criteria are defined as follows (OECD, 2019, pp. 5-12):

- **Relevance:** The extent to which the intervention objectives and design respond to beneficiaries', global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change.
- **Coherence:** The compatibility of the intervention with other interventions in a country, sector or institution.
- **Effectiveness:** The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups.
- **Efficiency:** The extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.
- **Impact:** The extent to which the intervention has, or is expected to, generate significant positive or negative, intended or unintended, higher-level effects.
- **Sustainability:** The extent to which the net benefits of the intervention continue or are likely to continue.

The definitions of relevance and effectiveness refer to the aid activity, and the general description is applied both to the macro and micro and outcome project perspectives.

Efficiency refers to the qualitative and quantitative measurement of outputs in relation to inputs and the achievement of least cost (Zidane & Olsson, 2017). The definition is quite broad and applicable to single projects, groups of projects and large-scale interventions.

Impact is measured at the macro-level of development intervention, and the result of direct and indirect, intended or unintended measured at the aid activity outcomes.

Sustainability has elements of quality criteria and design including environmental, financial and sustainability metrics for detail of inputs and planning and there is inconsistency in the application to micro and macro-level measurements (World Bank Group IEG, 2016). The OECD DAC evaluation criteria have remained relatively

unchanged since 1992, when they were first developed and after revision in 2019 (OECD-DAC, 2019). The majority of the OECD DAC criteria do not measure at the spending activity and procurement level in detail (Biscaye et al., 2017; Chianca, 2008).

As discussed in the following section, the definition of VFM includes similar terms such as economy, effectiveness, efficiency and sometimes, equity or the 4 Es and further defined in Section 3.4 (ICAI, 2011). This overlap of definition concerning DE and VFM further leads to inconsistency and difficulty in benchmarking for results in international development (Clements, 2020; Gulrajani, 2014).

Now that development effectiveness has been discussed in detail, the concept of value for money in development and procurement will be considered in depth.

#### **3.4 DEFINING VALUE FOR MONEY IN INTERNATIONAL DEVELOPMENT**

Similar to Development Effectiveness (DE), Value for Money (VFM) has disparate definitions. The Global Fund was one of the early pioneers of both DE and VFM, motivated by the goal of improved accountability for procurement funds disbursed towards international development in the health sector (Brugha et al., 2004). VFM has developed as a subset of DE and gained more attention and focus following the Paris Declaration of Aid Effectiveness of 2005 (Emmi et al., 2011). The Department for International Development (DFID) and the Global Fund have defined VFM to include the concepts of the 4 Es; economy, efficiency, effectiveness and equity (Glassman et al., 2013; Jackson, 2012). ICAI (2011) defines these as (ICAI, 2011, p. 4):

- Economy: getting the best value inputs
- Efficiency: maximising the outputs for a given level of inputs
- Effectiveness: ensuring that the outputs deliver the desired outcome
- Equity: ensuring that the benefits are distributed fairly

Biscaye et al. 2015 noted that many development organisations are unclear on the definition and measurement of cost performance and described this important measure as comparing project cost by budget, outputs or outcomes. The international development organisations often refer to either approved budget or Cost Estimate (CE) to describe the original planned project budget. In this research the term Cost Estimate (CE) is used to describe the original planned project budget and projected cost (Ahsan & Paul, 2018). Baccarini and Love (2014) define cost performance by the difference between the final project cost and the approved budget, or CE, and this definition is used in this research. The World Bank Group (WBG) and other Multilateral Development Banks (MDBs) include similar concepts for VFM in the New Procurement Framework reform initiative. The WBG defines VFM principles and its associated terms, which are used today as follows (World Bank Group, 2013, pp. 2,3; 2016, pp. 7,8; 2022):

- Value for Money: the principle of value for money means the effective, efficient and economical use of resources, which requires an evaluation of relevant costs and benefits, along with an assessment of risks, and non-price attributes and life cycle costs, as appropriate. Price alone may not necessarily represent value for money.
- Economy: the principle of Economy takes into consideration factors such as sustainability, quality, and non-price attributes and life cycle costs as appropriate, that support value for money. It integrates economic, environmental and social considerations into the procurement process that the WBG has agreed with the borrower. It also permits augmenting identified sustainability criteria with specific criteria in support of the borrower's own sustainable procurement policy.
- Integrity: the principle of integrity refers to the use of funds, resources, assets and authority according to the intended purposes and in a manner that is well

informed, aligned with the public interest, and aligned with the broader principles of good governance.

- **Fit for Purpose:** the principle of fitness for purpose applies both to intended outcomes and the procurement arrangements in determining the most appropriate approach to meet the project development objectives and outcomes considering the context and the risk, value and complexity of the procurement.
- **Efficiency:** the principle of efficiency requires that procurement processes be proportional to the value and risks of the underlying project activities. Procurement arrangements are generally time-sensitive and strive to avoid delays.
- **Transparency:** the principle of transparency requires that the borrower and the WBG enable appropriate review of the procurement activities, supported by the appropriate documentation and disclosure. Transparency requires; (i) that relevant procurement information is made publicly available to all interested parties, consistently and in a timely manner, through readily accessible and widely available resources at reasonable or no cost; (ii) appropriate reporting of procurement activities; and (iii) the use of confidentiality conditions in contracts only where justified.
- **Fairness:** the principle of fairness refers to (i) equal opportunity and treatment for bidders and consultants; (ii) equitable distribution of rights and obligations between borrowers and suppliers, bidders, consultants and contractors; and (iii) credible mechanisms for addressing procurement-related complaints and providing recourse. Open competitive procurement is the WBG preferred procurement approach, whenever possible, to maximise fairness of opportunity to bid. Whenever possible, the WBG requires the eligible individual and firms to be given the same opportunities to compete for WBG-financed activities.

ADB defines VFM with three parameters facilitating transparency (efficiency, quality and flexibility) and has developed an insightful figure for the concept shown in Figure 3.

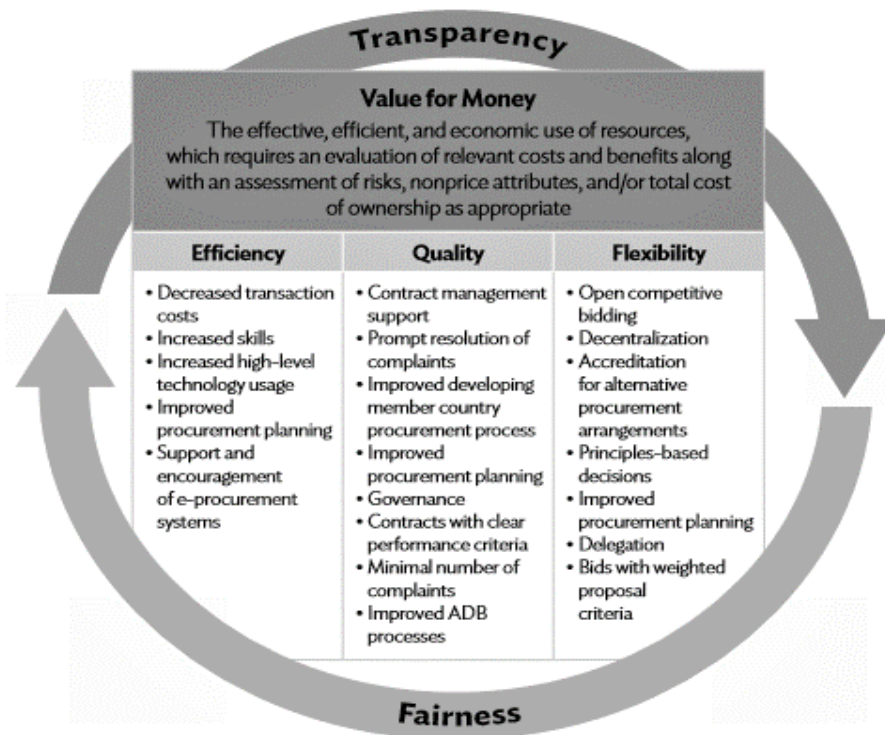


FIGURE 3. VALUE FOR MONEY DEFINED BY THE ASIAN DEVELOPMENT BANK (Asian Development Bank, 2019, p. vi).

The Chartered Institute of Procurement and Supply (CIPS) describes VFM as often situational and dependent on a variety of factors (Chartered Institute of Procurement and Supply CIPS, 2020). The CIPS definition of VFM involves evaluating the extent to which the proposed solutions will achieve the desired outcomes and reconciling those benefits with the total lifetime costs of realising those benefits. Typical factors that are taken into account in defining VFM include fitness for purpose, quality, total lifetime costs, risk, environmental and sustainability issues, and a variety of other factors relating to the contribution of solutions to the organisation’s overall goals. As corporate social responsibility is translated into congruent procurement processes, the definition of value for money needs to take into account a broad range of criteria (CIPS Australia, 2020).



The various definitions have many common principles and themes, and as VFM is applied to the procurement and project management processes, they provide essential definitions that are relevant for measurement of how money is spent (Easterly & Pfitze, 2008). The VFM emphasis and perspective helps to focus on factors that influence how and where money is spent in procurement and supply chain management (SCM) (Barr & Christie, 2015; Fleming, 2013; Lapede, 2000).

Common procurement assessment models, developed in the private sector, have a focus on efficiency, responsiveness and flexible procurement methods and therefore metrics are more aligned with the definitions of both DE, as discussed in the previous section, and VFM (Jackson, 2012; Arlbjørn & Freytag, 2012). Procurement assessment models used in international development, such as the OECD's Methodology for Assessing Procurement Systems (MAPS) and the WBG Alternative Procurement Arrangements (APA), have a primary emphasis on policy, governance, regulations and compliance aspects of procurement (Tadelis, 2012; World Bank Group, 2016).

Procurement performance measurement and the selection of applicable key performance indicators can improve the transparency and accountability for the significant proportion of development funds managed through procurement. Best practices from the public and private sectors for performance management of procurement can help provide insight into how funds are disbursed and improve DE and VFM in international development (Easterly & Pfitze, 2008; Tadelis, 2012).

The attention towards DE has led to increased interest in the concepts of cost performance and value as part of evaluation and measurement for development results (Khagram et al., 2009). The Global Fund and Department for International Development (DFID) and other members of the working group on VFM in global health, identify four

areas for the achievement of more health for the money (Glassman et al., 2013). These areas include; resources allocation, contracts, performance verification and costs (Department for International Development (DFID), 2011). The working group, rather lengthily, defines VFM as “creating and complying with rules or procedures for allocating resources that elicit the production and use of the health-maximising mix of services for the available donor, national and private resources” (Brugha et al., 2004; Glassman et al., 2013, p. 13). The history and emergence of VFM is often considered as a relatively recent event, but in reality, the concept has developed over some time as part of the evolving definition of cost performance (Barr & Christie, 2015). The Independent Commission for Aid Impact (ICAI) has developed a logical framework for the measurement of aid effectiveness and defines VFM using the 4 Es when considering overseas aid (ICAI, 2011). The measurement of VFM is an important aspect for procurement programmes and projects in international development, as it offers solutions, methods and answers for much of the criticism directed towards DE. VFM has six main methods for measurement, which include; cost performance, cost-utility analysis, cost-benefit analysis, social return on investment, rank correlation and basic efficiency resource analysis (Fleming, 2013). As with DE, challenges over VFM concern the current lack of tangible measurement of output level metrics and the relationship with outcomes, social and macroeconomic aspects of development (Fleming, 2013). The World Bank Group, OECD, DFID and other leading international development organisations have adopted compatible definitions and methods for VFM and have included the concept in their respective procurement procedures and guidelines. According to the OECD, there are three description levels of relevance for VFM; the global level, country program level and the project level (Jackson, 2012). The definition and metrics used in procurement best practices for VFM provide transparency and clarity

for funds disbursed through the procurement process (Yuan, Zeng, Skibniewski, & Li, 2009). Lessons learned from private sector procurement practices have led to the gradual adoption of more flexible procurement methods, measurement of value and efficiency in the public sector and development organisations (Roodhooft & Van den Abbeele, 2006).

For example, the WBG procurement reform procedures include VFM, and adopt other more modern and flexible methods derived from private sector practices, in the WBG New Procurement Framework (Renard & Lister, 2015; World Bank Group, 2016).

To conclude, the overarching theme from these definitions is that there are inconsistencies in the definitions of both DE and VFM as applied to international development. The inconsistency and overlap of terms introduces difficulty in the measurement and understanding project success and value in development projects.

Now that key terms of this research (Development Effectiveness and Value for Money) have been defined from multiple perspectives, the next section considers the different perspectives of the private sector when it comes to effective project performance.

### **3.5 PRIVATE SECTOR DEFINITIONS AND APPROACHES TO PROJECT EFFECTIVENESS**

International development organisations and the private sector are disparate in their application and measurement of project effectiveness. The OECD evaluation methodology, definitions and criteria are used by many international development organisations to measure development effectiveness for projects, as outlined in Section 3.3. Much of the more recent literature on the private sector's Critical Success Factors (CSFs) has concerned information technology projects (Iriarte & Bayona, 2020; Trigo & Varajão, 2020). CSFs have remained remarkably constant across time and industry with similar factors used in several industries to measure project effectiveness thereby substantiating ubiquitous acceptance (Elhaniash & Stevovic, 2016; Westerveld, 2003).

Shrnhur et al. (1997) proposed that project success in the private sector and industry is divided into four dimensions from project efficiency, customer satisfaction, business success and finally long-term considerations over time (Shrnhur, Levy, & Dvir, 1997). Biscaye et al. (2015) described four levels of the Hierarchy of Results Measurement and represented project result measurement at the bottom level. Lim and Mohammed (1999) classify projects into two categories describing them as macro and micro levels. The macro viewpoint is taken from a broad perspective at the level of the vision, goal and outcome. The micro viewpoint includes cost, time and quality and the two different viewpoints are often measured by different stakeholders and through a variety of indicators and metrics, as discussed in Section 3.5 and Section 3.10. The micro viewpoint includes cost, time and quality, which Atkinson (1999) referred to as the “iron triangle”. Other logical framework models distinguish project success into different concepts of project management and product purposes and success (Baccarini, 1999). Jugdev and Müller (2005) reviewed the development of definitions for project success over 40 years and suggested that measures of success should be set early in the project cycle and include indicators of efficiency and effectiveness. More recently, Bergmann and Karwowski (2018) described the literature on the evolution of modern project management and definitions of project success, Critical Success Factors (CSFs) and agile project management over the last decade. The authors identified six project success factors as management, process, project, organisational, people and technical factors. Frefer et al. (2018) maintained that the iron triangle Critical Success Factors (CSF) of cost, time and quality are key to project success. Moradi et al. (2020) mapped the evolution of project success research from 1992 to 2018 and identified the top five project success criteria as cost, time, quality, business success and technical performance. Picciotto (2020) argued for a new approach to project management in

development to include better definition of inputs, outputs and outcomes, adopt sensible evaluation criteria and combine self-evaluation with independent evaluation. The authors confirm the importance and relevance of the iron triangle CSFs across multiple industries and also recommend additional CSFs which will be further examined in this research. Projects that have been considered unsuccessful at the micro level due to overruns or delays, are ultimately considered successful from other perspectives, such as community impact over the passage of time (Shokri-Ghasabeh & Kavouosi-Chabok, 2009). For example, the Sydney Opera House has been considered successful due to its iconic architecture, reputation and long-term profitability despite large overruns in cost and time (Burke & Macdonald, 2014). From a project management perspective, there are commercial examples such as the Gorgon, Australia natural gas project that declared a budget overrun by over AUS\$21 billion from an original budget estimate of AUS\$37 billion in 2009 to AUS\$52 billion in 2012 (Olaniran, et al., 2015). The project experienced delays and significant increases in costs, nevertheless it is considered successful, due to long-term profitability and factors, such as favourable oil prices and positive economic growth (Johnson & Babu, 2018; Olaniran et al., 2017).

This paradox can occur the other way around and, as Jabeen (2016) noted, despite the success of individual projects, unintended influences, such as design flaws, politics, economics and external environmental factors can cause negative development impacts. The definition of the nature and level of measurement of project success can be complicated. It is therefore important to establish the definition and criteria for success at both the macro and micro levels of performance measurement for international development projects in a similar way to the private sector (Kerzner, 2017).

In the example of major infrastructure projects, CSFs include extensive planning and cost estimation calculated through detailed design and extensive financial studies,

environmental and social factors and contract management skills (El-Reedy, 2016). The focus on the iron triangle and CSFs in private sector projects includes project-specific performance metrics contained in the specific project scope, specifications, cost estimate, schedule and deliverables. Recognition of the importance of cost management has led to further analysis of the causes of overruns and underruns for private sector projects (Johnson, Leenders, & McCue, 2017). In private sector construction projects, reasons for poor performance have typically included inaccurate cost estimation, poor procurement and ineffective contract management (Bhargava, Anastasopoulos, Labi, Sinha, & Mannering, 2010; Pollock, Price, & Player, 2007). More recently, Johnson and Babu (2018) identified 30 causes of project time overruns and 20 causes of project cost overruns and these include variations from original design, unreliable project schedules and inaccurate cost estimation (Johnson & Babu, 2018). To promote project success, Frefer et al. (2019) identified nine common groups of CSFs to measure performance which include cost, time and quality factors plus factors such as customer satisfaction. Flyvbjerg et al. (2018) observed that for 258 transportation infrastructure projects in the United States, 86% experienced overruns (Flyvbjerg et al., 2018). In a similar way, Love et al. (2018) observed that out of a sample of 1,093 water infrastructure projects in the United Kingdom, there were 657 overruns and 436 underruns (approximately 60% and 40%) and only one project was delivered exactly on budget, or the Cost Estimate (CE). There are clearly differences by industry and yet current knowledge appears to be aligned and suggests that cost and time present challenges for project management and that there are applicable Critical Success Factors (CSFs) to mitigate such challenges. Araújo et al. (2017) argued that procurement is a crucial process and that the effective and efficient performance of every phase of the procurement process is makes a vital contribution to overall project performance and success (de Araújo et al., 2017;

Sundqvist, Backlund, & Chronéer, 2014). The procurement process plays an important role in the selection of cost-effective suppliers. The process evaluates specifications and scope for the selection of the most capable and competitive contractors for the implementation and delivery of the goods, works and services (Mir & Pinnington, 2014). The private sector commits investment in procurement and project management resources, leadership and process performance and this contributes to effective process management and successful project delivery (Mensah, 2016; Manyathi et al., 2021).

Under the OECD project evaluation methodology, the definition and distinction of macro and micro-levels are often unclear and inconsistent some project evaluations refer to macro indicators and others to micro indicators or a hybrid of both (Lamhauge, Lanzi, & Agrawala, 2012). For project evaluations with no clear, defined and identified micro indicators, or iron triangle CSFs, it is difficult to understand how donors' and taxpayers' money has been spent. The combination of the micro and macro perspectives suggest that there is at least two important levels of metrics for the measurement of project success (Frefer et al., 2018). In response to the question of how funds are disbursed, the CSFs, including the iron triangle and other criteria, may potentially provide effective performance indicators to measure project success and VFM (Vrchota et al., 2020).

This section provides some examples of the private sector approach to the understanding of project performance and Value for Money (VFM) in projects by the adoption and measurement of Critical Success Factors (CSFs) and the clear definition of success.

The private sector approach to the measurement of project performance provides examples of proven indicators to define project success, measure project performance and answer value-based questions on how effectively development funds are disbursed.

### 3.6 CRITICISM OF DEVELOPMENT EFFECTIVENESS; WHERE DOES THE MONEY GO?

Now that it is clear about the quantities of funding going through international development organisations, the research examines the levels of the hierarchy of performance measurement in [Figure 2](#), to understand whether development is effective. Easterly (2008) queried the value of development aid by asking the question of “where does the money go?” (Easterly & Pfutze, 2008). Much of the criticism and attention regarding DE is concentrated on perceived weaknesses regarding the disbursement of funds, transparency and delivery of results at the bottom two output and spend activity levels (Easterly & Pfutze, 2008). The difficulty in meaningful measurement of DE at the top two effectiveness and outcome levels is often attributed to the influence of external, global, regional and local factors. These may include economic growth and trade, competition, political factors, corruption and duplication from other donor organisations working in the same sector and geographic territory (Hansen & Tarp, 2000; Okafor & Udibe, 2021). International development organisations themselves came under scrutiny for uneven responses to commitment towards DE made at the Paris Declaration on Aid Effectiveness (2005), general lack of progress, weak compliance and poor alignment of commitments and goals (Brown, 2020, Gulrajani, 2014; Wood et al., 2011).

Other criticism includes claims of inconsistent evaluation methods, poor connectivity between impacts, outputs, outcomes and differences in the way that international organisations approach development evaluation (Denizer et al., 2013; Mitchell, 2019). There is variation in the approach and prioritisation for measurement of DE, and there are significant challenges with understanding the transparency and accuracy of data from the evaluation models (Bourguignon & Sundberg, 2007). International organisations approach performance measurement of DE at different levels of hierarchy; effectiveness, outcomes, outputs, and spend activity levels (Biscaye et al., 2015).



The relationship between outcomes and outputs is important, for example, the output measurement for a school built on time, to quality standards and budget, does not necessarily inform whether more children will be educated or reflect on the quality of future education outcomes (Riddell, 2009). Conversely, from the outcome measurement perspective in health; if a patient has not died, there is often a lack of evidence that aid inflows are making the patient better (Doucouliagos & Paldam, 2009; Mosley, 1986).

In the private sector, major capital or operations projects, such as infrastructure or information systems, have long been measured and monitored from start to completion by adherence to budget, schedule, quality performance factors (Jha & Iyer, 2007; Moradi et al., 2020). The combination of these performance metrics for individual projects combines to indicate the success or failure of the specific project and cumulatively towards the overall outcomes (Yuan et al., 2009). In the context of international development organisations, the connectivity between microeconomic and macroeconomic levels is often opaque, and there are contradictory results in DE performance metrics (Hansen & Tarp, 2000). As approximately half of ODA is estimated spent on goods and services through the procurement process, procurement performance metrics should serve as critical indicators for DE and VFM performance (Ellmers, 2011).

Best practices in procurement performance provide the opportunity to demonstrate the improved implementation of DE, VFM and reduce funds lost to corruption in the disbursement funds for development programs (Asiedu et al., 2021; Schiele, 2007).

### **3.7 CORRUPTION AND WASTE**

McMullan (1961) pointed out that corruption by definition is not only illegal but has many detrimental effects on societies, countries and government including political instability and repression in addition to mistrust of government and discouragement of

investment and enterprise (McMullan, 1961). In addition, and relevant to the research, inclusion of inefficiency and waste from corruption and poor practices negatively influences project success and the distribution of funds (Harnois & Gagnon, 2021).

Van Roy (1970) provided a working definition of corruption as “the use of power, preferment, or prestige, or for the benefit of a group or class, in a way that constitutes a breach of law of the standards of high moral conduct” (Van Roy, 1970, p. 86).

The former Secretary General of the United Nations, Ban Ki-moon is attributed with the statement that 30% of aid “failed to reach its final destination” owing to corruption (Ravelo, 2012). Corruption is difficult to measure precisely, and this percentage would represent approximately US\$4.8 billion of the US\$16 billion annual expenditure of the United Nations agencies. According to WBG estimates in 2010, \$20 billion to \$40 billion is stolen each year from Low-Income-Countries (LICs) (Baker, 2005; World Bank Group, 2010). The OECD had long argued that the loss of these funds, under the oversight of international development organisations, serves to embed corrupt practices further and sustain poverty (OECD, 2014). These figures may not include funds lost to other factors such as poor design, bad loans, waste, inefficiency and the loss and leakage is potentially even greater (Bardhan, 1997; Harnois & Gagnon, 2021; Kuipers, 2021). International development organisations spend considerable resources, time and effort in the discussion on transparency, fairness and competition to reduce the incidence of corruption (Arrowsmith, 2010). The MDBs and UN agencies have comprehensive policies on that govern the debarment of suppliers that are proven to have contravened the policies of the organisation (Seiler & Madir, 2012). The WBG has defined corruption, fraud, coercion, collusion and obstruction as practices that may lead to sanctions (World Bank Group, 2020). While these definitions are generally recognised

and subscribed to by other MDBs, they vary between organisations and the WBG definitions are defined as follows:

- *A corrupt practice* is the offering, giving, receiving or soliciting, directly or indirectly, of anything of value to influence the actions of another party improperly.
- *A fraudulent practice* is any act or omission, including a misrepresentation, that knowingly or recklessly misleads, or attempts to mislead, a party to obtain a financial or other benefit or to avoid an obligation.
- *A collusive practice* is an arrangement between two or more parties designed to achieve an improper purpose, including influencing the actions of another party improperly.
- *A coercive practice* is impairing or harming, or threatening to impair or harm, directly or indirectly, any party or the property of the party to influence the actions of a party improperly.
- *An obstructive practice* is: (i) deliberately destroying, falsifying, altering or concealing of evidence material to investigation or making false statements to investigators in order to materially impede a Bank investigation into allegations of a corrupt, fraudulent, coercive or collusive practice; and/or threatening, harassing or intimidating any party to prevent it from disclosing its knowledge of matters relevant to investigation or from pursuing the investigation; or (ii) acts intended to materially impede the exercise of Bank's inspection and audit rights.

(World Bank Group, 2022, pp. 1, 2)

The WBG provides insights on fraud and corruption in its financed projects and identifies procurement itself and specific areas of the procurement process where corruption may occur (World Bank Group, 2022).

The WBG identifies nine specific areas where procurement can be detected as it refers to red flags and recognised that corruption could take place during the procurement process and post-contract award when the project is under implementation. Debarment terms can vary from one to five years with or without conditions and serves as a deterrent for suppliers and contractors from corruption and breaches of policy (Seiler & Madir, 2012).

The United Nations' anti-fraud and anti-corruption framework for the United Nations' Secretariat defines fraud and corruption as follows:

“Fraudulent Acts include both fraud and corruption. Fraud encompasses any act or omission whereby an individual or entity knowingly misrepresents or conceals a material fact, to obtain an undue benefit or advantage for himself, herself, itself or a third party, or to cause another to act to his or her detriment. Corruption encompasses any act or omission that misuses official authority or that seeks to influence the misuse of official authority, in order to obtain an undue benefit for oneself or a third party” (World Bank Group, 2022, p. 1).

Corruption is assumed to have many negative effects on international development however, in some countries, it is argued that it is used to “grease the wheels” and may facilitate business transactions (Wei, 1999). Some of the effects of corruption in government, civil society, corporations, and community have far-reaching consequences for development and economic growth. As mentioned earlier in this section, these include injustice, inefficiency, mistrust of government, waste of public resources, discouragement of enterprise, political instability, lack of trust, repressive measures and restriction of government policy (McMullan, 1961). In many developing countries, where the institutions of government are generally weak, it is generally perceived that there is a relatively high incidence of corruption (Bardhan, 1997; Kolstad, Fritz, & O’Neil, 2008). However, Paulo Mauro (1995) found that there was a significant negative

association between corruption and the country rate of growth it can be assumed that corruption exists and has different of extremes and causes in every country and society (Mauro, 1995). The amount and nature of corruption can be the difference between success and failure of MDG goals or any specific project. There are differences between the level and scale of corruption that have a profound effect on DE and project success. In terms of levels of corruption, De Graaf (2007) lists six descriptions and attempts to explain the causes as follows:

1. Public choice theory; an individual makes a rational decision that leads to a pre-determined outcome.
2. The “bad apple” theory; individual corrupt agents as the cause of corruption.
3. Organisational culture; culture and structure of the organisation within which the agent (s) is working.
4. Clash of moral values; the distinction and potential clash between value and norms of society and the values and norms of the individual.
5. Public administration lapses; organisational integrity and the effect on public officials and individuals.
6. Correlation theories of multiple factors; the cause of corruption is correlated with all levels.

(De Graaf, 2007, p. 45)

Corruption can take place both in the procurement process in the selection of supplier or contractor and project management process in the implementation and delivery phases. In public choice theory, for example, an individual may decide to introduce bias into the procurement process to unfairly influence a particular bid. The objective may be to compensate for low wages in LICs and gain money from suppliers and contractors as a

means to supplement and compensate for poor incomes (Ampratwum, 2008). The corrupt activity can take the form of fraudulent or collusive behaviour where groups of suppliers can collude to determine a winner for a procurement selection process and share benefits at the award stage or during the project management stages (Ika, 2012). In contrast, the “bad apple” theory considers the possibility that there are individuals, groups and even entire organisations that engage in corrupt behaviour (Muzio, Faulconbridge, Gabbioneta, & Greenwood, 2016). This behaviour can leach out into wider behaviour and even become embedded into the culture of the organisation. Similarly, the organisation or government department may have a culture that engages in corrupt practices that becomes routine or normalised (Al-Jundi, Shuhaiber, & Al-Emara, 2019). In a similar way, public administration of a given country or region may have lapses in governance and compliance and there can be clashes between the values and norms of the society and of the individual. Such clashes can work both ways and bad apples can influence the organisation and conversely organisations can potentially influence the individual to engage in practices previously considered unacceptable (Ashforth & Anand, 2003).

A common tactic is to pressure, marginalise, demote and fire individuals who do not comply with corrupt activities of the organisation. Once engaged as a participant, the individual is implicated and thereby, socialised into future corrupt activity (Al-Jundi et al., 2019; Pinto, Leana, & Pil, 2008).

In terms of the scale and measurement of corruption, there are many indicators of corruption that use different criteria to assess corruption in any specific country. The most widely known example of these is the Transparency Corruption Index, which draws data from 13 data sources and captures the assessment of experts on corrupt behaviours in the public sector and active mechanisms for prevention of corruption (Transparency

International, 2021). The behaviours include bribery, diversion of public funds, use of public office for private gain, nepotism and state capture and the Corruption Perceptions Index (CPI) is expressed as a score in rank order by country. For the purposes of benchmarking, the CPI is standardised so that all of the sources are converted and expressed on a scale of 1-100, and 0 represents the highest level of corruption and 100 is the least corrupt (Transparency International, 2021). The indices have different methods of data collection and combine of quantitative or qualitative information however, they show high levels of correlation with each other (Lederman, Loayza, & Soares, 2005).

The Transparency CPI is widely utilised as a guide to relative levels of corruption and for strategic and business decision making. The argument for having a single indicator of corruption is understandable; however, the difficulty with arriving at an accurate and reliable index has proved challenging. The Transparency International (TI) aggregated data sources of the CPI each have their strengths and weaknesses, which can cause imbalances in the composite index (Kaufmann, Kraay, & Mastruzzi, 2011). The sources also differ in terms of scale, definition and type of corruption as there are so many forms of corruption (Misangyi, Weaver, & Elms, 2008). The aggregation from different sources of the CPI and other indices such as the World Governance Indicators (WGI) of the World Bank Group have experienced variation in measurement year on year. As a result, annual variations in the CPI and WGI may cause distortions in measurement from year to year (Rohwer, 2009). Transparency and corruption indicators serve as useful metrics, but caution is required before applying policy based on results, ranking or changes in time. The indices, such as the Worldwide Governance Indicators (WGI), are helpful as guidance and inputs alongside other measurements but should be used as output indicators in themselves (Harnois & Gagnon, 2021). Transparency International (2004) lists several political leaders and the proportion of public funds that they have embezzled

over recent history (Transparency International, 2004). The analysis shows that alarming numbers of political leaders are responsible for the embezzlement of substantial quantities of money and would undoubtedly have a significant adverse effect on public programs and development, particularly if the origin of some of the money is from international development organisations and donors (Baker, 2005). There are numerous examples of the “resource curse” of Angola where the oil and diamond industry have been marked by persistent cronyism and nepotism (Williams & Isaksen, 2016). In the example of Mozambique, over US\$4.9 billion is estimated lost between 2004 and 2014 (Department for International Development, 2015). Such cases are not necessarily restricted to developing countries and have been documented in the UK where corruption also occurs, although perhaps on a smaller scale (Murray, 2014). The argument has been made that corruption somehow increases the speed of business exchange and by implication improve efficiency for small “grease the wheel” transactions and to explain a means to unlock obstacles put in place by corrupt public officials (Wei, 1999). In public procurement, however, bribery not only encourages the appointment of poorly qualified bidders with weak capability, subsequent delivery, integrity, and effectiveness are all likely to be compromised. The corrupt practices also increase the costs to recover bribe amounts and further costs due to inefficiency, resulting in contract amendments and corrections, waste and rework (Bardhan, 1997).

In terms of procurement more specifically, fraud and corruption are considered widespread, and the World Bank Group (WBG) has estimated that approximately US\$1.5 trillion in public contract awards is influenced by corruption (Paterson, Changwony, & Miller, 2019). In addition, the volume of bribes exchanging hands for public sector procurement is estimated at US\$200 billion per annum (Kaufman, 2005).



Fraud and corruption creates major obstacles for organisations to achieve their objectives, and procurement-related corruption, particularly, is a serious issue in LIC and MIC countries (Raymond, 2008). As procurement involves the identification, qualification, selection and management of suppliers and contractors involved in the disbursement of development funds, it is a key process where corruption can occur and provides the opportunity for it to be prevented. There are some areas of procurement and project implementation that have been recognised as being vulnerable to corruption, fraud, and inefficiency, which can happen when due process is not followed correctly. Fraud and corruption in procurement may include bribes, kickbacks to cover price, quality, conflict of interest, substitutions, collusion and other examples. Fraud and corruption are distinct but can occur simultaneously and can be confused with each other. Fraudulent practices can also be difficult to detect and can involve internal and external parties or both. Such practices include bid collusion of contractors in the bidding process, false representation, conflict of interest and comprised evaluation process (Anderson & Katz, 1998; Matthew, Patrick, & Denise, 2013). While fraud can potentially take place at all points of the procurement process, the Integrity Vice Presidency of the World Bank Group (INT) identifies 12 specific areas where there is an increased risk. These are listed under the four phases of: planning, bidding process, evaluation and contract management. They include procurement planning, advertising, bidding documents, supplier shortlisting and qualifications, pre-bid conference, bid submission and opening, bid evaluation, evaluation report, contract drafting, delivery and contract changes (Integrity Vice Presidency (INT), 2020). The twelve areas may serve to help systematic reviews and audits but would not always catch all potential red flags or prevent incidents which are heavily influenced by the motivational and environmental factors of the organisation (Matthew et al., 2013). The criticism of Development

Effectiveness (DE), includes the effects of corruption, inefficiency and waste on project delivery (Ferry, Hafner-Burton, & Schneider, 2020). The next section discusses the reaction and commitment to improving DE and the disbursement of development funds.

### **3.8 COMMITMENTS TOWARDS IMPROVEMENTS IN DEVELOPMENT EFFECTIVENESS**

The Millennium Goals were established in 2000 by the United Nations with the participation of world leaders and multinational organisations and established eight comprehensive, time-limited and measurable targets to beat extreme poverty, known as the Millennium Development Goals (MDGs) (Fukuda-parr et al., 2013; Stijns, 2012).

The MDGs set high-level ideals, goals, and targets to be achieved by 2015 and unified a very broad range of organisations with the common aim of global objectives. The MDGs' ambitions include; eradicate extreme poverty; universal primary education; promote gender equality and empower women; reduce child mortality; improve maternal health; combat HIV/Aids, malaria and other diseases; environmental sustainability and global partnership (Stijns, 2012). The nature of the goals themselves appear ambitious as objectives for effective development and by their very nature, require careful measurement and monitoring across countries, sectors and across time. The definitions of the MDGs can also make it difficult to attribute the achievement of development success or failure and distinguish results from free market other causes (Powell, 2005).

The driver of the meteoric rise of Asian countries such as Thailand, Taiwan, Malaysia, Indonesia, Vietnam and Cambodia in the 1990s, until the subsequent Asian financial crisis, raises questions of how much was contributed by international development (Suryahadi, Hadiwidjaja, & Sumarto, 2012). Later the rapid and continued economic growth of China dwarfed the equivalent development and growth of other countries which received proportionately a far greater level of development funds (Sachs, 2012).

The MDBs were replaced by the Sustainable Development Goals (SDGs) in September 2015 under the resolution adopted by the General Assembly Transforming our world: the 2030 Agenda for Sustainable Development (United Nations, 2015). The SDGs are intended as a plan of action for people, planet and prosperity and to strengthen universal peace. There are 17 SDGs and 169 targets intended to build on the MDGs and expand to deliver what was not completed. The term sustainable has economic and environmental implications for long term durability of development towards people, planet, prosperity, peace and partnership (Hák, Janoušková, & Moldan, 2016). However, criticism of the SDG indicators concerns the difficulty of measurement of sustainability in general and inconsistency of the indicators themselves (Janoušková, Hák, & Moldan, 2018).

As outlined by the United Nations (2015), the SDGs are as follows:

- Goal 1: End poverty in all its forms everywhere
- Goal 2: End hunger, achieve food security and improved nutrition and promote sustainable agriculture
- Goal 3: Ensure healthy lives and promote well-being for all at all ages
- Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all
- Goal 5: Achieve gender equality and empower all women and girls
- Goal 6: Ensure availability and sustainable management of water and sanitation for all
- Goal 7: Ensure access to affordable, reliable and sustainable and modern energy for all
- Goal 8: Promote sustained, promotable and sustainable economic growth, full and productive employment and decent work for all

- Goal 9: Build resilient infrastructure, promote inclusive and sustainable industrialisation and foster innovation
- Goal 10: Reduce inequality within and among countries
- Goal 11: Make cities and human settlements inclusive, safe, resilient and sustainable
- Goal 12: Ensure sustainable consumption and production patterns
- Goal 13: Take urgent action to combat climate change
- Goal 14: Conserve and sustainably use the oceans, seas and marine resources for sustainable development
- Goal 15: Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss
- Goal 16: Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels
- Goal 17: Strengthen the means of implementation and revitalise the Global Partnership for Sustainable Development

(United Nations, 2015, p. 14)

The words sustained, sustainable or sustainably are used 14 times in the descriptions of the 17 SDGs and in SDG 8, 14 and 15, are mentioned twice. The challenge of simple descriptions and the difficulty of measurement is immediately apparent when considering the SDGs that include the broad and far-reaching indicators in common with the Universal Declaration of Human Rights. Such SDGs include the responsibility for all states to respect, protect and promote fundamental freedoms for all without distinction of

any kind as to race, colour, sex, language, religion, political or other opinions (Fukuda-Parr & McNeill, 2015). However, it is notable that SDB 17 recognises the need for strategic planning, implementation and effective reporting by including the commitment for collaboration, combined with the goal to strengthen the means of implementation (United Nations, 2015). As part of the SDGs, there is increasing demand and expectation for funds intended for environmental projects and this includes the unimaginable figure of US\$5.7 trillion anticipated future investment in green infrastructure (Ackom & Motty, 2020). This incorporates US\$100 billion by 2020 of annual financing for multilateral organisations, such as the Global Climate Fund (GCF) and US\$17 billion in grants, spent since 1991, by the Global Environmental Facility (GEF) (Cui & Huang, 2018).

For the measurement of the SDGs, there are 229 proposed indicators associated with the 169 targets of these 149 have consensus with the Inter-Agency and Expert Group (IAEG) and the remaining 80 do not have general agreement on the measures. There are five general categories of the indicators, distributed differently and can be categorised under: People, 40.6%; Money, 26.2%; Plans and Policies, 16.6%; Production and Consumption, 8.7% and Planet, 7.9% (United Nations, 2015). The indicators are mostly measurements aimed at the global, regional, country or sector level and are macroeconomic indicators making it possible to set targets and monitor progress (Atkisson, 2016). Pizzi et al (2020) carried out a systematic review of articles on the SDGs and technological innovation, business firm's contribution in developing countries, non-financial reporting and education (Pizzi, Caputo, Corvino, & Venturelli, 2020). The role of public procurement has been recognised as a means to improve sustainable practices towards the SDGs in the light of challenges experienced in pandemics and emergencies (UNCTAD, 2020).

Hak et al. (2016) noted that there are challenges with the measurement of DE when development goals are expressed at such a macro-level and that they can be ambiguous

(Hák et al., 2016). Some approaches are measured with a top-down strategy with the assumption that a “rising tide lifts all boats”, while others take a more bottom-up approach (Elalfy, 2021). There is also a substantial difference between estimation of expenditure from a top-down perspective compared to a bottom-up form of estimation of expenditure and programme and project costs. Top-down cost estimation is often calculated more broadly in the form of macro indicators based on an economic change or time-bound estimates with the expectation that detailed expenditure is worked out at a later stage. On the other hand, bottom-up measurement depends on a better understanding of the specific fund disbursement and breakdown of project scope and specifications and hence provides more detail on expenditure of funds (Stijns, 2012).

The SDG indicators do not effectively measure the micro-level and answer the stakeholders’ questions on how money is spent and whether development funds are spent effectively at the project level (Atkisson, 2016; United Nations, 2015).

International development organisations are increasingly concerned with demonstrating DE and Management for Development Results (MfDR) concepts partly due to demands from donors and from public interest on how funds are disbursed (Renard & Lister, 2015; World Bank, 2009). The Paris Declaration on Aid Effectiveness, 2005 and the Accra Agenda for Action, 2008 made commitments to ensure DE include the potential expanded use of procurement systems of borrowing countries. The importance of transparency and disclosure of information was added as a new commitments in the High-Level Forum for Aid Effectiveness in Busan, 2011 (Ellmers, 2011).

Questions over the disbursement of funds and over effective allocation of aid development have partly driven this interest in DE (Easterly & Pfutze, 2008; Gulrajani, 2015; Woods & Narlikar, 2001). The United Nations Conference on Financing for

Development held in Monterrey, Mexico, in March 2002 made commitments and set targets to improve DE and accountability (Stern, 2003). The conference resulted in 87 objectives, including the reform of aid practices and transformation of performance measurement and evaluation for results towards the Millennium Development Goals (MDGs) (Picciotto, 2002). These objectives have been accompanied by the adoption of MfDR and other performance-based management concepts developed from lessons from practices from the private-sector and various industries (Meier, 2003).

Different development organisations have taken alternative approaches to improve DE, have different definitions and have experienced varying degrees of progress and success (Gulrajani, 2014). International organisations such as the OECD and WBG are generally optimistic in their self-evaluation reporting and, while this may be partly motivated by self-interest, they acknowledge that measurement is involved and that aid disbursement has numerous political and administrative challenges (Collier & Dollar, 2004; Riddell, 2009). Despite the commitments and initiatives around DE, criticism persists over the lack of clarity between measurement directed at the macro, or outcomes, level and reports focussed on micro-output levels of cost performance, efficiency and VFM considerations (Bourguignon & Sundberg, 2007; Holzapfel, 2016; Mosley, 1986).

In addition, even after the commitments were made, the indicators for performance measurement vary between MDBs, UN agencies and bilateral organisations and the data remains poor and inconsistent (Biscaye et al., 2017). Measurement of outcomes requires longer periods of observation, a good understanding of outputs and clarity of the allocation of all donors and agencies operating in the same region and sector (Mitchell, 2019; White, 2005). McKee et al. (2020) argued that improvement in transparency and measurement of development was a priority for the commitments made at the Paris Declaration on Aid Effectiveness, 2005 (Wood, Kabell, Muwanga, & Sagasti, 2008).

This section has dealt with the current measurement and commitment towards improvements in delivery of effective development goals and outcomes. Against this background, the next section goes on to discuss challenges with the measurement of the SDGs and DE and VFM at different levels of evaluation.

### **3.9 MEASUREMENT OF DEVELOPMENT EFFECTIVENESS**

Now that the commitments towards improvement of DE have been discussed, it is important to determine the level, definition and measurement of DE to establish if the development goals of donors and development organisations have been effective or not. International development organisations have concerned themselves with the issue of DE through multiple studies and increasing demand for clarity from donors, stakeholders and the general public. There are multiple financial instruments for delivery of development which further differ depending on the type of development organisation (United Nations, 2019; World Bank Group, 2017). Different financial forms of development include loans, grants, technical assistance, programmes and projects. Development is directed towards multiple sectors including health, transportation, agriculture, infrastructure, sustainable projects, governance and education. The effectiveness of development goals and results can be influenced by the selection of the type of financial instrument, project or programme design and delivery mechanism (Asian Development Bank, 2019; Brech & Potrafke, 2014; World Bank Group, 2020).

International development organisations, such as Multilateral Development Banks (MDBs), often rely on macro attributes, or characteristics such as the basic description and value of the financial instruments to identify and monitor disbursement activities. However, it is necessary to rigorously monitor and evaluate progress effectively to understand development effectiveness and to make key decisions on future disbursement objectives (Kilby, 2009). The macro viewpoint of development evaluation are important



indicators for the performance of development objectives and are often used to determine future allocation of aid and disbursement of funds (Bourguignon & Sundberg, 2007).

The implementation of development activities are performed by a combination of development organisations, Implementing Partners (IPs) and other parties depending on the capacity of the recipient country government and departments.

The case for increasing aid disbursement to countries with good performance and sound economic policies led to improved growth in MICs and LICs was made as part of the justification for DE during the United Nations Conference on Financing for Development in Monterrey, Mexico in March 2002 (Burnside & Dollar, 2000). The conference led to further discussion on DE at the macro and micro-level and emergence of early debate and development of both DE and VFM concepts (Hansen & Tarp, 2000; Jackson, 2012).

The focus on macro indicators for disbursement to determine future allocation of aid, rather than the results and feedback on the ground, may be as the beneficiaries have little voice with local governments and much less influence with donors from High-Income Country (HIC) that control the aid agencies (McGillivray & Morrissey, 2000). The macro attributes are useful to identify the characteristics of the development instrument, but the measurement of detailed indicators and metrics are equally essential to understand the performance and effectiveness of the disbursement of development funds.

Easterly (2003) noted the WBG statement in its 1998 report on assessing aid; that despite resources committed to development goals, agencies primarily saw themselves as “being in the business of dishing out money” (Easterly, 2003; Milliband & Gurumurthy, 2015).

At the micro-level, project reports provide greater detail of expenditure, project progress and results, but many evaluation reports do not specifically measure performance of on the basis of cost and expenditure (Clements et al., 2008). Mir and Pinnington (2014) cite

several studies that argue that quality, planning and cost estimation of the strategic planning and implementation phases of development projects are a significant factor in the performance and success of the project and outcomes (Mir & Pinnington, 2014). The concept of VFM is increasingly being applied to development activity to measure DE and link inputs and costs with outputs and outcomes (Jackson, 2012).

Evaluation groups, such as IEG at the WBG, the Independent Evaluation Department (IED) at ADB and the Evaluation Department (EvD) of DFID include criteria for the measurement of the cost performance and disbursement of funds (Asian Development Bank, 2019; Department for International Development (DFID), 2016; IEG, 2014). In the private sector, by comparison, cost performance and the measurement of fiscal performance is frequently characterised as Return on Investment (ROI) amongst other financial metrics (Barr & Christie, 2015). While ROI is not always applicable to donor funds, the MDBs often utilise cost-benefit analysis and Economic Internal Rate of Return (EIRR) (Van Toan, Hà, & Chau, 2012).

However, there are challenges in the definition of financial returns and benefits related to many development projects which can lead to confusion of the calculation of returns and inconsistent measures of EIRR (Dixon, Carpenter, Fallon, Sherman, & Manipomoke, 2013). For example, in the case of the MDBs, the payment of loans by a borrowing government may not necessarily mean that the programme or project, or its component parts, were completed successfully or effectively (Ahsan & Gunawan, 2010).

Such measurement for financial benefit or EIRR in the case of some sectors such as education and health, for example, can often be difficult to determine and hence other forms of cost-benefit analysis is necessary (Esdadafal, 2014). In contrast, the private sector usually has a comparatively clear definition of financial return such as revenue generation, Net Present Value (NPV) calculations, toll payment or oil price (Kerzner,

2017). Nevertheless, the private sector process is structured to ensure that funds used for implementation are deployed with rigorous process controls and that costs are contained for project completion towards effective and efficient results (Moradi et al., 2020; Westerveld, 2003). In international development, in cases where EIRR is more difficult to determine, it is possible to quantify the volume and value of goods, equipment and services delivered according to the specifications, scope of services and the completion of services and works. In addition, it is possible to monitor funds committed and expensed, adherence to schedule and project management quality and standards (de Souza & Gomes, 2015). In the case of funds sourced from donors and public funds, there is less opportunity to recover losses than in the private sector through increased sales or production. As mentioned earlier in this section, development funds are largely received from donors and ultimately from taxpayers. Gulrajani (2014) argued that due to these factors there is an increasing importance, and perhaps obligation, to measure every dollar disbursed on behalf of donors and taxpayers (McKee et al., 2020).

### **3.10 EVALUATION OF DEVELOPMENT PERFORMANCE**

The independent evaluation process of international organisations follows procedures and guidelines that are based on principles outlined in the 1991 Organisation for Economic Co-operation and Development – Development Administrative Committee (OECD-DAC) for evaluation of development assistance (OECD, 2019). The principles incorporate standards and practices endorsed by the Evaluation Co-operation Group (ECG) of MDBs and other development organisations (Asian Development Bank, 2019).

The evaluation for development results framework used by the OECD’s Development Assistance Committee (OECD-DAC) is the primary evaluation standard for the evaluation of results. The international development organisations largely follow OECD-DAC recommended criteria, albeit with some minor variations (OECD, 2018). However,

project evaluations that use the OECD-DAC criteria often lack detailed analysis on factors such as procurement activity, cost performance and project schedules (Lamhauge et al., 2012). Project evaluation utilises common criteria for the overall evaluation of outcomes, and this measurement is taken as the indicator for project performance assessment. The OECD (2019) adjusted the criteria used to evaluate outcome ratings for several years to comprise six key criteria defined as relevance, coherence, effectiveness, efficiency, impact and sustainability (Morra-Imas & Rist, 2009; OECD, 2019).

The evaluation departments of international development organisations prepare several types of development evaluation reports from different perspectives (Asian Development Bank, 2019; European Bank for Reconstruction and Development, 2011; IEG, 2021).

The evaluation guidelines outline the procedure for evaluating completed public sector operation projects supported by loans, grants and technical assistance and each of the organisations has developed a specific framework in line with its own requirements.

The evaluation assessments use criteria including relevance, coherence, effectiveness, efficiency, impact and sustainability with differences between organisations (Biscaye et al., 2017; OECD, 2019). The OECD framework for defining of effectiveness is very broad and does not clearly distinguish between micro and macro-levels of measurement outputs and outcomes as discussed in Section 3.9, and shown in Figure 2 (Clements et al., 2008; OECD, 2019). The OECD definition of efficiency refers to inputs and outputs but does not include cost estimation and VFM (OECD-DAC, 2019). Private sector definitions and interpretations for effectiveness and efficiency are more detailed than development definitions in terms of measurement project performance, cost performance and delivery of outputs and outcomes (Fleming, 2013; Jackson, 2012; Kakwezi & Nyeko, 2019). The definitions of effectiveness and efficiency include both macro and micro measurements, and the relatively low incorporation of cost and quality data within

these criteria, may explain some of the difficulty in finding evidence for where funds are disbursed (OECD, 2019). As development evaluations are widely considered as the source of truth for the measure of success of DE the inconsistency does not provide visibility on cost performance (Crawford & Bryce, 2003; World Bank Group IEG, 2016).

WBG evaluation methodology, for example, includes different types of reports, including thematic and sectoral evaluations across projects, country programme specific and country cluster evaluations looking at country or country cluster performance against objectives. Project-specific performance evaluations include Project Completion Reports (PCRs) and Implementation and Completion Results (ICRs) validated using a set of evidence including literature reviews and site visits (Barr & Christie, 2015; IEG, 2013).

The development organisations follow evaluation methods based on OECD models, although they have variations in approach and process. The titles for the plethora of different reports, can be confusing and difficult to cross-reference; a brief example is provided in Table 9, which illustrates the types of different evaluation reports generated by two of the MDBs. The broad nature of the definitions of evaluation criteria and the different viewpoint in the measurement of project success and DE further add to lack of clarity project evaluation measurement and VFM (Frefer et al., 2018).

Despite the availability of reports, in seeking the answer to better measure DE, there are different schools of thought concerning the measurement at the country level macro-level and project-specific project, or iron triangle, level of measurement (Picciotto, 2020).

TABLE 9. EVALUATION REPORTS WORLD BANK AND ASIAN DEVELOPMENT BANK

World Bank Group (WBG)	Asian Development Bank (ADB)
Annual Corporate Thematic Reports	Annual Corporate Thematic Reports
Country Programme Evaluations (CPE)	Country Assistance Programme Evaluations (CAPE)
Cluster Country Programme Evaluations (CCPE)	Country Partnership Strategy Final Review (CPSFR)
Validation of Completion and Learning Reviews (CLRRs)	Sector Assistance Programme Evaluations (SAPE)
Implementation and Completion Results Reports Reviews (ICRRs)	Project completion Report Validation (PVR)
Implementation and Completion Results Reports (ICRs)	Project Completions Reports (PCR) Technical Assistance Completion Reports (TCR)
Project Performance Assessment Reports (PPARs)	Project Performance Evaluation Reports (PPER)

(Asian Development Bank, 2019; IEG, 2012)

These reports use similar rating and assessment methods using a four, five or six-level rating system consolidated to arrive at a score derived from the component evaluation criteria. The WBG rating scale is defined in Table 10 for the WBG IEG evaluations.

TABLE 10. INDEPENDENT EVALUATION GROUP RATING SCALE

Rating	Description
<b>Highly Satisfactory</b>	There were <b>no</b> shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
<b>Satisfactory</b>	There were <b>minor</b> shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
<b>Moderately Satisfactory</b>	There were <b>moderate</b> shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
<b>Moderately Unsatisfactory</b>	There were <b>significant</b> shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
<b>Unsatisfactory</b>	There were <b>major</b> shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
<b>Highly Unsatisfactory</b>	There were <b>severe</b> shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.

(World Bank Group, 2015, p. 14)

The descriptions for ratings are subjective and broad and, while the guidelines provide information on how to consolidate different ratings for an overall score, the process can be subject to positive bias, variable and inconsistent (Denizer et al., 2013; World Bank Group, 2015). To address this, the ADB converts the four-point scale evaluation ratings

to numerical scores and weights the evaluation criteria, which helps provide more objectivity and comparability for reporting purposes (Asian Development Bank, 2019).

The measurement of development performance at the country level is developed from data collated from aggregate aid activity and project performance expressed in the form of ratings (Howes et al., 2011). More than 100 papers have been published since the 1960s that address the issue of aid effectiveness; these have differences in the level of measurement (Doucouliagos & Paldam, 2008). Measurement at the project level is widely practised and effective projects lead to successful delivery of corporate and organisational objectives (Cooke-Davies, 2002; Frefer et al., 2018). However, there are minor differences in the criteria definitions and rating scales between different organisations and this research provides a new approach to comparison of project ratings.

It has long been recognised that there are differences in performance evaluation results between the macroeconomic data and microeconomic data. Mosely (1986) noted that there are cases where microeconomic indicators show that projects are successful, while the macroeconomic indicators are more discouraging (McKee et al., 2020; Mosley, 1986). For example, Roodman (2007) explained that it is easily observable to tell if a road is paved, but it is more difficult, if not impossible, to determine if the road infrastructure, raises total output per capita for a national economy (Roodman, 2007).

Macroeconomic metrics and indicators on development-funded projects are instruments of choice for many development organisations, and yet poor performance remains widespread. The WBG found that over 50% of projects were unsuccessful in Africa and up to 2,000 projects with similar statistics for IFC (Ika et al., 2012). Further to this, the IEG claimed that 39% of WBG projects were unsuccessful or less than satisfactory in 2010 (Chauvet et al., 2010).

The concept of ranking multilateral and bilateral donors according to performance in aid quality greatly depends on the consistency and quality of the sources of data (Easterly & Pfütze, 2008; Mosley, 1986; Roodman, 2006). The definition and interpretation of performance measurement for aid quality differs greatly depending on the perspective of studies, and various studies measure different criteria (Birdsall, Kharas, Mahgoub, & Perakis, 2010; Biscaye et al., 2017). The recommendations of the Paris Declaration for Aid Effectiveness, 2005 list selectivity, alignment, harmonisation and specialisation as key measures for success towards aid effectiveness and DE objectives (Lonsdale, 2016; McKee et al., 2020). The recommendations assume that improved aid quality can be achieved by targeting aid, alignment with country policy, coordination with other donors and specialisation of development by geography and sector (Knack & Rahman, 2007). The measurement parameters recommended by the Paris Declaration for Aid Effectiveness reflects the development priorities, objectives and macro design aspects of the donor policy. However the evaluation criteria do not always provide detailed measurement of cost, time and quality in projects (Nielsen, 2010; Roodman, 2006).

The previous section reviews the evaluation and measurement methods for development activity used in international development organisations. These methods are used to assess whether development goals, programmes and projects are effective or otherwise and hence serve as the measurement of DE. The following section focuses on procurement performance measurement as distinct from the measurement of overall development goals, programmes and projects (Papadaki et al., 2014). The procurement process is an essential mechanism for the selection of suppliers and contractors for goods, services and works in development projects and hence contributes towards project performance and DE (de Araújo et al., 2017; Ellmers, 2011; Sundqvist et al., 2014).



### 3.11 MEASUREMENT OF PROCUREMENT PERFORMANCE

Despite the effort of Low Income Countries (LICs), Middle Income Countries (MICs) and partners like the World Bank Group to improve procurement performance, public procurement still experiences poor delivery, quality, cost and compliance issues (Rönnbäck, 2012). Public procurement performance has been the subject of attention from governments, donors, practitioners, academics and project managers for many years (Ahsan & Paul, 2018). In 2004 the European Institute of Purchasing Management (EIPM) held a conference dedicated to the subject. The EIPM conference recognised that, in many public sector, performance indicators are based on financial indicators and compliance rather than efficiency, value and strategic value more common in private sector industry (Kakwezi & Nyeko, 2019; KPMG, 2012). Financial measures sometimes ignore market dynamics, complexity and technical and commercial innovation of goods, services and equipment (Lardenoije, van Raaij, & van Weele, 2005). It is possible to have financial measures and metrics that indicate successful outcomes and yet projects and deliverables may still not be completed satisfactorily on the ground. Challenges to projects may be caused by numerous other factors such as managerial efficiency, benchmarking and the measurement of project performance (Iyer & Banerjee, 2016). Successful delivery of development goals and major projects in the private sector is the result of effective delivery, or a summary of the parts, of all the discrete project and procurement elements of the overall development objective (Kerzner, 2017). The definition, design, schedule and cost estimation for each project serve as the standards and indicators for the monitoring of cost, schedule and quality for overall project performance and objectives (Cooke-Davies, 2002; Henchie, 2007; Tadelis, 2012). The success of individual project outputs requires quality procurement, project management, planning, design and implementation (Howes et al., 2011; Mosley, 1986). In the same

way as with the private sector, the success of development projects, with substantial procurement content, likewise, depends on effective procurement and project management process expertise and performance (Assaf & Al-Hejji, 2006).

In pursuit of the answer to the questions on how money is spent, the effective implementation of discrete project outputs must be achieved on schedule, adhering to budget estimates and according to design, specification and scope. To achieve effective implementation, outcomes and DE, it is important that overall development objectives and the component projects have been well designed, planned and coordinated (Picciotto, 2020). For example, improvements in primary education for both genders can only be achieved if the schools are built, education materials delivered, teachers are trained and there is adequate attendance. In the health sector, for example, malaria and other diseases can be reduced and eradicated when the medicines, mosquito nets, medical facilities, medical and health staff and essential services are in place (Howes et al., 2011). Without the fundamental inputs of any development project, overall success is almost impossible.

It is therefore possible to have positive outcome indicators for overall development evaluation indicators and yet the implementation, and delivery of individual projects can be unsuccessful and ineffective. Conversely, it is also possible to have negative overall outcome indicators and yet equipment and services may be delivered effectively (Ahsan & Gunawan, 2010). DE is frequently measured at the macro outcomes level with insufficient connectivity and traceability to the specific projects and activities that make up the development goals (Chianca, 2008). In addition, distortions to the measurement of development can occur due to overlapping territory and activities of different international organisations, economic, political and project factors that influence the measurement indicators and evaluations (Khagram et al., 2009).

The procurement process captures the record and measurement of quality of goods and services, implementation milestones and the amount and time to complete the process (Lindstrom, 2014). This is particularly important given that a substantial proportion of development funds are ultimately sourced from donor contributions, government and public taxes. In response to the question introduced in Section 3.1, on how funds are disbursed, procurement indicators and principles provide the opportunity for data capture and measurement of what is procured, its delivery and impact on the project concerned (Chianca, 2008; Escadafal, 2014). The procurement process is comprised of several distinct phases and relies on data collection and collaboration with multiple internal and external stakeholders and Subject Matter Experts (SMEs) (Knudsen, 1999; Lardenoije et al., 2005). The structured procurement process, procedures and guidelines facilitate collection of quality and time performance measurement documentation and data. The requirement for obtaining cost estimates, commitments and payments in the contract monitoring process enables tracking and measurement of expenditure and how funds are disbursed. The procurement process therefore provides the opportunity to capture cost, time and quality and scope metrics and indicators and potentially answer the question raised by Easterly (2008); “where does the money go” (Easterly & Pfutze, 2008).

In measuring project performance, the critical concept of efficiency is often represented by “doing things right” whereas effectiveness is “doing the right thing” (CIPS Australia, 2020; Kakwezi & Nyeko, 2019). This is a comprehensive description, and other major international organisations such as DFID define efficiency as; achieving outputs for inputs, while bearing in mind quality and defines effectiveness as; achieving programme outcomes, while bearing in mind equity (Fleming, 2013). However, neither of these definitions appear to do full justice to the key question of how development funds are spent or provide a reliable measurement of the concept of Value for Money (VFM).

The logic of such questions may be complemented by the private sector approach to efficiency, which combines the concepts of design, budget, cost and schedule (De Wit, 1988). In the private sector approach, effectiveness is often connected to customer satisfaction, quality, productivity and ROI (Sundqvist et al., 2014).

Procurement has a critical role in organisational governance and is expected to ensure supplier qualification, standards and quality and safeguard the organisation from fraud and corruption. Secondly, in the natural resources industry and other sectors procurement is required to safeguard the organisations from risks associated with Health, Safety and Environment (HSE) and Corporate Social Responsibility (CSR) including aspects, such as the Modern Slavery Act 2015, United Kingdom (International Labor Organisation (ILO), 2018; LeBaron & Rühmkorf, 2017; McCrudden, 2007).

Studies have shown that the preferred measure of procurement performance in industry is quality followed by cost, innovation, reliability and responsiveness (Knudsen, 1999; Kumar, Ozdamar, & Peng Ng, 2005). The performance measures in the international development industry for procurement performance are more aligned with the regulatory environment, process compliance and governance than efficiency and effectiveness as defined in the private sector (OECD, 2010; World Bank Group IEG, 2016).

The procurement process in both public and private sectors, include many similar data points such as design, budget and schedule, which are compatible with the evaluation criteria used in development evaluations, such as relevance, effectiveness, efficiency and sustainability (Schiele, 2007; World Bank Group, 2015). The major international development organisations utilise the OECD-MAPS measurement system and variation of the system, and there are several popular private sector PMA models (Lockamy III & McCormack, 2004; OECD, 2010). Procurement metrics and indicators offer an

opportunity to improve transparency and visibility of the disbursement of development funds and the measurement of DE and VFM both at the project and outcomes levels.

Current PMAs such as the OECD-MAPS, and WBG APA, claim to measure value for money, economy, efficiency, integrity, fit for purpose and transparency and fairness (OECD, 2015). However, as discussed in Sections 3.5 and Section 3.6, and considering the broad definition and subjectivity of the measurement of these concepts, they fail to provide detail of cost performance and value. Private sector PMAs, on the other hand, are geared towards the measurement of value and efficiency (Safari et al., 2021). Hence, there is a strong argument for an amended PMA for the assessment of development organisations to incorporate other procurement criteria. To introduce a method to measure VFM for procurement, it is possible to apply weightings for VFM to a modified PMA model, and use a numerical scoring method to arrive at a more objective and comparable measure of cost and VFM in development projects (Chianca, 2008). The procurement capacity and performance of international development organisations is essential for the delivery of successful projects and DE. A modified PMA, in this way, can potentially provide a more effective means to measure both DE and VFM. Such a model can also lead to improved transparency and more effective implementation for the estimated 50% of ODA volume disbursed using the procurement process (Ellmers, 2011). The following section describes existing PMA models used in the private sector and the development industry and discusses current knowledge on common models.

### **3.12 PROCUREMENT MATURITY ASSESSMENT MODELS**

Due to the requirement for transparency, fairness and competition in procurement and the requirements for internal and public auditable records, the procurement function manages and maintains detailed policies, procedures and guidelines (Arrowsmith, 2010).

International and public sector organisations have requirements for strong regulatory and

compliance aspects of the procurement process and procedures partly due to public accountability and national regulations (Odhiambo & Kamau, 2003). Private sector policy and procedures equally, have a requirement for governance and transparency and also have a pragmatic focus on efficiency, quality and supplier performance management (Arlbjørn & Freytag, 2012; Tadelis, 2012). The requirements for fairness, transparency, compliance and oversight in the procurement function across all sectors, also mandates strong governance and encourages effective performance measurement. Procurement processes, procedures and strategic guidelines provide a good source of comparative information through which the procurement function of any organisation can be measured and compared (Brandmeier & Rupp, 2010). Considering Kaplan's (2009) principle; what is not measured cannot be improved, baselines, standards and clear metrics are required for measurement, benchmarking and continuous process improvement (Kaplan, 2009). For performance measurement, procurement maturity assessment models have emerged over several years, and these have been developed and amalgamated from different project management, supply chain and procurement disciplines (Estampe, Lamouri, Paris, & Brahim-Djelloul, 2013). Rendon (2008) uses the term contract management to refer more broadly to the procurement process which can potentially cause confusion (Rendon, 2008). For clarification, contract management is a term more frequently used in both public and private sectors, specifically for the post-award phase of the procurement process. In addition, the definition and application of concepts such as supply chain management, procurement, logistics and contract management have evolved and varied in different organisations, sectors and industry. In the public sector, the term Procurement is used to refer to the entire procurement process, whereas private sector organisations are more familiar and use SCM to refer procurement, contracting and logistics as a whole (Giunipero, Hooker, Joseph-Mathews,

Yoon, & Brudvig, 2008). As a result, PMAs have differences depending on whether they are oriented to the procurement process alone, or more widely, to Supply Chain Management (SCM) (Aulia & Isvara, 2021; Estampe et al., 2013; Lahti et al., 2009).

Process maturity models emerged as part of the rapid development of total quality management (TQM) and other similar quality management assessment programs. The concept of process maturity and benchmarking resulted from the requirement to define standards, stages and levels of best practices for process efficiency and effectiveness. An early maturity model is the Capability Maturity Model Integration (CMMI). It was developed by the Software Engineering Institute (SEI) for engineering efficiency and effectiveness (Chrissis, Konrad, & Shrum, 2003). The CMMI model is concerned with quality-based process assessment; the process description and the abbreviations are remarkably similar to the CMMM model. The CMMI model is considered to be one of the best-known maturity models and designed on the basis of the description of process and performance required to achieve a higher level of maturity (Estampe et al., 2013).

There are a number of procurement and supply chain maturity models that have been developed from business process orientation (BPO) models that assess process management, measurement and controls (Lockamy III & McCormack, 2004). The business process maturity concept measures steps in the process or project lifecycle and has become widely used in project management performance and benchmarking. The procurement or SCM process is closely related to the earlier BPO and project models, but has a greater focus on procurement, supplier selection, evaluation criteria, regulatory compliance and expenditure analysis and reporting (Collier & Evans, 2020).

There are a limited number of procurement maturity assessment models used in international development, and many are developed from the OECD – MAPS assessment

model with shared characteristics (Lloyd et al., 2014). The procurement assessment models based on OECD-MAPS methodology have strong regulatory, governance and compliance focus and less emphasis on the VFM concepts of economy, efficiency and effectiveness (Barr & Christie, 2015; Netland, Alfnes, & Fauske, 2007; Picciotto, 2020). There are, however, variations in the models used by international development organisations, some follow the OECD-MAPS model very closely, and others have adapted the model for their specific requirements, Table 11.

The main difference concerns the ranking and scoring process, and in some cases, different scales including three, four, five and six-point scales are used as adapted by the evaluation group concerned. The development organisations and the name of the specific PMA with levels of maturity of the organisations are listed in Table 11 (Asian Development Bank, 2022; OECD, 2010; United Nations International Children's Emergency Fund, 2018; World Bank Group, 2016). Secondly, the models in Table 11 are primarily used to assess government department level procurement capacity and infrastructure of developing countries and are geared mainly at the policy, legal and regulatory framework level and not towards performance assessment.

The concept of maturity has developed in both procurement and project management literature and procurement maturity is described according to one, two, three, four and five levels, as shown in Table 11, Table 12 and Table 13 (Estampe et al., 2013; Lockamy III & McCormack, 2004; Schiele, 2007).

This research builds on current knowledge outlined in the literature review and has adapted and developed a new maturity model shown in Figure 4 and the Integrated Procurement Maturity Model (IPMM), represented by the procurement process diagram shown in Figure 5, as applicable to the research objectives.



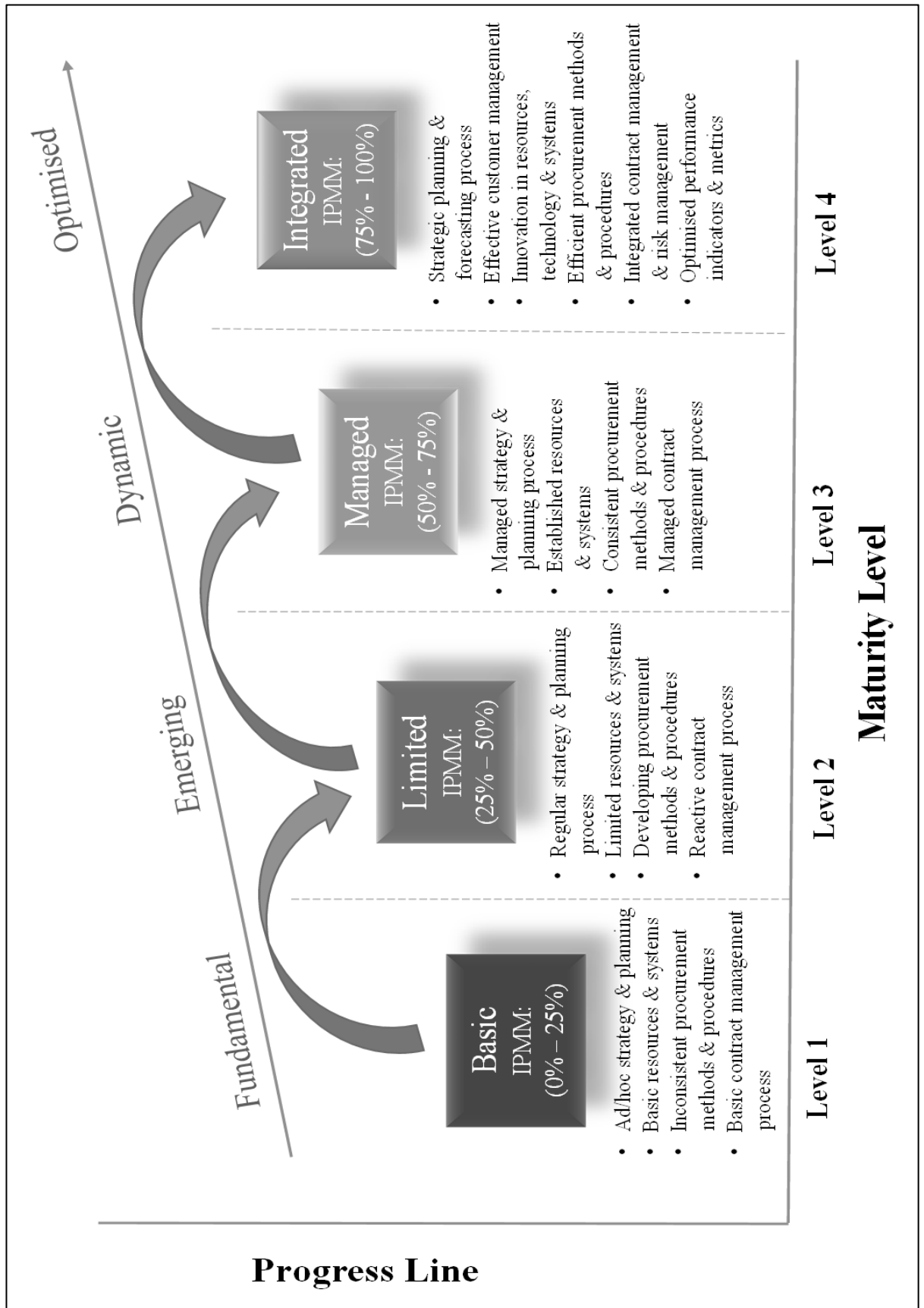


FIGURE 4. PROCUREMENT MATURITY LEVELS

TABLE 11. PROCUREMENT ASSESSMENT TOOLS USED IN DEVELOPMENT ORGANISATIONS

Abbr.	Organisation	Name of PMA	Levels*
OECD	Organisation for Economic Co-operation and Development (OECD, 2010)	Organisation for Economic Co-operation and Development Methodology for Assessing Procurement Systems (OECD-MAPS)	4
WBG	World Bank Group (WBG) (World Bank Group, 2016)	Alternative Procurement Arrangements (APA)	4
ADB	Asian Development Bank (ADB) (Asian Development Bank, 2022)	Alternative Procurement Arrangements (APA)	4
IADB	Inter-American Development Bank (IADB) (Inter-American Development Bank, 2018)	Acceptance of the Use of Country Procurement Systems	4
EBRD	European Bank for Reconstruction and Development (EBRD) (European Bank for Reconstruction and Development, 2012)	Public Procurement Assessment	4
AFDB	African Development Bank (AFDB) (African Development Bank, 2015)	OECD-MAPS	4
UNDP	United Nations Development Programme (UNDP) (United Nations Development Program, 2010)	Public Procurement Capacity Assessment	4
JUI	United Nations Joint Inspections Unit (JUI) (Joint Inspections Unit of the United Nations System, 2015)	Contract Management – Procurement Assessment Model (CM-PAM)	4
JICA	Japanese International Co-operation Agency (JICA) (JICA, 2013)	Study of National Procurement System (Crown Agents)	NA
UNICEF	UNICEF (United Nations International Children's Emergency Fund, 2018)	Procurement Capacity Assessment	3
UNOPS	United Nations Office for Project Services UNOPS (United Nations Office for Project Services, 2018)	Procurement Efficiency Assessment Tool (PEAT)	4
DFID	Department for International Development, DFID UK (DFID, 2007; OECD, 2015)	Procurement Capability Review Model and Standards Framework version 1.0	4
DFAT	DFAT (OECD, 2015)	Procurement Assessment Diagnostic Tool	4

\* Note: The levels are the levels of maturity from one to three, one to four or one to five

A substantial body of literature has developed on process maturity models with different procurement criteria and maturity levels; many of these listed in Table 11, Table 12 and Table 13. In some examples, the models propose inter-organisational supply chain relationships in addition to internal integration factors and multi-chain, or societal process factors, as the highest level of performance (Pache & Spalanzani, 2007).

In Table 12, the name of the author or group that has developed the model, is listed next to the name and source document (s). Procurement maturity can be defined as the level of performance effectiveness, efficiency and organisation of a procurement department, compared to best practices, as shown in Figure 4 (Bloch, 2011; Schiele, 2007; Van Weele, Rozemeijer, & Rietveld, 1998).

TABLE 12. ARTICLES ON PROCUREMENT MATURITY ASSESSMENT MODELS

<b>Name of PMA</b>	<b>Source document(s)</b>	<b>Level</b>
Best in Class Maturity Framework	(Aberdeen Group, 2006)	3
Levels of Procurement Development	(Anderson & Katz, 1998)	4
Purchasing Performance Excellence	(Schreiber Bernd, 2018)	5
A Development Model for Effective MRP Procurement	(Barry, Cavinato, Green, & Young, 1996)	3
Purchasing Excellence	(Roland Berger, 2014)	3
Business Process Re-Orientation Model	(Lockamy III & McCormack, 2004)	5
BSC: Balanced Scorecard	(Kaplan & Norton, 2001)	NA
CAPS Research	Benchmarking (CAPS Research, 2010, 2012, 2014)	NA
Fitting Purchasing to the Strategic Firm: Frameworks, processes and values	(Cavinato, 1999; Freeman & Cavinato, 1990)	5
Contract Management Maturity Model (CMMM)	(Rendon, 2008)	5
Purchasing Function Maturity	(Cousins, Lawson, & Squire, 2006)	4
GSCF Framework	(Croxtton, Garcia-Dastugue, Lambert, & Rogers, 2001)	3
PSM Drivers and Firm Performance	(Hartmann, Kerkfeld, & Henke, 2012)	3
Stages of Procurement Maturity	(Cammish & Keough, 1991; Keough, 1993)	5
Stages of Procurement Sophistication	(Kraljic, 1983)	4
Levels of Strategic Purchasing	(Paulraj, Chen, & Flynn, 2006)	3
Purchasing Integrated Maturity Model	(Potage, 2017)	5
Strategic Stages in Purchasing	(Reck & Long, 1988)	4
SASC; Strategic Audit Supply Chain	(Gilmour, 1999)	4
SCALE: Supply Chain Advisor Level	(Bossu et al., 2004)	4
SPM: Strategic Profit Model	(Stapleton, Hanna, Yagla, Johnson, & Markussen, 2002)	NA
Testing the procurement-performance link	(Schiele, 2007)	4
Assessment of Excellence in Procurement	(Schuh et al., 2017)	2
Management-oriented purchasing analysis	(Schweiger, 2015)	4
Purchasing and Supply Development Model	(A. J. Van Weele, 2010)	4
Purchasing Empowerment	(Voegele & Schwientek, 2002)	3
WCL: World Class Logistics Model	(Chandes et al., 2003)	4

Estampe et al. (2013) presents 16 different supply chain performance measurement models, and the paper describes the concept of process maturity and tabulates and compares the origin, type, attributes and indicators of each model.

The public sector procurement maturity models have a greater emphasis on matters of governance and process compliance private sector models have a greater emphasis on measurement of quality, efficiency, cost performance and delivery schedules (Tadelis, 2012). There is a great variation in the procurement criteria between the assessment models as they are adapted to the specific sector, industry and purpose of assessment (de Souza & Gomes, 2015; Ibbs & Kwak, 2000; Wendler, 2012).

There are several commonly used PMA models which have emerged both from academic and commercial demand for procurement assessment reporting and are used in the private and public sectors and some of these are shown in Table 13.

TABLE 13. COMMON PROCUREMENT MATURITY ASSESSMENT MODELS

<b>Abbreviation</b>	<b>Name of PMA</b>	<b>Reference</b>	<b>Levels</b>
APQC	American Productivity and Quality Center (APQC)	Article (PMA)(American Productivity and Quality Center APQC, 2018)	5
CEB	Corporate Executive Board (CEB)	(Gartner, 2018)	4
EFQM	EFQM: Excellence Model	(EFQM, 2018; Rönnbäck, 2012; Samardžija & Kralj, 2010)	3
EVALOG	Global EVALOG (Global MMOG/LE)	(AIAG, 2010; Estampe et al., 2013; Odette, 2020)	3
KPMG	KPMG PMA	(KPMG, 2012)	5
IPMM	Implement Procurement Maturity Model	(Implement Consulting Group, 2018)	5
SCOR	SCORmark SCC	(APICS, 2010; Bolstorff, 2001)	4
ROSMA	The Purchasing Chessboard (ROSMA)	(Schuh, Kromoser, Strohmer, Pérez, & Triplat, 2009)	4
360°	360° Procurement Performance Analysis	(Horvarth & Partners, 2018)	4

SCOR itself is a popular PMA model for supply chain applications and has four levels of maturity classification that include internal integration, functional integration, external integration and inter-company collaboration. Many models have adapted the earlier PMA models to include, or exclude, process areas to address research objectives and differences of the industry of concern (Poluha, 2007). Authors compare and review different PMA models and observe common criteria with overlap and duplication between different PMA models (Safari et al., 2021; Schiele, 2007; Schweiger, 2015).

Many of the leading models such as APQC, ROSMA and SCOR have been further adapted and modified and are widely adopted by major consulting organisations such as KPMG, PricewaterhouseCoopers and A.T. Kearney (Huang, Sheoran, & Keskar, 2005). Organisations in the public and private sector use PMA models for procurement assessment and benchmarking purposes and these incorporate a wide range of procurement criteria depending on the industry and requirements (American Productivity and Quality Center APQC, 2018; Kearney, 2016; KPMG, 2012; Poluha, 2007).

The key procurement phases also differ between different organisations, industries and the private and public sector. The SCOR model identifies four important process steps that are relevant between different organisations in the industry but need some adaptation for use in service industries and public organisations which do not make or manufacture products. However, the relative simplicity and relevance of the key process areas to a wide range of industries, make the model popular and widely used in conjunction with other models such as balanced scorecards (BSC) (Kaplan & Norton, 2001; Min, Thakkar, Kanda, & Deshmukh, 2009). Procurement maturity assessment models are helpful for benchmarking against similar organisations for comparative analysis. In addition, they set standards for best practices, define procurement criteria and advance organisational capability, procurement maturity and performance (Lockamy III & McCormack, 2004).

Current research on procurement maturity assessment models has been limited to the comparison of models designed for specific industries in the private sector rather than more broadly applied to international development (Van Looy, De Backer, Poels, & Snoeck, 2013). The functions responsible for the evaluation of development practices for development, such as the Independent Evaluation Group (IEG) of the WBG, state the importance of the measurement of VFM and risk in public procurement. However, they do not include detailed procurement indicators for VFM and risk factors as significant components in the evaluation process, criteria and reports (Fjeldstad & Isaksen, 2008). Similarly, while the OECD-MAPS procurement assessment models, are designed to assess the procurement process, they do not include metrics for VFM (OECD, 2010). The increased emphasis of procurement process metrics, including VFM, in the evaluation for development results, can improve transparency and performance of management and implementation of DE (Mena et al., 2014; Wood et al., 2011).

The previous section reviews the role of the procurement and project management processes in international development and the concept of measurement and performance in relation to DE and VFM. Based on the knowledge considered in this chapter, the following section introduces a preliminary theoretical framework to highlight the relationship between the procurement and project management processes and the measurement of project performance in international development.

The comparison and cross reference of all the procurement maturity assessment models reviewed in Table 11, Table 12 and Table 13, led to the development of the IPMM maturity model, Figure 4, and procurement process diagram, Figure 5, including 12 key procurement criteria under four procurement pillars, or phases, indicated as follows:

1) procurement planning and strategy phase, 2) organisation and resources phase, 3) regulatory framework phase and 4) contract management phase.

The 12 key procurement criteria identified in this research are illustrated in the procurement process diagram in Figure 5 below.

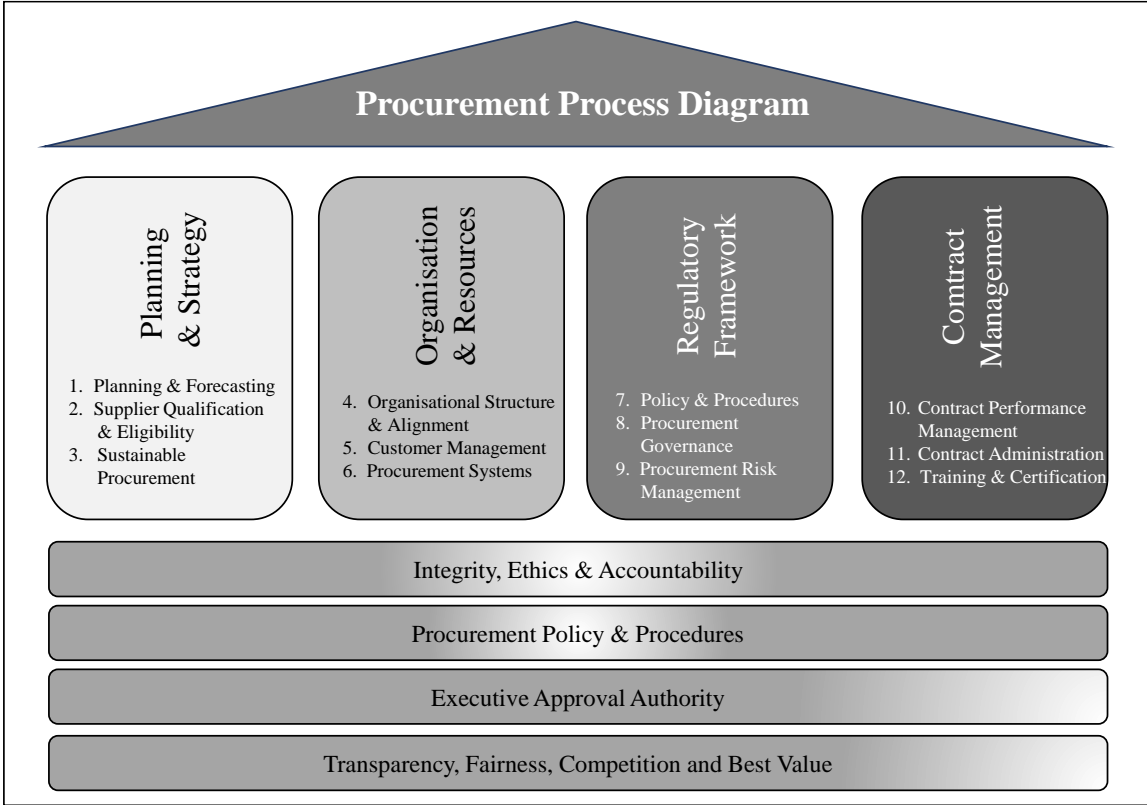


FIGURE 5. PROCUREMENT IPMM PROCESS DIAGRAM

The Planning and Strategy phase includes all procurement areas that take place early in the procurement process including planning, preparation of budget and strategic inputs to the procurement process. The Organisation and Resources and the Regulatory Framework phases include all elements of the process and structure required for the implementation of the procurement process. The Contract Management phase concerns the post-award activities and the implementation, management and performance of procurement contracts and deliverables (Cherono & Chepwoy, 2021).

The base of the diagram is represented by common principles of procurement stated by international development organisations including integrity, ethics and accountability,

procurement policy, binding authority and transparency, fairness, competition and best value (Asian Development Bank, 2018; World Bank Group, 2021).

### 3.13 PRELIMINARY THEORETICAL FRAMEWORK FROM THE LITERATURE

The theoretical framework is a preliminary framework that incorporates the concepts described in the literature review in Chapter 3 of this research, the proposed framework illustrated in Figure 6 shows how the combination of procurement, project management and performance measurement lead to DE and VFM. It has been demonstrated that the procurement process is critical for the effective implementation and delivery of development projects, as discussed in Section 3.11 (Assaf & Al-Hejji, 2006).

Furthermore, the point was made in Section 3.10 that project performance evaluation is measured using project evaluation criteria including relevance, coherence, effectiveness, efficiency, impact and sustainability (OECD, 2019). The effective and efficient management of procurement will lead to improved implementation, delivery, project management and performance. In addition, if projects are successful, then development outcomes will be more effective and improve DE as discussed in Section 3.11.

The theoretical framework in Figure 6 illustrates the measurement of procurement maturity, project evaluation and the relationship with project performance, DE and VFM.

The Integrated Procurement Maturity Model (IPMM) refers to the assessment of procurement maturity and performance and pertains to Research Question 1. The project performance evaluation refers to the evaluation of project performance and relates to Research Question 2, while Research Question 3 examines the relationship and association between procurement assessment and project performance evaluation.

Research Question 4 seeks to understand the performance measurement of the procurement and project management processes and their contribution to DE and VFM.



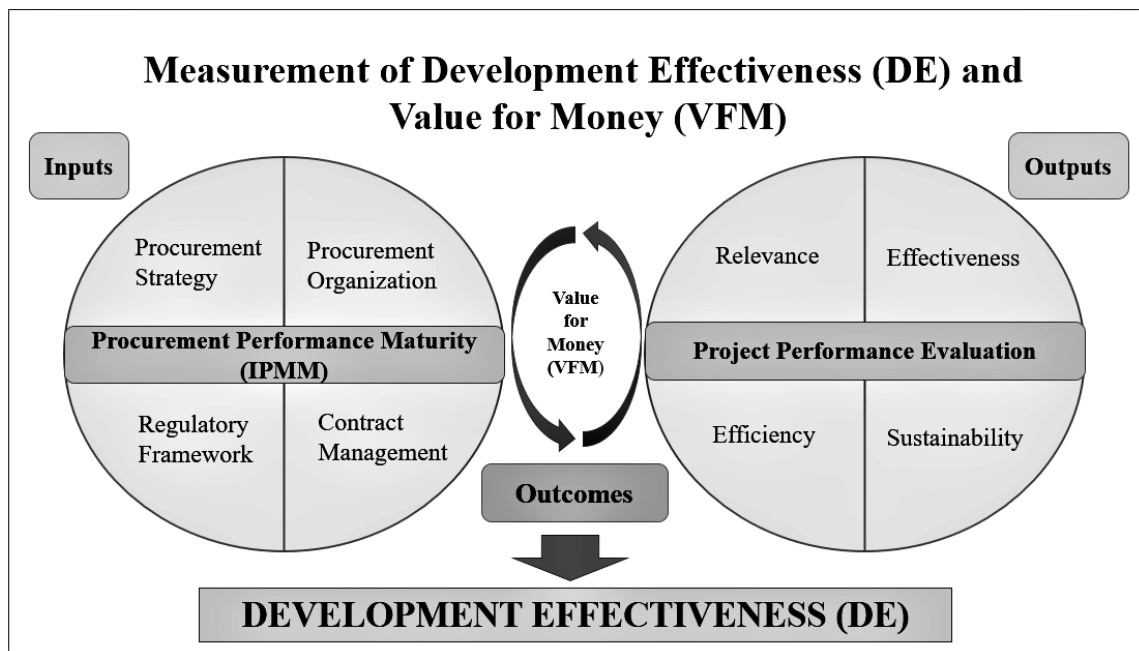


FIGURE 6. THEORETICAL FRAMEWORK SUGGESTED BY THE LITERATURE

The functions and process of procurement, project management and evaluation represent the key mechanisms for the development of design, specifications, scope, implementation, delivery and measurement of project performance. In addition, there is considerable overlap between procurement, project management, contract management and they are often considered to be components of the same process (Batenburg & Versendaal, 2008; Rendon, 2008). The three functions of procurement, project management and evaluation have equivalent organisational groups in international development organisations. The procurement function has separate and distinct structural and process approaches and arrangements within different organisations (Rendon, 2015; Schiele, 2007).

Figure 6 represents the culmination of the theory and practice discussed in Section 2 and Section 3. It illustrates the theoretical framework used in this research, and represents the relationship between the measurement of procurement maturity, project evaluation and the achievement of improved project performance and contribution to DE and VFM.

Based on the findings in the literature review, it has been determined that inputs and outputs of the three delivery mechanisms of procurement, project management and evaluation ultimately determine DE and VFM through projects. More specifically, Section 3.11 and Section 3.12 showed that many key inputs are made through the procurement process which is a combination of four phases, or pillars, including procurement strategy, procurement organisation, regulatory frameworks and contract management, shown in Figure 5. Section 3.4 describes several definitions of Value for Money (VFM), which combine the concepts of efficiency, quality, resources and costs, which are captured in the procurement and project management processes. The development organisations, in this research, evaluate project performance using the four criteria: relevance, effectiveness, efficiency and sustainability. The evaluation criteria measure project performance and it is posited that the achievement of DE and VFM is the result of the measurement and delivery and combination of effective procurement and project management processes.

Overall, the literature suggests that the success of development investment is dependent upon inputs, outputs and outcomes in the assessment of development effectiveness and gauging value for money. The procurement and project management processes serve as the critical mechanisms that capture the inputs, outputs and outcomes. The theoretical framework in Figure 6 shows how these concepts contribute towards DE and VFM, through the procurement and project management processes, towards development objectives in projects.

This research seeks to understand the relationship and performance of procurement, project management and measurement of DE and VFM in international development.

### 3.14 RESEARCH QUESTIONS

Against the background of this extensive literature review, the overall objective of this study is to examine the effectiveness and efficiency of money spent on development aid by international development organisations through procurement. It is estimated that over half of ODA spent is in the form of procurement on goods, services, works and consultants (Ellmers, 2011). The procurement process represents an essential mechanism through which a significant proportion of development aid is managed and spent (Ellmers, 2011; Ika, Diallo, & Thuillier, 2010). The procurement process includes established metrics for the measurement of procurement performance and nature, value, and volume of the flow of funds (Jackson, 2012). The measurement of procurement performance provides the ability and opportunity to identify and measure VFM as part of existing development evaluation methodology. The OECD-DAC project evaluation methodology used by development organisations in this research, evaluates development effectiveness from the perspective of the individual project as the unit of measurement.

More specifically, the following research questions are proposed:

1. What is the comparative procurement maturity of leading international development organisations from the perspective of DE and VFM?
2. How does existing project evaluation assessment methodology measure DE and VFM in leading international development organisations?
3. What is the relationship between procurement performance measurement and project performance evaluation in international development organisations?
4. Does procurement maturity and project management performance contribute towards DE and VFM in international development projects?

### 3.15 RESEARCH AIMS AND OBJECTIVES

The research questions were addressed in line with the following aims and objectives:

1. Develop an Integrated Procurement Maturity Model (IPMM) to measure procurement performance and VFM for international development organisations.
2. Analysis of project performance evaluation to understand how project evaluation reports are used to measure procurement performance, DE and VFM.
  - 2.1 Review the relationship between the procurement performance and project evaluation results by statistical analysis and correlation tests.
3. Conduct further analysis of the procurement performance and project evaluation results to examine how the two processes contribute towards DE and VFM.

This research reviewed common existing Procurement Maturity Assessment models (PMAs) and recognises the requirement for a new model more suitable for international development organisations and the measurement of VFM; existing models are either incomplete or lack specific relevance to their purpose. Procurement assessment models should provide a useful method to measure procurement performance for international development organisations and discussed in the following chapters.

Secondly, project evaluation is measured by leading international development organisations by a common OECD project evaluation methodology. The ten selected international development organisations follow the OECD project evaluation methodology, with some individual adaptations, and hence provide secondary data for project performance that is comparable for further analysis. To understand the influence of DE and VFM for procurement and project management this research adapts the IPMM and drills down to further explore the project evaluation data.

### 3.16 CONCLUSION

This chapter reviews the role of procurement, project management and project performance evaluation in international development projects. The chapter also outlines the criticism of international development to encompass questions on where and how the funds are spent, and concerns of transparency, effectiveness, the efficiency of the disbursement process for development funds. International organisations made commitments and set targets for the improvement of DE as part of the Paris Declaration on Aid Effectiveness, 2005 (Wood et al., 2011). The commitments were accompanied by the parallel development of concepts such as VFM, performance-based management methods and MfDR (Renard & Lister, 2015). However, leading international development organisations acknowledge that there is a disparity between the stated commitments for improvements in DE and tangible evidence of progress towards the development objectives (Collier & Dollar, 2004). In addition, project evaluation methods are inconsistent and do not include comprehensive indicators for cost performance and procurement activity (White, 2005). Given the significant proportion of ODA and other forms of aid spent on the procurement of goods and services (Ellmers, 2011), it is important to maintain high standards of governance, transparency and procurement performance to ensure project delivery, DE and VFM (Aulia & Isvara, 2021). Best practices in procurement and project management processes from both the private and public sectors provide examples of successful methods and techniques. They can also reveal proven methods for improving project management and performance for development projects (Moradi et al., 2020).

For the measurement of procurement performance, leading PMA models present clear baselines for the assessment, benchmarking, and standard setting of DE and VFM for procurement managed directly or indirectly by international development organisations.

Private sector and industry procurement maturity models include attributes designed to measure cost performance and VFM concepts, and these attributes can be incorporated to existing PMA models to improve the measurement of DE and VFM (Schiele, 2007).

Examples of process discipline and the focus on transparency and governance, intrinsic to the procurement and project management processes, can also be adopted and utilised to benefit other forms of aid disbursement, implementation and delivery methods.

Improved transparency, monitoring of results and attention to DE and VFM, in turn, potentially reduces the incidence of funds lost to corruption, waste and inefficiency.

Development organisations can then maximise every dollar of public and taxpayers' money committed towards development (Easterly & Pfitze, 2008; McKee et al., 2020).

Organisations that incorporate procurement and project management performance measurement into their development evaluation models further support the commitments made towards DE and VFM under the Paris Declaration on Aid Effectiveness (2005).

The next chapter takes the critical analysis of procurement assessment models and project performance evaluation and introduces the design of this research. The research philosophy, sampling, data collection and data analysis are all discussed.

## **4 RESEARCH DESIGN**

The previous two chapters introduced the industry and theoretical aspects of this research which led to the development of the theoretical framework in Figure 6, and the four research questions. The gap in the literature review concerning the measurement of procurement performance, project performance evaluation and the link to DE and VFM that was identified will be addressed through this research. Chapter 4 outlines the research design, and this encompasses research philosophy and then moves into sampling, data collection and data analysis. The first two research questions, proposed in Section 3, are addressed via two data collection phases that align with the theoretical framework of the research and measure procurement and project management process performance. The second two research questions are addressed by the analysis of the primary and secondary data to understand the relationship between the two processes and their relative contribution towards DE and VFM.

### **4.1 INTRODUCTION**

This research takes a pragmatist approach and acknowledges that both subjective and objective approaches are critical for understanding the complexity of roles and process in specific context and use of multiple methods of inquiry seeking practical explanation in addition to observation (Goldkuhl, 2012). The research, therefore, uses an abductive approach to understand the role that procurement plays in the effectiveness of the disbursement of international development funds in development projects.

The literature review in Chapter 3, showed a dearth of current knowledge about assessing procurement effectiveness and Value for Money (VFM) in development agencies. The research takes an exploratory approach to develop an applicable PMA model for comparing procurement performance and maturity of leading international development organisations. The new model incorporates key procurement criteria

illustrated in Figure 5, which can be weighted to assess Value for Money (VFM) delivered by the procurement process based on the definition for VFM used by leading MDBs, such as the ADB shown in Figure 3 (Asian Development Bank, 2019).

The integrated procurement maturity model (IPMM) is developed from critical analysis of the literature research in Chapter 3 on multiple PMA models from the public and private sectors with reference to current knowledge on procurement performance measurement. The weighting of criteria is calculated by multiplication by VFM factors, developed from expert input from a panel of professional procurement experts. The procurement criteria are evaluated by the panel, in order of priority for relative influence of each criteria towards VFM during the procurement process. The consolidated results lead to the development of the relative weight of each criteria for contribution towards VFM used as a multiplier as part of the IPMM questionnaire. The IPMM questionnaire is then used to assess procurement performance for both unweighted and weighted procurement criteria. The results are analysed to provide a comparative analysis for the influence and relative contribution of the procurement process performance and criteria towards VFM.

The key data for the IPMM questionnaire are both qualitative and quantitative and collected from the questionnaire, which is designed to gather and explore information directly from each of the selected international development organisations (Van Maanen, Sørensen, & Mitchell, 2007). The questions are based on a Likert scale, or multiple response format, and capture the numerical responses to questions on the procurement function of the organisation for statistical and comparative analysis (Hodgson, 2010).

As discussed in Section 3.10, in international development, development effectiveness for projects is currently measured using evaluation methods based directly or on



variations of the OECD evaluation methodology. OECD evaluation methodology uses standard evaluation criteria to measure the overall outcomes of development projects, as discussed in Section 3.10. The project evaluation reports are published by each of the participating international development organisations in annual reports, specific evaluation reports and are available on the internet as secondary data, for review and comparative analysis.

Considering that a significant proportion of Official Development Assistance (ODA) is disbursed through the procurement process, the research provides an analysis and comparison of the procurement performance results from the IPMM for each of the organisations (Ellmers, 2011). The IPMM results for the development organisations are compared against the aggregate project performance results by year, sector and country.

The research seeks to address the research questions set out in Chapter 3, by providing an analysis of the IPMM procurement performance results and the project evaluation performance results for the selected international development organisations. The research process is illustrated in Figure 7, and the flow diagram shows the two distinct, but complementary, phases of the research: Phase 1, development, issue and application of the IPMM model and Phase 2, collection and collation of the project evaluation results. The two phases then come together in Phase 3, to form the basis for addressing the third and fourth research questions; to examine the relationship between procurement and project management of the theoretical framework and contribution to DE and VFM.

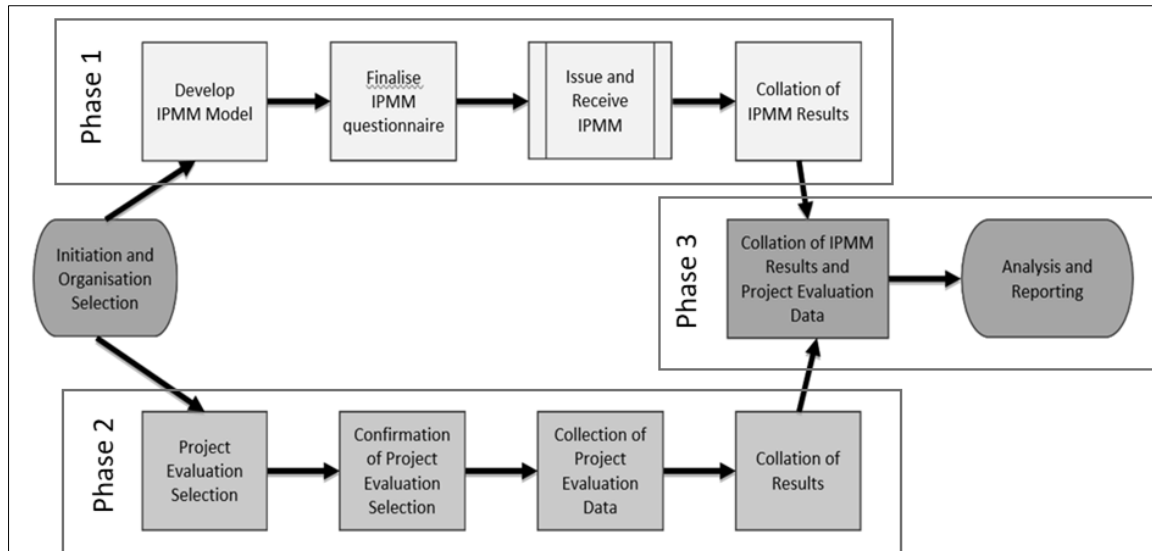


FIGURE 7. FLOW DIAGRAM OF THE RESEARCH PROCESS

## 4.2 METHODOLOGICAL STEPS IN THE RESEARCH

### 4.2.1 *Research philosophy*

The research philosophy is key to the understanding of the theory, evidence and practical considerations of the changing nature and differences of organisations and their processes. The philosophy aids in creating the appropriate context as the researcher searches for the truth in the research questions (Astley & Van de Ven, 1983). There are practical and philosophical obstacles that cause challenges for researchers, so a thorough understanding of the philosophy and practical factors of the research is important to ensure that research objectives are achieved (Miller & Tsang, 2011).

In management research concerning different international development organisations, there are numerous variables and multiple causations which need to be considered in testing, analysis, comparison and correlation of the results to ensure meaningful review and conclusive findings (Fabian, 2000).

There are several approaches to research philosophy and therefore great care is required to adopt approaches to research that are best aligned to address the research questions.

Saunders, Lewis and Thornhill's (2016) ubiquitous research onion in Figure 8, illustrates the key parameters in designing research and is used in justifying this research.

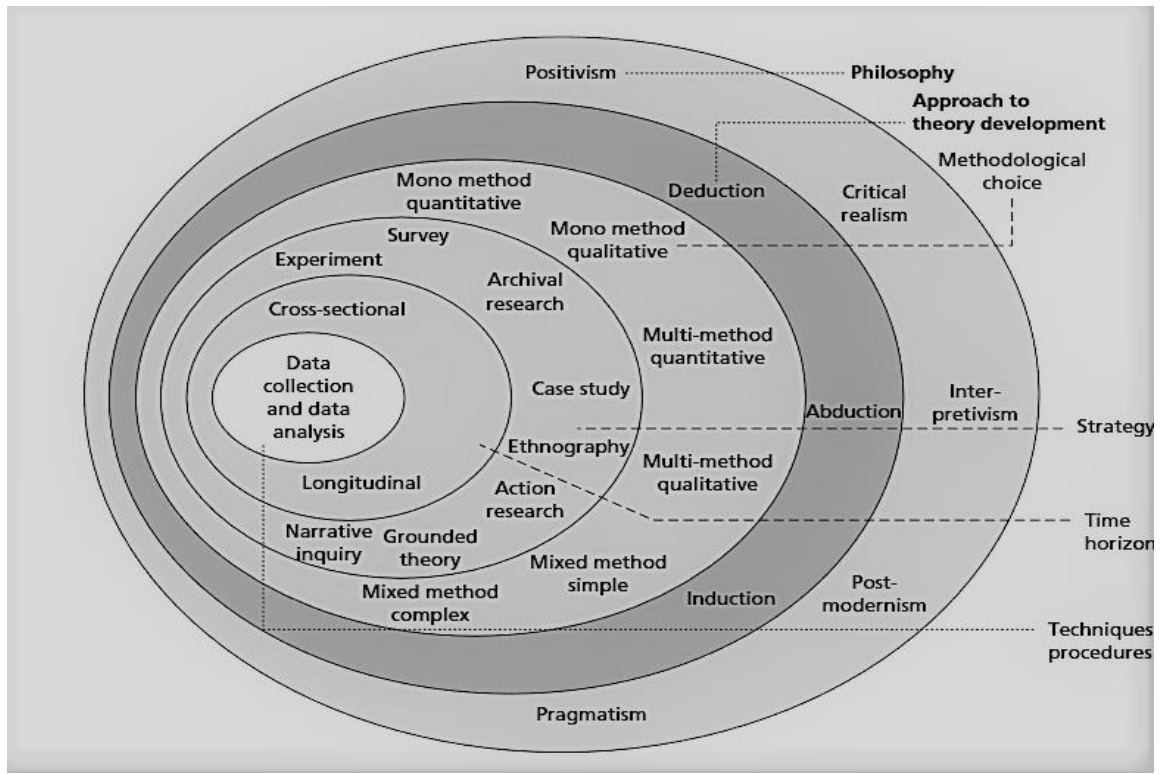


FIGURE 8. THE RESEARCH ONION

(Saunders, Lewis, & Thornhill, 2016, p. 164)

As explained by Table 14, there are four research philosophies (positivism, realism, interpretivism and pragmatism) that are applied to ontology, epistemology and axiology (Saunders, Lewis, & Thornhill, 2009). Ontology is concerned with the researcher's view of the nature of reality, epistemology refers to the researcher's view of what constitutes acceptable knowledge and axiology is the researcher's view of the role of the value in the research. Data collection techniques are included in the table to provide guidance on applicable technique and methods for handling data collection.

TABLE 14. COMPARISON OF RESEARCH PHILOSOPHIES IN MANAGEMENT RESEARCH

	<b>Positivism</b>	<b>Realism</b>	<b>Interpretivism</b>	<b>Pragmatism</b>
<b>Ontology:</b> <i>the researcher's view of the nature of reality or being</i>	External, objective and independent of social actors	Is objective. Exists independently of human thoughts and beliefs of knowledge of their existence (realist), but is interpreted through social conditioning (critical realist)	Socially constructed, subjective, may change, multiple	External, multiple, view chosen to best enable answering of research questions
<b>Epistemology:</b> <i>the researcher's view regarding what constitutes acceptable knowledge</i>	Only observable phenomena can provide credible data, facts. Focus on causality and law like generalisations, reducing phenomena to simplest elements	Observable phenomena provide credible data, facts. Insufficient data means inaccuracies in sensations (direct realism). Alternatively, phenomena create sensations which are open to misinterpretation (critical realism). Focus on explaining within a context or contexts	Subjective meanings and social phenomena. Focus upon the details of situation, a reality behind these details, subjective meanings motivating actions	Either or both observable phenomena and subjective meanings can provide acceptable knowledge dependent upon the research question. Focus on practical applied research, integrating different perspectives to help interpret the data
<b>Axiology:</b> <i>the researcher's view of the role and values in research</i>	Research is undertaken in a value-free way, the researcher is independent of the data and maintains an objective stance	Research is value laden; the researcher is biased by world views, cultural experiences and upbringing. These will impact on the research	Research is value bound, the researcher is part of what is being researched, cannot be separated and so will be subjective	Values play a large role in interpreting results, the researcher adopting both objective and subjective points of view
<b>Data collection:</b> <i>techniques most often used</i>	Highly structured, large samples, measurement, quantitative, but can be use qualitative	Methods chosen must fit the subject matter, quantitative or qualitative	Small samples, in-depth investigations, qualitative	Mixed or multiple designs, quantitative and qualitative

Source: (Saunders et al., 2009, p. 119)

After consideration of the research philosophies in management research and the view of the researcher, pragmatism was determined to be the most appropriate methodology to sufficiently address the research questions. This approach supports the position that neither positivist, realist nor interpretivism is applicable to the research and subject matter. The research has some dependency on social factors, is not subjective or socially constructed in all aspects and does not exist independently of human thoughts and beliefs (Morgan, 2007). Pragmatism allows the application of a more flexible philosophy and

involves both quantitative strategies and consideration of qualitative factors in order to address and answer research questions. Examples in the information technology and systems industry have demonstrated the use of qualitative research and quantitative research for a powerful combination of theory and action research methods (Goldkuhl, 2012). Due to the highly practical nature of this research and its grounding in international development practices, pragmatism is by far the most appropriate and suitable methodology herein. The practical and mixed methods philosophy caters for the rigour and relevance of diverse data and integrated methodology and includes the four key themes of: experience, enquiry, habits and transaction, applied in this research (Elkjaer & Simpson, 2011; Feilzer, 2010).

#### *4.2.2 Research approach*

Further to the selection of pragmatism as the most appropriate research philosophy, three research approaches are considered and these are identified as deductive, inductive and abductive. In the case of the deductive approach, the researcher adopts a theory and hypothesis from a set of premises and the research strategy is designed to test whether those premises are true. The deductive approach emphasises the need to focus on scientific principles, utilise data and causal relationships between variables and structured quantitative data (Saunders, Lewis, & Thornhill, 2012).

In the inductive approach, the theory follows the information and analysis as the researcher collects data to develop a theory derived from the data analysis. The inductive approach emphasises collection of qualitative data, a more flexible structure and an understanding of the meaning humans attach to events (Kovács & Spens, 2005).

The third research approach is a combination of deductive and inductive or an abductive approach is where data are collected to explore a phenomena, themes and patterns to

generate new or modify existing theory which is then analysed and tested through additional data collection (Dubois & Gadde, 2002).

The abductive approach was deemed the most appropriate for this research as it is largely quantitative in nature but modified through the collection of limited and minor qualitative inputs to develop weighting for use in data analysis to explore the phenomenon, generate theory and testable conclusions. By the introduction of intuition into the scientific approach abduction is different from previous method of scientific explanations (Kovács & Spens, 2005). A comparative description of the three approaches is illustrated in Table 15.

TABLE 15. COMPARISON OF DEDUCTIVE, INDUCTIVE AND ABDUCTIVE APPROACHES

Description	Deduction	Induction	Abduction
Logic	In a deductive inference, when the premises are true, the conclusion must also be true	In an inductive inference, known premises are used to generate untested conclusions	In an abductive inference, known premises are used to generate testable conclusions
Generalisability	Generalising from the general to the specific	Generalising from the specific to the general	Generalising from the interactions between the specific and the general
Use of Data	Data collection is used to evaluate propositions or hypotheses related to an existing theory	Data collection is used to explore a phenomenon, identify themes and patterns and create a conceptual framework	Data collection is used to explore a phenomenon, identify themes and patterns, locate these in a conceptual framework and test this through subsequent data collection and so forth
Theory	Theory falsification or verification	Theory generation or building	Theory generation and modification;

Source: (Saunders et al., 2016, p. 145)

#### 4.2.3 Research methods

In the design of the research, quantitative, qualitative or mixed method or methodology choice is necessary to examine and apply. Quantitative research involves collection of

largely numerical data and may involve the use of questionnaires and other means of collection of quantitative data such as secondary data sources and published information.

Qualitative research includes other forms of data collection which may be more subjective and open to human thoughts and beliefs and very often both of these methods have some element of combination and are defined as mixed methods (Du Plessis & Majam, 2010).

An argument for mixed methods and the integration of quantitative and qualitative data is that it casts aside some of the traditional methods of research and takes the approach that the most important issue is to answer the research questions (Feilzer, 2010). The approach does not mean that a pragmatism should imply lower standards of research or that expediency is more important than accurate data and vigilance. On the contrary, mixed methods can expand the quality of data and information for analysis (Denscombe, 2008). There is also the opportunity to gain insights into unexpected relationships between data from different organisations and responsive to different process and environmental factors (Driscoll, Appiah-Yeboah, Salib, & Rupert, 2007). Mixed methods are increasingly being used to introduce innovation, add value and gain greater insights and reduce some of the constraints perceived by researchers in research of complex business areas (Cameron & Molina-Azorin, 2011). Mixed methods research can also potentially answer questions that cannot be answered by quantitative and qualitative data alone and researchers can use a wider range of methods to address a research problem (Du Plessis & Majam, 2010).

Table 16 explains the numerous strategies for combining quantitative and qualitative methods for optimised research design (Greene, Caracelli, & Graham, 1989).

TABLE 16. JUSTIFICATION OF COMBINING QUALITATIVE AND QUANTITATIVE RESEARCH

<b>Reason</b>	<b>Explanation</b>
Triangulation	Convergence, collaboration, correspondence or results from different methods. In coding triangulation, the emphasis was placed on seeking collaboration between quantitative and qualitative data.
Complementarity	Seeks elaboration, enhancement, illustration, clarification of the results from one method with the results from another.
Development	Seeks to use the results from one method to help develop or inform the other method, where development is broadly construed to include sampling and implementation as well as development decisions.
Initiation	Seeks the discovery of paradox and contradiction, new perspectives of frameworks, the recasting of questions or results from the other method.
Expansion	Seeks to expand the breadth and range of enquiry by using different methods for different enquiry components.

Source: (Bryman, 2006, p. 105)

This research adopts a mixed method design with both quantitative and qualitative data collected through questionnaire design and secondary data, derived from published information from the selected international development organisations.

The quantitative and qualitative data are processed through the three phases in this research which include the distribution and collection of the questionnaire responses, the collection of secondary data and the analysis of both sets of results, as shown in Figure 7.

Quantitative data are collected directly from the IPMM questionnaire and also included in the content of the published secondary project evaluation reports. Qualitative data are collected by the IPMM questionnaire and are also contained in the secondary data.

Questionnaires on procurement, for example, often collect quantitative and qualitative procurement data related to organisation, procedures, process and systems (Adams & Cox, 2008; Liebetrueth, Melneck, & Pils, 2016). The questionnaire includes Likert and multiple response questions and the secondary project evaluation reports include evaluation results, expressed as numerical scores and ratings and used for comparison and quantitative analysis (Hodgson, 2010; Patton, 2002). The use of results from different methods enables triangulation, convergence collaboration and analysis between



source of qualitative and quantitative data inputs (Bryman, 2006). Quantitative procurement data from the IPMM questionnaire, for example, may reveal positive or negative relationships with project success and performance. The data is then subjected to statistical tests such as correlation analysis to examine the nature of the relationships. The combination of the two strategies enables complementarity and illustration of results from both qualitative and quantitative data types. The methods allow development and conversion of subjective data and observations to objective comparison and analysis (Hodgson, 2010). In addition, the method encourages initiation of new perspectives of the procurement and project management frameworks and expansion of breadth of enquiry to explore DE and VFM. The research is expected to lead to the discovery of new insights and alternative analysis may reveal unexplainable results or conclude that further research or additional data may be required (Bell, Bryman, & Harley, 2018).

The research purpose and the way research questions are asked are defined as exploratory, descriptive, explanatory or evaluative research (Yin 2003). Exploratory studies ask open questions to explore what is happening and gain insights on about the subject of interest. Descriptive research seeks to gain an accurate understanding of events, individuals and situations by description of the subject and information.

Explanatory studies aim to study a situation or problem to explain the relationships between different variables. Evaluative research, on the other hand, makes comparisons between events, situations, groups, places or periods (Fowler Jr, 2013).

Based on this rationale, the research design is comprised of mixed methods and a combination of explanatory and evaluative research studies. This includes a combination of a questionnaire and archival, documentary information and the data are used for comparison, evaluation and analysis in this research.

The mixed methods research includes both quantitative and qualitative methods and the creation of an IPMM questionnaire, shown in Appendix 1, that compares procurement performance with project evaluation performance data collected from secondary sources.

#### *4.2.4 Time horizons*

Time horizons research is defined as Cross-sectional Studies and Longitudinal Studies. Cross-sectional studies refer to events or phenomena at a specific time relating to factors and may be applied to comparison of different organisations. Longitudinal studies are defined as charting progress and variables over time and providing information on the progress or variable. Longitudinal studies may involve secondary data and the comparison of snapshots of information over a selected time period to compare events over different years or time periods (Greener, 2008).

The purpose of the research is to compare procurement performance of international development organisations against project evaluation performance over selected time periods, geography and sector. Therefore, both definitions of cross-sectional and longitudinal studies are used in this research. The cross-sectional aspect relate to the comparison of different organisations, processes, systems and organisational structure at a specific time. The longitudinal studies, in this research, include secondary data and variables over selected time periods relating to project performance and the evaluation data (Hewson, Vogel, & Laurent, 2015).

### **4.3 DATA SOURCES AND DESCRIPTIONS**

The primary data, in this research were collected via a procurement questionnaire designed and adapted for international development organisations. The secondary data were collated from both traditional and electronic published sources and organisation web sites as available in the data collection phase of this research from 2017 and 2019.

A key part of the research onion, shown in Figure 8, deals with the research's data collection and how the data are analysed; several methods are identified, including:

- Sampling (Saunders et al., 2016, p. 276)
  - a) Probability or representative sampling where the chance for selection from the population is known and has an equal likelihood of selection. Due to these characteristics, probability sampling is frequently used for survey and experimental research strategies
  - b) Non-probability sampling where the probability for each case being selected is not known and it is not possible to make statistical inferences from the data on the characteristics of the population sampled.
- Secondary data (Saunders et al., 2016, p. 318)
  - a) Documentary form of data in a written form or non-written data in the form of audio, visual or sourced from electronic or physical document storage or archives.
  - b) Multiple source data based on geographical area under the collection strategy or data collected over a specified time-period.
  - c) Surveys such as censuses often carried out by governments and are mandatory or obligatory. Continuous and regular surveys repeated over a time with data collection taking place on a continuous basis. Ad hoc surveys carried out when required or for a specific need usually as a single event.
- Primary data (Saunders et al., 2016, pp. 356, 436)
  - a) Participant observation as qualitative data to derive meaning from the actions of participants. Structured observation as quantitative data collected from frequency on participant actions. Interviews in the form of structured interviews using questionnaires based on standard questions and used for

descriptive research rather than explanatory purposes. Semi-structured interviews for explanatory purposes and include themed questions that may vary depending on the interview. Unstructured interviews with no standard questions and conducted in an informal manner for exploratory purposes.

- b) Questionnaires can be self-administered through postal delivery, internet, intranet and other manual forms of delivery. Questionnaires can also be administered by telephone or through structured interviews.

The following sections describe the research design, selection of organisations and administration used in this research.

#### *4.3.1 Selection of the international development organisations*

In alignment with the research questions, this research assesses and compares the procurement procedures and processes of leading international development organisations including MDBs, UN agencies and bilateral development organisations.

Criteria for identifying organisations for research includes procurement volume and annual expenditure of the MDBs, UN agencies and bilateral development organisations.

Secondly, it is necessary to have comparable regional and country data across multiple sectors. Third, for applicable secondary data collection quality and volume, the international development organisation need to share performance measurement methods and have comparable rating scales. Finally, the volume and availability of the secondary data are required to have sufficient breadth and volume for the purposes of collection, comparison and statistical analysis (Walliman, 2017).

Recognizing the large number and type of international organisations, a preliminary search was carried out for MDBs, UN agencies and bilateral development organisation on the basis of procurement volume of more than US\$1 billion per annum. The second selection requirement was that each organisation implement the OECD-DAC project

evaluation methodology. The third requirement was the availability and volume of published public project data to allow effective comparability with the other organisations as required to answer the research questions and this is illustrated in Table 17. The selected international development organisations follow industry best practices in procurement processes, open data initiatives and evaluation methodology. Both the procurement and evaluation development groups of the international development organisations collaborate with organisations such as such as OECD, and continue to review, develop, improve and adopt common standards and methodologies (OECD-DAC, 2020). However, differences remain in the adoption of OECD methodologies, comparative performance in development outputs and outcomes, collection of data and measurement of performance among many of the leading organisations themselves (Gulrajani, 2014).

The selection of the international development organisations, for suitability of participation in this research, is based on a combination of factors including procurement attributes, such as volume, and the availability of comparable project evaluation data. Many international development organisations have procurement volumes suitable for participation in the IPMM. However, a relatively lower number of development organisations follow the OECD-DAC project evaluation methodology, as an essential requirement and condition for the comparability of the secondary data in this research. A summary of the selection requirements and parameters are illustrated in Table 17.

TABLE 17. PARAMETERS FOR THE SELECTION OF THE DEVELOPMENT ORGANISATIONS

<b>Organisation requirements</b>					
<b>Name</b>	<b>Type</b>	<b>Size</b>	<b>Procurement Value 2017 (US\$)</b>	<b>OECD- DAC Methodology</b>	<b>Availability of DATA</b>
WBG	MDB	Global	14.3	√	√
ADB	MDB	Regional	6.5	√	√
IADB	MDB	Regional	4.5	√	√
EBRD	MDB	Regional	2.1	√	√
AFDB	MDB	Regional	1.7	√	√
ISDB**	MDB	Regional	5.77	√	√
UNICEF	UN Agency	Global	3.5	√	√
UNPD	UN Agency	Global	3.2	√	X
UNHCR	UN Agency	Global	1.2	√	X
UNDP	UN Agency	Global	1.7	√	√
WFP	UN Agency	Global	2.9	√	X
USAID*	Bilateral	Global	19.35	√	X
JICA**	Bilateral	Global	7	√	√
GIZ	Bilateral	Global	1.4	√	X
DFID	Bilateral	Global	1.4	√	√

*Note\*:* USAID volume is defined as commitments to international development and includes procurement, projects, programmes, training and contribution to other agencies.

*Note\*\*:* ISDB and JICA volumes includes loans, grants with a major proportion of procurement activity.

In the case of the MDBs many of the leading organisations have large procurement volumes, a global or regional footprint, follow OECD-DAC evaluation methodology and publish evaluation data through data sets or by annual evaluation reports.

The WBG is the only MDB that has a global geographic presence and WBG publishes both procurement and evaluation data on its external website and annual reports (World Bank Group, 2019). The ADB, EBRD, AFDB and IADB also follow the selection parameters in Table 17, although the collection of evaluation secondary data requires navigation through several different data sources and reporting systems (African Development Bank, 2018; Asian Development Bank, 2017; European Bank for Reconstruction and Development, 2017; Inter-American Development Bank, 2019).

In the case of Islamic Development Bank (ISDB), for example, the procurement reporting is combined with data on disbursement for project loans and grants published in the annual ISDB development effectiveness report (Islamic Development Bank,

2017). The United Nations agencies' procurement volumes are documented in the Annual Statistical Report (United Nations Office for Project Services, 2020).

Many of the UN Agencies follow OECD-DAC methodology, although there are differences in the consistency and content of procurement and development evaluation reporting. For example, the reporting of secondary data for development evaluation is much more dispersed. Some data are reported through organisational publications and websites and others jointly published by the United Nations Evaluation Group (UNEG).

UNDP publishes the largest volume of evaluation data, compared to other UN Agencies, and UNICEF similarly includes large volumes of data, but has differences in reporting format and content (United Nations International Children's Emergency Fund, 2018).

Other smaller UN agencies do not follow OECD-DAC evaluation methodology in a consistent manner, and therefore, data are not available for analysis in a comparable format as reflected in the United Nations Annual Statistical Report (ASR) (United Nations Office for Project Services, 2020).

Bilateral development agencies also have different approaches to evaluation methodology and differences in availability and quality of data. DFID and JICA publish data in their evaluation and annual reports in a detailed and consistent manner (Department for International Development (DFID), 2018; JICA, 2013).

USAID has a great deal of data available and a Development Experience Clearinghouse (DEC) covering comprehensive development and evaluation information. The evaluation methodology does not consistently follow the OECD-DAC format and is difficult to compare with the other bilateral organisations and hence, is not included in this research (USAID, 2020).

GIZ follow the OECD-DAC evaluation methodology and publish data in the annual evaluation report but do not include sufficient depth of data for comparability against the other development organisations and is not included in this research (GIZ, 2016, 2017).

Ten international development organisations are listed in Table 17, which have similarities in that they are leading international development organisations with comparable procurement volumes. These organisations also follow OECD-DAC evaluation methodology and have availability of published secondary project evaluation data in both quantity and quality for effective comparative analysis in this research.

#### *4.3.2 Selection of the procurement expert panel*

Based on the guidance of Dalkey and Helmer (1963) about expert panels, the procurement expert panel for this research aimed to obtain the most reliable consensus of opinion about VFM on procurement by a selected group of procurement experts. The procurement expert panel provides inputs on the relative influence of procurement criteria towards VFM in the procurement process to develop VFM factors. The VFM factors are converted to weights in the IPMM to produce the weighted IPMM results and the VFM factors are shown in the following sections. The creation and use of weights in the IPMM is commonly used in private sector and development procurement and serves as a form of sensitivity analysis for VFM. The method recognises that each criteria may have variable influences and is used in procurement vendor evaluations (Banda, 2019). The identification and development of the differences into VFM factors is derived from the inputs from the procurement expert panel used for the sensitivity analysis (Beecham, Hall, Britton, Cottee, & Rainer, 2005). This research recognises that the VFM factors are based on validation and feedback from the procurement expert panel and the factors may be refined and adjusted in a future exercise by expansion or alteration of the panel participants. The VFM factors are used in the research to test the IPMM questionnaire



results specifically for characteristics for VFM by the sensitivity analysis. Hakim (1987) noted that expert panels can be used to test explanations at early stages of research (Hakim, 1987). Previous studies have included expert panels in numbers of 10 to 30 participants to gain expert advice, conduct reviews on various subjects and develop criteria for model development (Dyba, 2000). Expert panels can be drawn from different organisations and industries to provide feedback for research purposes or information on issues of concern (El Emam & Madhavji, 1996; Soares, Toffoletto, & Deschênes, 2006). For example, Verbrugge et al. (2019) used expert panels comprised of multiple experts for the validation and judgement of data in conjunction with questionnaires.

In this research, a Request for Information (RFI) document was designed to collect feedback from expert procurement professionals to verify and then prioritise the 12 procurement process criteria in rank order for their comparative impact towards achieving VFM, and the request for information format is shown in Appendix 2.

The request for information provides information and expert professional feedback and input on the procurement criteria based on the respective expertise of the participants.

The intention of the technique was to obtain the opinion of a group of experts and collate quantitative and qualitative feedback, which can be used to develop weighted criteria.

The research combines individual expert panel feedback to develop weighted criteria that are applied to the IPMM, to reflect the influence of VFM on the procurement process of the international development organisations (Rohde et al., 2020; Verbrugge et al., 2019).

The factors or weights are derived from the feedback from the expert panel based on the panel's subjective expertise and experience and hence, are used for sensitivity analysis as a supplementary part of the IPMM questionnaire (Beecham et al., 2005; Mullen, 2003).

The participants are selected for their expertise in procurement and as leaders and experts from different industries. The requirement in this research is for the expert participants to

provide feedback, including a percentage rating, to indicate the relative importance of VFM for each of the 12 procurement criteria in the IPMM, as shown in Appendix 1.

The use of an expert panel in this way, provides a method to obtain and collate expert feedback to introduce further information and provide answers to a complex problem (Linstone & Turoff, 1975). The expert panel provides qualitative feedback, to prioritise the procurement criteria of the IPMM and the qualitative feedback is converted into quantitative data or percentages. The quantitative data are then used to create factors, or weights, to test and understand the relative impact of the IPMM procurement performance results on VFM (Driscoll et al., 2007; Fafchamps & Woodruff, 2016).

There is no standard method to determine sample size of participants for expert groups although 15 has been suggested as a common number (Campbell & Cantrill, 2001; Holsapple & Joshi, 2000). MacMillan, King and Tully (2016) argued that increasing the number of participants may increase the diversity of responses but can also potentially lead to decreased value and relevance of responses (McMillan, King, & Tully, 2016).

The request for information from an expert panel generates results based on the expertise and contribution of the participants and the panel should fulfil the following requirements (Adler & Ziglio, 1996):

- Knowledge and expertise of the subject concerned
- Capacity and willingness to participate
- Sufficient time to participate
- Effective communication skills

In this research the subject matter experts and participants are senior and experienced representatives from procurement organisations in international development, private sector, public sector, consultancies, professional and academic institutions. The

participants are drawn from different global organisations introduces different perspectives, experience and knowledge to apply to the purposes of the objectives.

TABLE 18. EXPERT PROCUREMENT PANEL PARTICIPANTS

<b>Organisation</b>	<b>Number</b>
International Development Organisations	5
Private sector industrial sector	2
Consultancy	2
Academic institution	1
Professional procurement organisation	2
<b>TOTAL</b>	<b>12</b>

The procurement experts are required to comment on their wide perspective on the procurement process in relation to Value for Money (VFM) and it is important to obtain feedback of leading experts from different backgrounds and industries. Development organisations have more recently adopted the VFM concept from academic institutions, professional procurement organisations and private sector (World Bank Group, 2016).

The request for information in this research is distributed by email to the experts as shown in Table 18 accompanied by the participants’ invitation letter, consent form and instructions and received by email and collated with the responses.

The determination of the agreed weights for the procurement criteria is the result of assessment of each different percentage weight input by the participant to the question sheet. After discussion on any anomalies or confirmation, to ensure that the requirements have been fully understood, a simple average is taken for all scores per procurement criteria. The resultant agreed percentage weight is taken as the average of all scores for procurement criteria with a total of one hundred percent (Geist, 2010).

$$\text{Percentage criteria} = \frac{\text{Participant score a} + \text{Participant score b} + \text{Participant score x}}{\text{Participant Number}} \times 100$$

4.3.3 Secondary data collection

There are many sources of secondary data that can be used in research and as part of the data collection process. For research that requires comparison between international

organisations, secondary data are an important source of information to answer research questions (Fowler Jr, 2013). There are quantitative and qualitative types of secondary data used in descriptive and explanatory research and may be comprised of raw data or data drawn from other sources (Zikmund et al., 2013).

Secondary data are drawn from documentation from written materials such as public records, information from databases, administrative records, minutes, letters and reports. Documents can also include journals, articles, books and papers and non-written form documentaries, video, recordings and electronic sources and provide combinations of quantitative and qualitative information. Secondary data in this format can be public information or may require agreement, subscription and/or payment to secure the information required (Hair Jr, Wolfinbarger, Money, Samouel, & Page, 2015).

Secondary data can be derived from existing surveys or questionnaires and a combination of these forms of data from different sources. In international development, for example, there are some sources that consolidate all these types of information for research and analysis purposes (OECD, 2019). Many international development organisations follow open data initiatives which advocate publication and disclosure of both operational and administrative information. This ensures that the organisation, activities and handling of public funds are perceived as transparent and open and enable a high degree of research of a wide variety of secondary data (Open Data Institute, 2021; Open Knowledge Foundation, 2021).

Secondary data have the advantage of saving time and resources in the case that information is public and relatively easy to access and collect. Public data also have the advantage in that there is little interaction with the source organisation and origin of the data and more time can be spend developing analysis and comparative studies.

Challenges involved with secondary data include the large number of potential sources and applicability of the data for research objectives and questions. A great deal of time and effort may be necessary to filter different sources and types of secondary data for quality, validity and relevance (Ahsan & Gunawan, 2010). Reliability of the data is also an important factor that requires discipline of the data collection, entry, analysis and reporting process in the organisation. This may also require elimination of bias as far as possible. The project managers responsible for oversight and reporting for development projects are often called Task Team Leaders (TTLs) (Asian Development Bank, 2019; World Bank Group, 2012). Project evaluation results often self-reported by the TTLs, and therefore it has been suggested that there can be an increased possibility, or a tendency, for positive bias (Denizer et al., 2013; Ika et al., 2010). The research takes a sample of ( $n = 1,920$ ) all published projects between 2015 and 2017 of international development organisations that use the OECD project evaluation methodology and other attributes relevant to this research. The sample of projects from organisations selected from the larger population of the total international development organisations summarised in Chapter 2 and the parameters for selection are described in Section 4.3.1. The selected international development organisations all follows the OECD project evaluation methodology and criteria for evaluate project performance. The definitions of the criteria are broadly the same, however differences in the scale of overall evaluation rating with WBG, AFDB, EBRD, IADB and UNDP using a six-point scale. ADB, UNICEF, ISDB and JICA use a four point scale and DFID uses slightly different description and a five-point scale as discussed in Section 3.10. The comparison of different scales in requires adjustment and this is discussed in a following section. Secondary data were mainly sourced through publicly available data and documents disclosed on the organisation's web sites. Other data sources include annual development

reports, annual evaluation reports published on the web sites, and data sourced directly from the international development organisation. WBG, ADB, AFDB, UNDP, UNICEF and DFID have developed web-based databases for the disclosure of development data (African Development Bank, 2020; Asian Development Bank, 2019; Department for International Development, 2020; United Nations Development Program, 2020; United Nations International Children's Emergency Fund, 2020; World Bank Group, 2021). Web sites and databases for these organisations include links to the project evaluation reports for the review and examination of detailed project information and narrative.

The data and project evaluation reports collected for EBRD, IADB, ISDB and JICA were sourced from annual evaluation reports and project evaluation reports published through their respective websites (European Bank for Reconstruction and Development, 2020; Inter-American Development Bank, 2018; Islamic Development Bank, 2018; JICA, 2020). Open data has been recognised as important for driving the Sustainable Development Goals (SDGs), encouraged by organisations such as the WBG and United Nations (World Bank Group, 2022). Disclosure of government data has been increasing over the last 50 years, the Open Knowledge Foundation was founded in 2004, the United States launched the data.gov portal in 2009, and the Open Data Institute was established in 2012 (Open Data Institute, 2021; Open Knowledge Foundation, 2021). The amount of available project evaluation data for this research has increased since 2012 and shown more consistency from 2015 in response to the increased focus and attention (Verhulst & Young, 2017). The WBG has been fairly consistent with the number of annual project evaluation reports since 1998 although the content, volume and structure of the reports has evolved substantially (World Bank Group, 2022). The ADB roughly doubled the number of annual project evaluation reports disclosed after 2006 and there has also been gradual improvement of content and volume of the project evaluation reports (Asian

Development Bank, 2019). The project evaluation databases of the two organisations are similarly well-organised and well updated for the research data required between 2015 and 2017. The African Development Bank (AFDB) experienced a number of changes and major disruption and relocation of the whole organisation from Côte D'Ivoire to Tunis due to the civil war between 2004 and 2011 and returned in 2015 (African Development Bank, 2021). The compilation and disclosure of some required data were affected by the organisational disruption and was collated in this research by request to the Chief Evaluation Officer of the AFDB.

The EBRD publishes project evaluation data in its Annual Evaluation Review (European Bank for Reconstruction and Development, 2020). The project evaluation reporting format changed from 2015 to conform with the OECD methodology, hence only 2016 and 2017 data were available at time of data collection for the research. IADB produces an Annual Project Completion Report and again only 2015 data were available and published at time of data collection (Inter-American Development Bank, 2018).

Similarly to EBRD and IADB, the ISDB publishes an Annual Evaluation Report and 2015 and 2016 data were available in the evaluation reports at the time of data collection for this research. UNDP publishes project evaluation data in its Evaluation Resource Centre and data were collected by utilizing the advanced search function for 2015, 2016 and 2017 (United Nations Development Program, 2020). UNICEF publishes project evaluation data in the Global Evaluation Reports Oversight System (GEROS) and annual meta-evaluation reports (United Nations International Children's Emergency Fund, 2020). The UNICEF data includes overall project evaluation ratings but at the time of data collection these reports did not contain the breakdown of the ratings for the individual evaluation criteria. DFID publishes project evaluation data in its development tracker database on the Foreign, Commonwealth and Development Office page and data

were collected for 2015, 2016 and 2017. At the time of data collection the data did not include breakdown by individual evaluation criteria (Department for International Development, 2020). JICA publishes data in its JICA Annual Evaluation Report and data were available for the years 2015, 2016 and 2017 at the time of data collection for this research (JICA, 2020). The data from all the different organisations were comparable with the exception of the absence of breakdown of individual evaluation criteria for UNICEF and DFID. Secondly, the three different ratings scales were normalised for comparison as described later in this section. Andrić et al. (2019) observed that there was correlation between project size and cost variation, most notably cost overruns, in East Asia and South Asia but no such correlation in Central Asia. To examine the relationship between project size and cost variation in this research a Pearson's correlation test was performed on project size and cost variation as described in the results Chapter 5.

#### *4.3.4 Primary data collection*

The term questionnaire is used for alternative types or methods of data collection which include face to face, telephone or self-completion by traditional paper format or electronically (Rowley, 2014). The design of questionnaires is important to ensure that there is a good response rate and that results of the questionnaire provide clarity and answers for the purposes towards which the questionnaire has been designed, and ultimately to address the project's research questions. Saunders et al. (2016, p. 439) state that the effectiveness of questionnaire implementation can be maximised by:

- Careful design of the individual questions;
- Clear and pleasing visual presentation;
- Lucid explanation of the purpose;
- Pilot testing;
- Carefully planned and executed delivery and return of completed questionnaire.



Individual questions in this research were designed to align with each of the procurement phases, process criteria and sub-criteria with a description of the purpose of the question. Questionnaires are generally used for descriptive or explanatory research to enable the review and examination of relationships or correlation between variables. Interviews and open-ended questions are considered appropriate for more in-depth exploratory research and to identify additional issues that may not be revealed by the questionnaire alone (Adams & Cox, 2008; Adler & Ziglio, 1996).

The type of questionnaire depends on the method of delivery, return and contact required with the selected participants. Self-completed questionnaires can be distributed, by hand, through the postal service or online; the latter can be responded through the computer or from a mobile device. The type of questionnaire selected also depends on many factors related to research questions and objectives such as the size, characteristics and availability of the respondents. The types and number of questions influence the time required and likelihood of completing the questionnaire. The delivery, collection and communications are also important to ensure that answers are not influenced or distorted by other respondents, peers, managers or other factors that can contribute to biased responses (De Vaus, 2013). Contamination and bias of the respondents' answers reduces reliability of the data and participants with little knowledge of the subject may have the tendency to guess and produce an uninformed response. Others may seek to answer according to their impression of the popular or socially desirable response and so questionnaire design and selection of participants needs to consider these aspects carefully to ensure reliability and integrity of data (Dillman, Smyth, & Christian, 2014). There are several techniques to gather information from experts or panellists such as the Delphi and Table 19 illustrates some of these types of survey (Hasson & Keeney, 2011).

TABLE 19. DELPHI SURVEY DESIGN TYPES

Design Type	Aim	Target panellists	Administration	Number Rounds
Classical	To elicit opinion and gain consensus	Experts selected based on aims of research	Traditionally postal	Three or more rounds
Modified	Aim varies according to project design, from predicting future events to achieving consensus	Experts selected based on aims of research	Varies, postal, online etc.	May employ fewer than 3 rounds
Decision	To structure decision making and create the future in reality, rather than predicting it	Decision-makers selected according to hierarchical position and level of expertise	Varies	Varies
Policy	To generate opposing views on policy and potential resolutions.	Policy makers selected to obtain divergent opinions	Can adopt a number of formats including bringing participants together in a group meeting	Varies
Real-time consensus	To elicit opinion and gain consensus	Experts selected based on aims of research	Use of computer technology that panellists use in the same room to achieve consensus in real time rather than post	Varies
e-Delphi	Aim can vary depending on the nature of the research	Expert selection can vary depending on the aim of the research	Administration of Delphi via email or online web survey	Varies
Technological	Aim varies according to project design, from predicting future events to achieving consensus	Experts selected based on aims of research	Use of hand-held keypads allowing responses to be recorded and instant feedback provided	Varies
On-line	Aim varies according to project design, from predicting future events to achieving consensus	Experts selected based on aims of research	Implementation of the technique on any online instrument such as a chat room, or forum	Varies
Argument	To develop relevant arguments and expose underlying reasons for different opinions on a specific single issue	Panellists should represent the research issue from different perspectives	Varies	Varies
Disaggregative Policy	Constructs future scenarios in which panellists are asked about their probable and the preferable future	Expert selection can vary depending on the aim of the research	Varies	Varies

(Hasson & Keeney, 2011, p. 1697)

In this research a panel of experts was used as the approach to gather information and expert opinion on Value for Money (VFM) and applied to develop the VFM weighting factors is support of the IPMM questionnaire, as described in Section 4.3.2.

There are three types of questionnaire data described by Dillman et al. (2014) collected by questionnaires and these are:

- Factual or demographic
- Attitudes and opinions
- Behaviours and events

(Dillman et al., 2014, p. 457)

Factual and demographic is comprised of data that are readily available to the participant and may include name, dates, organisation information as required by the research and considering confidentiality and ethics factors. Attitude and opinion variables contain data that require review, consideration before response and may include questions on the organisation, process, procedures, systems, organisational structure and personal sentiments. Behaviours and events variables similarly are influenced by context and include behaviours and performance of individuals, organisations and events and can include questions on frequency, relevance and comparison (Dillman et al., 2014).

In this research the questionnaire is built from criteria derived from critical procurement process phases and the participant is invited to provide responses on factual data, opinions, behaviours and events related to the procurement criteria as it applies to the organisation. The questionnaire does not require information on behaviours and events. Zikmund et al. (2013) described three common approaches to assessing the reliability of a questionnaire: 1) test re-test, 2) internal consistency and 3) alternative form. The test re-test involves comparison of responses by sending the questionnaire twice which can be time-consuming and not always recommended as a first choice. Internal consistency involves comparison and consistency of answers to different questions within the questionnaire. An alternative form is a method of comparing different forms of the same

questions or groups of questions and it is important to use this method sparingly as participants often spot this method and become confused or fatigued by the repetition (Zikmund, Babin, Carr, & Griffin, 2013).

In the design of individual questions researchers have three approaches according to Rowley (2014), where questions used in other questionnaires can be adopted, adapted or entirely new questions can be developed (Rowley, 2014).

In the situation where there are existing questionnaires used for similar and relevant assessment purposes, questions can be adopted or adapted and on more detailed inspection of such existing questionnaires the opportunity for improvement often exists. While there is no guarantee that existing questions are very good one benefit of adopting or adapting existing questions is that they have usually been tested for reliability and validity so are more likely to yield better results than new questions (Fowler Jr, 2013).

In the case of existing PMA models explored in the literature review in this research 13 procurement assessment models in Table 11 used in international development organisations and nine common PMA models in Table 13 were reviewed for relevance to the research objectives. The questions have similarities in areas of enquiry but great differences in context, content and methods used and hence a high degree of question adaptation has been used in this research. The result is a questionnaire that is tailored to the research environment, populated with questions of established validity and reliability. Most questionnaires include both open and closed questions (De Vaus, 2013). Open questions are often used in depth or semi-structured interviews in exploratory research where there is flexibility and high degree of variability of answers.

Saunders et al. (2009) lists six different types of closed questions:

- List questions provide a list of options from which the participant can select;

- Category questions are designed to provide only one category and are useful for collecting data on behaviour and attributes;
- Ranking questions asks respondents to place items in rank-order in terms of importance or relevance;
- Rating questions, in which a rating options or scales are used to record responses;
- Quantity which requires a number or an amount in response;
- Matrix questions, where responses to two or more questions can be recorded using a table or grid.

Rating questions are frequently used to collect options and often use Likert-style rating scales where participants are requested to respond to a series of statements on a four, five, six or seven-point rating scales. Rating questions often follow a familiar pattern to ensure that participants are not confused and follow the series logic. In some case an even number of points, such as four, is selected rather than an odd number. This forces more of a choice and avoids the potential for participants to select a mid-point or neutral position and this depends on the nature of the question and research. A ten-point numeric rating scale can be used to reflect the feeling of a respondent on a scale defined by each extreme. A variation of this scale is a semantic differential rating scale where the participant is asked to rate an object or idea in a series of bipolar scales. Each bipolar scale is described with opposite adjectives and participants are requested to position their response on a scale between the extremes.

Fast - - - - - Slow

Unfriendly - - - - - Friendly

Value for Money - - - - - Over-priced

(Fowler Jr, 2013; Saunders et al., 2009, pp. 375, 381, 382)

For analysis of the data by computer the responses need to be coded for subsequent analysis; some software systems such as Survey Monkey and Qualtrics, used in this research allow for coding to be embedded into the questionnaire relatively easily. The layout can be configured to be attractive and clear and colour can be used to effect particularly if the questionnaire is administered electronically to keep printing costs manageable (Qualtrics, 2020; Survey Monkey, 2020). Open questions are more difficult to code and may need some careful thought where there are many possible responses.

For most questionnaires the early sections include list questions to establish important information such as the participant's organisation and position even if names are not required. The header for each section is also useful to orientate the participants and inform them of the section area and purpose of the questions and help with order and flow (Dillman et al., 2014; Greener, 2008; Saunders et al., 2016).

Research on the ideal length for questionnaires suggests that different and short questionnaires may be easy and convenient but suggest that the research may be superficial. However, long questionnaires may be cumbersome and reduce response rate for some participants (De Vaus, 2013). Techniques and variation in the question type such as matrix question can maintain interest, focus and reduce the apparent length of the questionnaire for the participants (Edwards et al., 2002; Zikmund, 2013).

The questionnaire includes a clear introduction section and covering letters for all the methods of distribution and administration. A message or email was also issued to inform participants or the organisation of the pending questionnaire issue. The covering letter included useful information for the participants on the requirements and purpose of the questionnaire and other relevant information such as university ethics, administration and answers to frequently-asked -questions. The questionnaire instructions also

contained information of what will happen next and to provide on how the questionnaire is used in closing the questionnaire (Dillman et al., 2014). In this research, these details are described in the research ethics and the administration included in Section 4.4.

It is important to pilot both the test the response time and ensure to create the opportunity to receive feedback which can be used to refine and improve the questionnaire. A further short questionnaire might be helpful to gather information on how long the process took, clarity, relevance and any ambiguous questions (Phellas, Bloch, & Seale, 2011).

The IPMM used in this research is presented in Appendix 1 and was aligned to the procurement process of the organisations concerned for relevance and the duration was kept to 20 to 25 minutes. The response options included single and multiple responses on a four-point Likert or multiple response scale in response the feedback from the pilots, the questions were randomised to make the coding and scoring of answers less predictable for participants. For the multiple response type questions, a single point is assigned for each answer and the respondent has the option to tick any of the options for up to four points maximum. There are five ‘net promoter’ type questions and the allocation of scores in Qualtrics assigns a score from one to four points depending on the position selected on the scale. In addition, participants of the pilots indicated some concern on the confidentiality of the questionnaire, additional instructions were included in the communications and questionnaire. The instructions, ethics information and invitations advised participants that the questionnaire was voluntary and that the information would be kept confidential by email and IPMM instructions. The head of procurement of each organisation was requested to distribute the IPMM to procurement staff and staff with appropriate knowledge of the organisation’s procurement process and systems to respond authoritatively to the questionnaire. Responses from 42 participants were received: 38 submissions from the ten international development organisations, and

four from private-sector benchmark organisations in the oil and gas industry. As recommended by Rowley (2014), when the data collection was complete, the data were cleaned and it was revealed that nine of the submissions were incomplete or unfinished, leaving 33 completed IPMM submissions (Rowley, 2014). The total sample of usable data was from 33 questionnaire results, 29 from the ten international development organisations and four from the three private sector organisations. This represents approximately 12% of procurement senior managers and experienced professionals in development organisations with relevant knowledge of the procurement process.

Brookes et al. (2014) argued that the protocol for the number of respondents for maturity assessment models in project management and procurement is not always consistent. Cooke-Davies and Ayzmanov (2003) use single respondents to represent each organisation and Mullaly (2006) used an average of respondents to arrive at a single results where there was more than one respondent from each organisation. Ibbs and Kwak (2000) and Backlund (2014) applied different methods and issued questionnaires to project managers using joint answers if required for clarity of response for difficult questions in the questionnaires. The authors argued that results tend to improve when the respondents are motivated and knowledgeable of all aspects of the criteria included in the tool (Mullaly, 2006). Grant and Pennypacker (2006) recognised that there may be limitations with the adoption of single responses that must be balanced against the requirement for knowledgeable participants. For example, Essman and du Preez (2009) received up to 30 questionnaire responses from some organisations and three or less responses from others, and combined the statistical analysis for all the responding organisations to interpret the group results. The nature of the privileges and immunities in the international development organisations, limitations on obtaining data and the



challenge in receiving multiple responses and is a possible factor that may have restricted or reduced the number of respondents in this research (Miller 2006).

#### *4.3.5 IPMM questionnaire design*

International development organisations generally use the OECD-MAPS methodology (OECD, 2010), or variations of the model, to assess the procurement performance of country procurement systems and implementing partners (OECD, 2010; World Bank Group, 2016). The most common objective is to establish and assess the capacity of implementation partners for procurement projects and activity to be performed on behalf of the development organisation and many of these models are listed in Table 11. To assess procurement performance for the international development organisation themselves it would seem logical to measure the same procurement criteria used on the OECD-MAPS models. This is because the procurement process and procedures, required by the international development organisations, are also mandatory process steps for the Implementing Partners (IPs). The OECD-MAPS methodology has been adapted by many of the international development organisations to develop models, such as the Alternative Procurement Arrangements (APA) of the WBG and ADB. Such models have differences from the OECD-MAPS model towards internal organisational procurement procedures, processes and procurement reforms. The OECD-MAPS and modified PMA models are intended and designed specifically to assess borrowing country and their Implementing Partners (IPs) procurement capability to manage donor funds (Asian Development Bank, 2018; Department for International Development (DFID), 2011; United Nations Office for Project Services, 2018; World Bank Group, 2016).

The literature review in Chapter 3 reviewed articles, academic papers and studies on PMA models in the public and private sectors; the procurement criteria included in the assessments and 27 of these authors and references are listed in Table 12. There is

considerable similarity and overlap with procurement criteria with the OECD-MAPS and related models and those included in the 27 models. However, there are several differences, and these are largely in the areas of strategy and market intelligence, customer management, systems and technology and risk performance management. The most notable difference is the increased emphasis on compliance and governance in public sector and international development organisations (OECD, 2010; World Bank Group, 2016). The private sector organisations have a greater emphasis on procurement systems, spend analysis, risk management and performance measurement (Poluha, 2007; Schuh et al., 2009). DE and VFM are under various stages of measurement, in the processes and procedures of the development organisations, and are not well reflected in the OECD-MAPS derived and adapted PMA models, used in development.

Following the literature review and analysis of the articles, papers and studies, nine common PMA models were reviewed as listed in Table 13. The models have significant differences in structure, scoring and methodology but maintain similarity, alignment and overlap in the procurement criteria. Some of the PMA models listed in Table 13, such as SCOR are widely used in supply chain management applications in the public sector while the ROSMA and APQC models have a very broad application in public sector and across industries allowing them to develop benchmark information across industries (American Productivity and Quality Center APQC, 2018). A model that has been used in the WBG and AFDB, is the Gartner (CEB) model and therefore, limited work has been carried out in procurement assessment in development organisations (Gartner, 2018).

The common procurement process phases and key procurement criteria, contained in the different articles, papers, studies and models in Table 11, Table 12 and Table 13, and discussed in Section 3.12, have been grouped and listed in Table 20. The reference for

the articles, papers and PMA models, where the procurement criteria are referenced or utilized in practice, are indicated against each of the procurement phases and criteria.

The comparative review illustrated in Table 20 demonstrates that the most widely used PMA models include the greatest range of common procurement criteria. This also shows that the international development organisations include compatible procurement criteria, as identified in this research, shown in Figure 5 in their adopted PMA models.

In particular, both the APQC and ROSMA PMA models incorporate all the criteria identified in Table 20. However, the APQC and ROSMA models include procurement criteria, such as marketing and inventory management, that are not substantial aspects of procurement in international development. In addition, these PMA models are difficult to adapt towards the assessment of the procurement process of international development organisations (American Productivity and Quality Center APQC, 2018; Kearney, 2016).

The OECD-MAPS models are primarily designed to assess Implementing Partners (IPs) or government departments handling the required procurement activity, and not to assess their own internal processes and procedures. However, if an international development organisation expects its IPs to perform against specific performance standards, this research takes the approach that the organisation will hold itself to the same performance standards expected from the IPs. There are additions to the procurement processes and criteria that are included in procurement reforms, but not yet included on OECD-MAPS derived PMA models. These include aspects such as category management, eProcurement technology, procure-to-pay, supplier risk and performance management.

TABLE 20. REFERENCES FOR KEY PROCUREMENT MATURITY CRITERIA

Reference PMA Models	Development PMA Models/Process				Common PMA Models											
	OECD MAPS	MDBs	UN	Bilateral	APQC	ROSMA	EFQM	SCOR	SCMA	CEB	EVALOG	KPMG	IPMM	360		
Key Criteria 1 <b>Planning &amp; Forecasting</b>	√	√	√	√	√	√	√	√	√	√	√	√	√	√		
Key Criteria 2 <b>Supplier Qualification</b>	*√	√	√	√	√	√	√	√	√	√	√	√	√			
Key Criteria 3 <b>Sustainable Procurement</b>		√	√	√	√	√							√			
Key Criteria 4 <b>Procurement Organisation</b>	√	√	√	√	√	√	√	√	√	√	√	√	√	√		
Key Criteria 5 <b>Customer Management</b>	**√	**√	**√	**√	√	√	√	√	√	√	√	√	√	√		
Key Criteria 6 <b>Procurement IT Systems</b>		***√		***√	√	√	√	√	√	√	√	√	√	√		
Key Criteria 7 <b>Policy &amp; Procedures</b>	√	√	√	√	√	√	√	√	√	√	√	√	√	√		
Key Criteria 8 <b>Procurement Governance</b>	****√	****√	****√	****√	√	√	√	√	√	√	√	√	√			
Key Criteria 9 <b>Risk Management</b>	√	√	√	√	√	√	√					√				
Key Criteria 10 <b>Contract Management</b>	√	√	√	√	√	√	√	√	√	√	√	√	√	√		
Key Criteria 11 <b>Contract Administration</b>	√	√	√	√	√	√	√	√	√	√	√	√	√	√		
Key Criteria 12 <b>Training &amp; Certification</b>	√	√	√	√	√	√	√	√	√	√	√	√	√	√		

Note: \* PMA model does not incorporate Market Intelligence and Supplier Management

Note: \*\* Procurement process and PMA models do not incorporate Category Management

Note: \*\*\* Procurement process and PMA models have limited detail on eProcurement and reporting tools

Note: \*\*\*\* Development procurement process and PMA models have emphasis on corruption, fraud and integrity factors

Conversely, the private sector procurement processes, procedures and popular PMA tools do not have as much emphasis on compliance and governance as their counterparts in international development organisations (Tadelis, 2012). The inclusion of the procurement criteria in Table 19 and Figure 5, across the different articles, papers, studies and models in Table 11, Table 12 and Table 13 confirms the importance of these specific procurement criteria for procurement in both the private and public sectors. As there is no current PMA model with all of these procurement criteria, as relevant for development organisations, an Integrated Procurement Maturity Model (IPMM) is developed in this research for the assessment of international development organisations.

The procurement process diagram introduced in the literature research, and shown in Figure 5, is developed in this research from common procurement processes and key procurement criteria emerging from the articles, papers, studies and models listed in Table 11, Table 12, and Table 13. The four procurement pillars, or phases, are organised by sequencing the 12 procurement criteria following the procurement process as best as possible. Secondly, by categorising some procurement criteria by their organisational and process properties. The procurement phases are represented by the procurement strategy phase, the central organisational and regulatory framework phases and the contract management phase of the procurement process.

The IPMM is comprised of the four procurement phases, 12 key procurement criteria and 37 sub-criteria and the structure is used to assess relative procurement capacity and performance of the international development organisations in the IPMM questionnaire.

The IPMM questionnaire, is adapted from the constructs and references shown in Table 19 and Figure 5, for the assessment of the procurement process of the international development organisations and comparison with private sector benchmarks.

#### *4.3.6 VFM factors and weighted IPMM questionnaire*

The structure and purpose of the IPMM questionnaire is intended to provide a comparison between selected leading international development organisation for procurement capacity and performance. The procurement criteria have been developed by reference to the standards, policy and requirements of the international development organisations and informed from articles, papers and studies including review of leading PMA models as shown in Table 13. The procurement process diagram, illustrated in Figure 5, represents the basis of the IPMM questionnaire structure, categorisation and layout. The procurement process diagram includes the 12 key procurement process criteria and 37 procurement sub-criteria. The assumption of the research is that cumulative good performance of each of the component procurement criterion results in better overall procurement performance.

The definitions used for the measurement of DE and VFM performance have similar and overlapping terms. For example, project evaluation methodology is comprised of the evaluation criteria including relevance, coherence, effectiveness, efficiency, impact and sustainability (OECD, 2019). The definition for VFM includes similar definitions, such as the 4 Es of Economy, Effectiveness, Efficiency and Equity as discussed in Section 3.4 (ICAI, 2011; Jackson, 2012). For example, the effectiveness and efficiency criteria are included in both sets of definitions and this demonstrates the connection in definitions used for procurement and project performance measurement and evaluation.

Each of the four procurement phases and 12 procurement criteria vary according to their individual contribution on procurement performance and consequently towards project performance, DE and VFM. The individual differences of 12 procurement criteria, therefore, might be expressed in terms of a relative ranking, or weights, that represent the contribution, or influence, of each of the criteria on project performance and VFM.

The IPMM results, for each of the organisations are compared by an analysis of, both the original IPMM results and IPMM results weighted for VFM, for additional sensitivity analysis of the influence of the procurement process on VFM (Mimović & Krstić, 2016).

The procurement expert panel of 12 procurement professionals from MDBs, UN agencies, professional procurement institutes and private sector organisations, shown in Table 20, was tasked to provide a percentage and rank order for each criterion. The procurement expert panel was requested to prioritise each procurement criterion by allocating a relative percentage to each procurement criteria as shown in Table 21 (Luzon, 2016; Soares et al., 2006). The resulting percentage value was used to create a relative VFM factor for each of the procurement criteria which are multiplied by the original IPMM results to create the modified weighted IPMM results as shown in Table 21 (Pizzol et al., 2017). For example, as with the case of the original IPMM procurement criteria, where criteria are considered having equal importance, the VFM factor is indicated with a weight of 1. However, in the case that the procurement expert panel considers that the criterion has double the comparative influence towards VFM, then the VFM factor is indicated by a weight of 2. Consequently, if the influence on VFM considered half the comparative influence towards VFM, then the VFM factor is indicated as 0.5, respectively. The difference between the original IPMM results and the weighted IPMM results is due to the multiplication of the original results by the VFM factors. The multiplication of VFM factors increases the weighted IPMM ratings and results for criteria which emphasise VFM in the procurement process, while it decreases the IPMM criteria rating and results for those organisations which do not have an emphasis on VFM. The technique creates two versions of the IPMM assessment results; one original, unweighted, IPMM version and a weighted IPMM version. The weighted

IPMM results serve as a sensitivity analysis and adjusts the original IPMM score to reflect the approach of the organisation's procurement process towards achieving VFM.

Mimović and Krstić (2016) and Pizzol et al. (2017) argue that weighting is a common practice used to prioritise the quality to price ratio of the criteria in order of importance, or desirability, in evaluating firms during procurement and other selection processes.

There are several terms used for this procurement practice, such as comparative analysis and economic advantage techniques, used to include evaluation of quality, value and non-cost criteria (Surasaksa & Shin, 2019). To arrive at relative priorities and weights for criteria, several methods are used including the use of expert panels selected by Delphi method or based on a panel of experts as in this research (Luzon, 2016; Quynh, 2014).

Comparative analysis and economically advantageous weighting processes are used to prioritise criteria for evaluation in the oil and gas, construction, medical and Life Cycle Impact Assessment (LCIA) industries (ISO 14044, 2022; Toutouchian et al., 2018).

The weighting technique in this research is applied to the IPMM to prioritise criteria in relation to Value for Money (VFM). The resultant weighted IPMM results help highlight the procurement processes, phases and criteria that influence and contribute to VFM.

The resulting VFM factors are shown in the following Table 21.



TABLE 21. VALUE FOR MONEY FACTORS

Procurement Phase	VFM factor	Procurement Criteria	VFM factor
Procurement Strategy	1.33	Procurement Strategy	2.14
		Procurement Qualification	1.19
		Sustainable Procurement	0.66
Procurement Organisation	0.82	Procurement Organisation	1.03
		Customer Management	0.71
		Procurement Systems	0.72
Regulatory Framework	0.99	Policy & Procedures	1.26
		Procurement Governance	1.07
		Procurement Risk	0.64
Contract Management	0.86	Contract Performance	1.23
		Contract Administration	0.58
		Training & Certification	0.78

Table 21 shows the VFM factors which are calculated from the relative percentages received from the expert procurement panel. Assuming that all the 12 procurement criteria are of equal weight then they would be allocated 8.33%. The average percentage value derived the expert procurement panel feedback, for the procurement strategy criterion is 17.79% and, division by 8.33% results in a VFM factor of 2.14. Similarly, the average percentage value for the procurement governance criterion, derived from the expert procurement panel, is 8.92% resulting in a VFM factor of 1.07, and the remaining VFM factors are calculated using the same method. The resulting weighted IPMM factors, shown in Table 21, reflect the procurement process capacity and maturity from the perspective of the relative influence of VFM on the procurement phases and criteria. To address the research second and third question, the research provides analysis of the correlation and relationship IPMM results with the project evaluation results, and by extension the relationship between procurement, project performance, DE and VFM. The definition for VFM used in this research is the ADB definition shown in and the VFM factors are developed from this definition by the procurement expert panel.

#### 4.4 ADMINISTRATION OF THE QUESTIONNAIRE

Completed questionnaires for research purposes must be administered through the selected and specified channel and means of distribution; it is also important to ensure that introduction and covering letters follow university professional code of ethics.

Questionnaires are increasingly administered by email, internet, intranet and via combinations of electronic means. There are some standards or general operating guidelines that concern the use of internet such as:

- Ensuring emails and posting to user groups are relevant and are not mis-posted as junk emails (spam);
- Remembering that invitations to participate sent to over 20 user groups at once are deemed as unacceptable by many internet vigilantes and so researchers should not exceed this threshold;
- Avoiding sending the email to multiple mailing lists as this is likely to result in individuals receiving multiple copies of your email (this is known as cross-posting);
- Avoiding the use of email attachments as these can contain viruses.

(Hewson, 2003, 2015; Saunders et al., 2009, p. 400)

Postal, collection and delivery, telephone questionnaires and structured interviews are all common methods of administering questionnaires and have challenges with ensuring good response rates, clarity and understanding. The physical and direct nature of some of these methods introduces additional challenges in reliability of response and receptiveness of the participants particularly with telephone interviews. Structured interviews and collect and delivery methods help increase response rates due to their direct nature although administration can involve resources and costs (Bell et al., 2018).

In the case of this research the direct approach to the development organisations both through email and via the internet reduces the need for much physical intervention. However, follow up and regular reminders are necessary to improve response rates. The IPMM questionnaire was sent by email as there was not the opportunity or the time to travel to each international development organisation to administer structured interviews directly. The questionnaire was circulated by email to collect primary data to all the selected international development organisations firstly through communication and agreement with the procurement leadership of the organisation (Saunders et al., 2016). Once the agreement to participate is confirmed, the leader of the organisation is requested to circulate the IPMM to the procurement staff and to ensure re-iterate that the IPMM is fully confidential, voluntary and that there is no obligation to respond. It is acknowledged that participants should not be under any obligation, coercion or influence of the management of the organisation and hence the submissions and communications between researchers and participants is confidential. The objective is to achieve over three to ten participants per organisations and a total of approximately 30 - 50 participants. The purpose is to gain expert procurement assessment feedback from knowledgeable procurement professionals in each of the development organisations. However, in practical terms it is necessary to seek permission and collaboration with the heads of procurement departments in the international development organisation and access to participants and the response to questionnaire may vary between organisations. The questionnaire was distributed and received first responses in July 2019 and the final responses were received in December 2019 with 33 of the 42 responding participants successfully completing the questionnaire.

The challenges involved with achieving participation and responses from all the organisations were significant and regular follow-up was managed through email communication and where possible, verbal conversation by phone call and discussion.

#### 4.5 DATA ANALYSIS

The analysis of the quantitative data is critical for developing the conclusions and answering the research questions. The collection of quantitative and qualitative data is designed to maximise the opportunity to perform the analysis required in the research. The type, scale and range of data is important to enable comparison and to result in meaningful analysis and results (Brown & Saunders, 2007). In addition, data can be categorised to make it easier to collect, organise, compare and perform analysis as descriptive or nominal data and categorisation can be dichotomous and ranked to facilitate the analysis and reports (Walliman, 2017). Ranked data can also be adapted and used in numerical terms and this can also be helpful in comparison. Numerical data are more quantifiable and can be compared easily. In some cases, as in this research, there are different scales which have or are adapted for regular intervals used for comparison and analysis of relationships between entities (Comiskey & Dempsey, 2013).

There are four different and recognised types of categories including interval variables, ordinal variables, nominal variables and dichotomous variables. Interval variables are a fixed space variable such as numbers of staff, volume of transactions and procurement expenditure. Ordinal variables are frequently rank ordered but the space between them is not fixed across the range. For example, if there is a scale range between 50-60Kg, 60-70Kg, 70-80Kg, 80-90Kg and over 90Kg, the last category changes the set to ordinal rather than interval variables (Greener, 2008, p.56). Nominal variables cannot be rank ordered and may be of a more subjective nature such as colour or opinion based.

The collection and analysis of quantitative data, once identified, are organised into a defined layout on many types of software which can be transferred to other systems for comparative analysis. The use of scales such as Likert scales and other matrix scales in questionnaires can be coded to enable comparison of quantitative data (De Vaus, 2013). Numerical and categorical data can both be collected in the form of numerical codes and these facilitate analysis and organisation of the data. Numerical data may require adjustment by rounding or harmonising, in the example of different currencies, and may further be grouped and categorised if required. Categorical data also require coding schemes and early or pre-set coding decisions both make later analysis more effective and accurate and careful coding assists comparability of the data. Some questionnaire software enables setting codes across all questions and assist in calculating totals and comparison of individual questions and criteria. The review for errors in coding is necessary to ensure there is no or minimal instances of data distortion, that may affect the analysis (Comiskey & Dempsey, 2013).

For secondary evaluation data collected from international development organisations, there are similar categorical and quantitative rating methods, with differences in ratings scales used, including four, five, six or seven rating scales. The scales are used to evaluate projects using the evaluation criteria described in Section 3.3. Such scales need to be coded, or normalised, to make possible comparison of data and results between organisations (Asian Development Bank. ADB, 2018; IEG, 2021). In this way, the different scales are adjusted through the use of a consistent calculation to enable comparison between sets of data with different ratings scales. The adjustment of the different six-point, five-point and four-point scales and the transformation to numerical values of the project evaluation reports in this research are shown in Appendix 3.

In addition, to further examine and understand project cost performance and how money is spent, the actual cost for each project is reviewed against the project cost estimates. Five of the ten international development organisations include secondary data on project cost estimations and actual project costs in the published project evaluation reports. A set of thresholds, or bands, relative to the cost estimate is established to gain perspective on the degree of variation from cost estimates in the research. The thresholds include bands from 0 to 10%, 10 to 25%, 25 to 50%, 50 to 100% and more than 100%, for actual project cost above and below the cost estimate shown in Section 5.2.10.

#### *4.5.1 Normalisation of the different ratings scales*

Organisations using a four-point scale include ADB, ISB, EBRD, UNICEF and JICA for overall ratings, while WBG, AFDB, IADB and UNDP use a six-point rating scale. As mentioned, DFID is the only organisation using a five-point scale shown in Figure 9.

The normalised scores have some differences and distortions and, particularly, the five-point scale, and consequently, the mean result for DFID has a reduction from first to the tenth rank as a result of the normalised scores and the different five-point rating scale (Dawes, 2008). However, for the purposes of this research, the normalised ratings calculation has been used, shown in Table 23, to enable comparative analysis of the results. The ratings scales for six and four-point scales are defined in Table 22.

TABLE 22. RATINGS SCALES (SIX-POINT AND FOUR-POINT SCALES)

Six-Point Scale	Description	Four-Point Scale	Description
Highly Satisfactory (HS) - 6	All parameters fully met and no shortcomings	Highly Satisfactory (HS) - 4	Expectations exceeded
Satisfactory (S) - 5	All parameters fully met with minor shortcomings	Satisfactory (S) - 3	Successful with no major shortfall
Partially Satisfactory (PS) - 4	Parameters partially met with minor shortcomings in the evaluation report	Partially Satisfactory (PS) - 2	Major benefits achieved with significant shortfalls
Partially Unsatisfactory (PU) - 3	More than one parameter was unmet, and there were major shortcomings	NA	NA
Unsatisfactory (U) - 2	Most parameters were not, and there were major shortcomings	Unsatisfactory (U) - 1	Project considered a technical and economic failure
Highly Unsatisfactory (HU) - 1	None of the parameters were met, and there were severe shortcomings	NA	NA

(Asian Development Bank, 2014; United Nations Development Program, 2019)

The rating scale for DFID is shown in Figure 9.

A++ = Outcome substantially exceeded expectations

A+ = Outcome moderately exceeded expectations.

A = Outcome met expectation.

B = Outcome moderately did not meet the expectation.

C = Outcome substantially did not meet the expectation.

FIGURE 9. DEPARTMENT FOR INTERNATIONAL DEVELOPMENT RATING SCALE

(Department for International Development, 2012)

In order to make it possible to compare the different scales for different organisations, the scores are normalised for comparison according to the following formula:

$$N_{orm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

(Leung, 2011)

The calculated normalised scores are shown in Table 23.

TABLE 23. PROJECT EVALUATION NORMALISED RATINGS

6-point scale	Normalised 6-point scale	5-Point scale	Normalised 5-point scale	4-point scale	Normalised 4-point scale
HS	1	A++	1	HS	1
S	0.800	A+	0.750	S	0.666
PS	0.600	A	0.500	PS	0.333
PU	0.400	B	0.250	U	0
U	0.200	C	0		
HU	0				

*Note: Highly Satisfactory (HS), Satisfactory (S), Partially Satisfactory (PS), Partially Unsatisfactory (PU), Unsatisfactory (U), Highly Unsatisfactory (HU)*

As discussed later in Section 5.2.2, the different rating scales introduce distortions when the evaluation results are compared for each development organisation (Dawes, 2008).

The adjusted normalised ratings allow for more meaningful comparison, correlation, and regression tests between the project evaluation mean results (Jajuga & Walesiak, 2000).

The DFID rating scale is defined differently from the other organisations, for example, the rating of A, which is the equivalent of satisfactory in the four and six-point scales, is defined by “outcome met expectations” in the DFID five-point scale (Department for International Development, 2012). DFID also splits its satisfactory equivalent and above ratings into three levels of A++, A+ and A. As a result, and to manage this anomaly, A++ and A+ are treated as equivalent to highly satisfactory, thus increasing the percentage of highly satisfactory ratings in Figure 14 and Table 35. The treatment for the five-point scale results has the effect of increasing the percentage results for DFID, and order to manage distortions of the different rating scales, a normalisation equation is used to calculate adjusted mean results as shown in Table 23.

An alternative normalisation equation was identified in the literature review and applied to the rating scales (IBM, 2020). In addition, the same statistical analysis, correlation and regression tests were performed on the project evaluation results using the alternative normalisation equation, and a comparison of the results is shown in Appendix 3. Both normalisation equations have similar calculations based on proportions. The comparison



of the equations is to determine whether the different equation produces different results, rank order and conclusions. The alternative transformation equation is expressed as:

$$Y = (B - A) * (x - a) / (b - a) + A$$

In this transformation equation, A and a refers to the minimum score of the scale and B and b refers to the maximum score and used to transform any value from one scale to the other and the different results are summarised and shown in Appendix 3 (IBM, 2020).

The calculated mean results for the alternative normalisation equation show slightly higher values, but the rank order and statistical analysis results do not significantly change. Similarly, the correlation and regression results have slightly different values, but the statistical findings and conclusions do not change.

In summary, the alternative transformation equation results have minor differences due to the slightly different method of rescaling the different rating scales. However, the statistical analysis and all other results indicate the same findings and conclusions.

Therefore, while the different transformation equations lead to slightly different values, the statistical results, findings and conclusions remain the same as shown in Appendix 3.

#### **4.6 STATISTICAL TESTING AND ANALYSIS**

Statistics are a vital part of data analysis for research questions and objectives and describe variables from the perspective of central tendency or the dispersion of the data and information under analysis. Central tendency is useful to understand data commonality in the sense of values that occur most frequently, the mode, mid value or mid-point, the median, or the calculation of the average, the mean (Brown & Saunders, 2007). The numerical coding scheme for a questionnaire, for example, can be calculated in terms of mean scores for each respondent and displayed for comparison by category, sub-group or for analysis of the individual questions. International development

organisations evaluate project with six performance criteria and derive an overall performance rating by consolidating or calculating a total rating from the criteria.

In addition, to analysis by central tendency it is very useful to understand how the data are dispersed around the mode, median or mean and the two main ways this is measures is by inter-quartile range or standard deviation. De Vaus (2013) described the most basic range as the difference between highest and lowest value and additional information is provided when the data are divided into quartiles known as the inter-quartile range.

Other possible range intervals can be in the form of percentiles and deciles. To indicate and analyse the extent to how the data are dispersed around the mean the standard deviation is used. An important factor is developing standard deviation calculations is have the relative spread of different distributions in a comparable form and this is often achieved by dividing the standard deviation by the mean to arrive at the coefficient of variation. To compare dispersion around the mean for different organisations once the data are in comparable form, each organisation can be compared against a base value to create an individual index (De Vaus, 2013).

To address the research questions and objectives, the relationships between different variables are often core components of the research design. It is prudent to examine whether the relationship could occur by chance or if there is a correlation and this is called significance testing. If there is no significance it may be concluded that the relationship is purely random and testing is necessary to verify and understand the nature of statistical relationships (Bezzina & Saunders, 2014). In the example of procurement maturity, previous studies on the relationship with performance, has been found to be significant by process of interviews (Lockamy III & McCormack, 2004).

Testing significance is facilitated by analytical software and Jenkins and Anderson (2003) advise that it is important to understand how the formulae are working and the hypothesis being tested. The establishment of the hypotheses and the variables under consideration should relate closely to the research questions and objectives. It is very difficult to arrive at conclusions when testing with small sample sizes, and conversely, very large sample sizes can make the test over sensitive and it is necessary to collect an appropriate sample size for the research, especially with multiple variables (Jenkins & Anderson, 2003). In common with analysis of similar questionnaires, statistical analysis is performed on the IPMM results largely by analysis of the median and means providing useful information for comparison of the organisations and processes (Murray, 2013). The limitations of statistical analysis of Likert type scales requires care and consideration for the interpretation of results. Such consideration is acknowledged in this research and the statistical analysis provides helpful comparative analysis of the IPMM and project evaluation results (Boone & Boone, 2012). For the project evaluation results, statistical analysis is used in addition to correlation and regression tests which are conducted for the normalised continuous data, as discussed in Section 4.5.1 (Parker et al., 2013). The different statistical methods to examine relationships, differences and trends are summarised in Table 24.

TABLE 24. SUMMARY OF RELATIONSHIPS, DIFFERENCES AND TRENDS BY DATA TYPE

Applied to Research Questions (RQ)	Categorical		Numerical	
	Descriptive	Ranked	Continuous	Discrete
The test whether two variables are associated:	Chi-Square (data may need grouping)		Chi-Square if variables grouped into discrete classes	
RQ3: Chi-squared ( $\chi^2$ ) test of independence. To test association for cost variation data.	Cramer's V Phi (both variables must be dichotomous)			
To test whether two groups (categories) are different:		Kolmogorov-Smirnov (data may need grouping) or Mann-Whitney <i>U</i> test	Independent t-test or paired t-test (often used to test for changes over time) or Mann-Whitney <i>U</i> test (where data skewed or a small sample)	
RQ1: Kolmogorov-Smirnov to test IPMM results RQ2: Kolmogorov-Smirnov to test project evaluation results				
To test whether three or more groups (categories) are different:			Analysis of variance (ANOVA)	
RQ2: To test different organisations (categories) for project evaluation results				
To assess the strength of relationship between two variables:		Spearman's rank correlation coefficient (Spearman's rho) or Kendall's rank order correlation coefficient (Kendall's tau)	Pearson's product moment correlation coefficient (PMCC)	
RQ2: To test cost variation and corruption data for correlation RQ3: To test project evaluation results for correlation				
To assess the strength of a relationship between one dependent and one independent variable:			Coefficient of determination (regression coefficient)	
To assess the strength of a relationship between one dependent and two or more independent variables:			Coefficient of multiple determination (multiple regression coefficient)	
To predict the value of a dependent variable from one or more independent variables:			Regression analysis (regression equation)	
RQ3: To understand predictors for project evaluation results RQ4: To understand how individual evaluation criteria contribute to project evaluation results				
To examine the relative (trend) over time:			Index numbers	
To compare relative changes (trends) over time:			Index numbers	
To determine the trend over time of a series of data:			Time series: moving averages or Regression equation (regression analysis)	

Adapted from Saunders et al. (2009, p. 451)

To test if two variables are associated the chi square ( $\chi^2$ ) test is used and this test depends on the observed values expected where the variables are independent from each other. In other words, the null hypothesis, or proposition that there is no significant difference between two or more variables, is applicable. The observations for each category should align to a single category or class interval and the Chi-Squared ( $\chi^2$ ) test calculates that the relationship could occur by chance. In terms of the numerical expression of results, a probability of 0.05 means that there is only a 5% chance of the data occurring by chance is this is considered statistically significant (Hair Jr et al., 2015). Cramer's V is sometimes calculated by statistical software simultaneously and measures the association of between two variables on a scale between 0 and 1 and 0 represents no association and 1 represent perfect association. Phi is another alternative that is most often used for comparison of dichotomous variables (Quinlan, Babin, Carr, & Griffin, 2019).

For testing whether two groups are different and whether a group of observations of a category differ from a specified sample, the Kolmogorov-Smirnov and *t*-test are examples of frequently used tests. For example, project performance results in the health sector, may have different results and characteristics from the other ten development sectors. The Kolmogorov-Smirnov test is an appropriate test for ranked data of this kind, and used in this research to compare the cumulative data from the observed values from the specified population (Kanji, 2006).

For numerical data that can be defined in two groups the *t*-test provides an effective statistical method to compare the means of the two groups. A statistically significant result is indicated by a large *t* statistic which has a probability of  $p < 0.05$ . For data that measure the same sector such as health for different years it is possible to use a paired *t*-test. The *t*-test assume a normal distribution and for skewed data that or small sample sizes, the Mann-Whitney *U* test is an appropriate alternative (Dancey & Reidy, 2007).

For different sectors of development, in the research, such as health, education, transportation and agriculture if the numerical variable has three or more groups, difference can be assessed using a one-way ANOVA. In the ANOVA test, if the likelihood of difference between sectors is low is represented by a large F ratio with a probability of  $p < 0.05$  it is considered statistically significant.

A useful statistical test to assess the strength of the relationship or correlation is the correlation coefficient to measure the relationship between ranked or numerical variables. A value of +1 represents positive correlation and -1 represents negative correlation and 0 represents perfect independence. For numerical data the Pearson's Product Moment Correlation Coefficient (PMCC) to measure the strength of the relationship. In case one or more of the variables contain ranked data, a correlation coefficient is required calculated from the ranked data. The two most widely used rank correlation coefficients are Spearman's rank correlation coefficient (Spearman's rho) and Kendall's rank correlation coefficient (Kendall's tau) and these statistical tests are interpreted in the same way (Hair Jr et al., 2015).

In the case of international development data from IPMM, scores are derived in a numerical form and can be ranked for comparison. Numerical data on individual procurement criteria can be ranked for comparison purposes and it is possible to select any of the above statistical tests depending on the purpose of the analysis for correlation.

To measure the relationship between IPMM results for an international organisation (the dependent variable) and one or more of the project evaluation data results (independent variable) the coefficient of determination or regression coefficient may be applied. The coefficient of determination ( $r^2$ ) measures the proportion of the variation between the dependent variable and one or more independent variables with values between 0 and +1.

For the calculation of regression equations using two or more independent variables multiple regression analysis is applied and software is often used to enable complex calculation and analysis (Curtis, Comiskey, & Dempsey, 2016).

This research conducts and explores the correlation between procurement performance of international development organisations and project performance by organisation, region country and sector. Data collection includes both numerical and ranked data and there are a wide variety of statistical tests that are applied for analysis of data.

#### **4.7 RESEARCH VALIDITY AND INTEGRITY**

The data for the IPMM model are multi-qualitative and collected with the use of a questionnaire designed to gather responses directly from each of the selected international development organisations. The relative procurement performance is potentially an indicator of good practice, cost performance and efficiency. The comparative assessment of procurement in leading international development organisations provides evidence on the effectiveness, efficiency and Value for Money (VFM) of the disbursement of development funds by sector, geography and organisation.

The data for project performance for Development Effectiveness (DE) used by development organisations is drawn from disclosed project monitoring and evaluation completion reports. The research performs comparative analysis to understand the relationship between the IPMM results and project performance and DE by organisation, region, country and sector. The overarching research objective is to understand and evaluate the correlation of IPMM procurement capacity and the organisation self-evaluated project performance in international development organisations. The data are non-identifiable and is encrypted and stored in a secure environment for seven years after completion of the research and back-ups are made on a weekly basis. Ethics approval was granted by Curtin University's research office (number HRE2018-0649).

#### **4.8 PRESENTATION OF DATA**

There are numerous ways to present data in the form of tables, graphs and diagrams in a way that explains and represents the data as effectively as possible. The simplest way to represent data is by table useful for frequency distribution or categorisation and relationships between columns and rows that reflect research questions and objectives. An effective way to present highest and lowest values and comparative value by group is with bar charts and pictograms such as histograms with categorical data to present data as simply and clearly as possible. Another common display of extreme values is the box plot, which shows the relationship of the mid median point and quartiles of the distribution of data. For trends, it is helpful to present data in the form of a line graph which gives a representative across time. Histograms and line graphs can also be used in combination where the comparison of values across time need emphasis as a compound bar chart. Pie charts are another type of diagram that show proportionate segments is a clear and simple manner. The distribution of values can be distributed in many ways as polygon or histogram bar structures with a positive or negative skew. The flatness or gradient of the curve is referred to as kurtosis, and a negative flat curve is known as platykurtic, and a more extreme curve is mesokurtic (Doane & Seward, 2011).

Diagrams for the comparison of two or more variable, in this research, such as trends multiple line graphs in international development data are frequently applied to display continuous data and trends. Comparison of totals, values and variables can be represented by stacked bar charts, comparative proportional pie charts, compound scatter diagrams and various types of box plot (Walliman, 2017).

#### **4.9 APPROACH TO DATA ANALYSIS IN THIS RESEARCH**

This data analysis in this research includes statistical frequency analysis by comparison of the means, median, standard deviations, kurtosis and skewness for both primary and



secondary data. Due to the large sample of ( $n = 1,920$ ) project evaluations, the statistical comparison of groups was performed using the Kolmogorov-Smirnov, Shapiro-Wilks and Bonferroni tests and the analysis of variances by the ANOVA test. These tests are used to answer the first and second research questions to compare procurement maturity and project evaluation methodology used in international development organisations. The Pearson’s correlation test and regression analysis are conducted to answer the third and fourth research questions on the relationship between the procurement and project management processes and to explore how the procurement influences DE and VFM in international development projects.

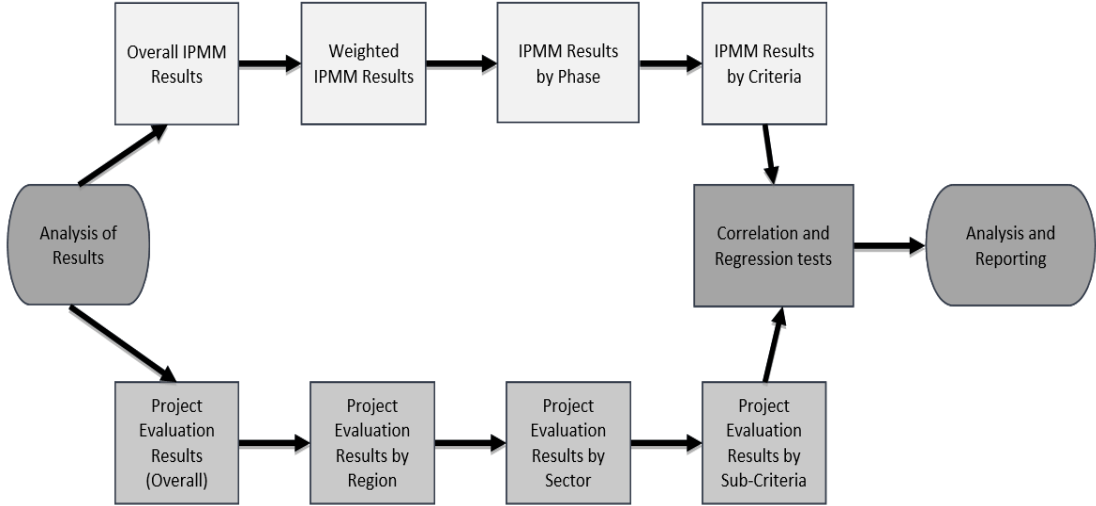


FIGURE 10. ANALYSIS OF RESULTS

The analysis of the results in answer to the research questions is discussed in Chapter 5 and illustrated in Figure 10 and Figure 11.

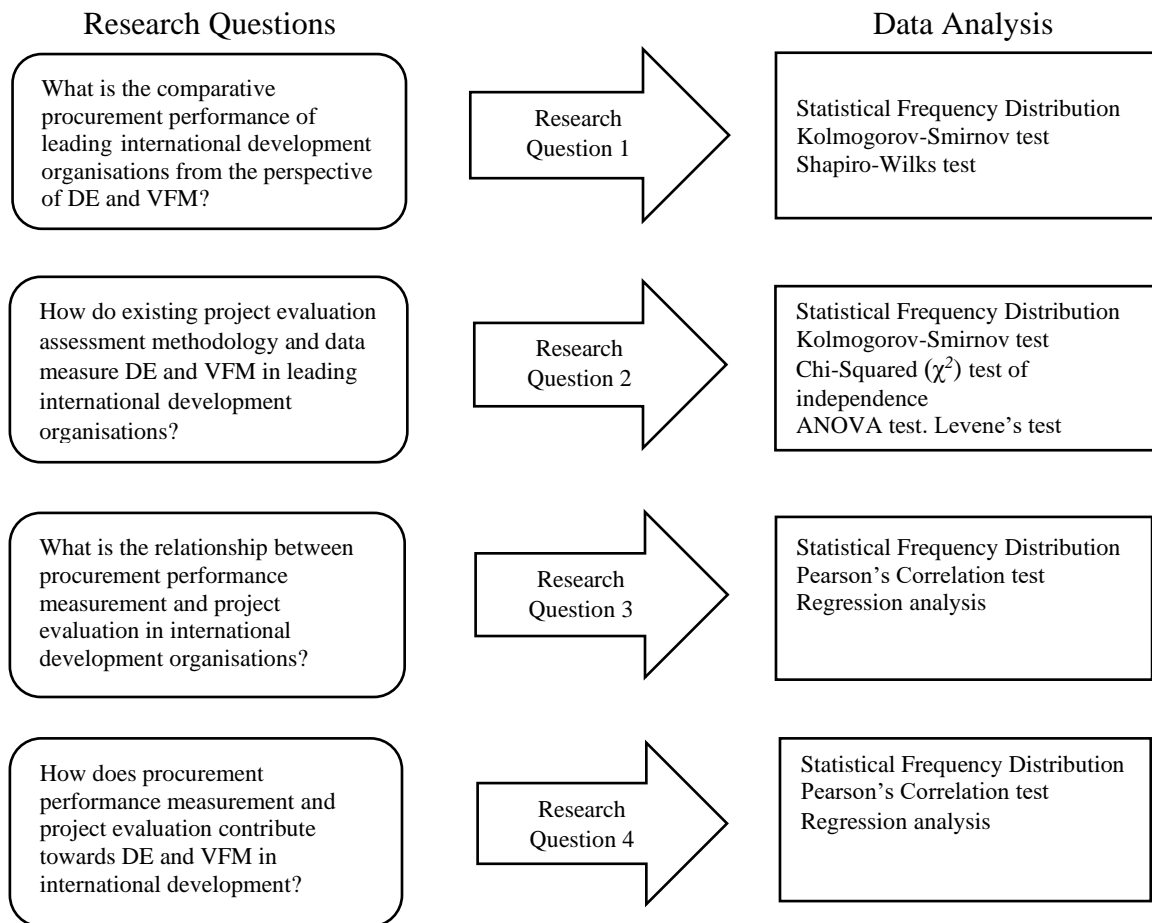


FIGURE 11. DATA ANALYSIS

Both the IPMM questionnaire and project evaluation reports include qualitative data and use rating scales to convert responses to numerical quantitative values. This research follows a pragmatism philosophy, comprised of qualitative data and quantitative data. Qualitative data are collected from the IPMM questionnaire results, procurement expert panel inputs and project evaluation reports. The quantitative data are also derived from the IPMM questionnaire, project evaluation reports and the Likert-type and multiple response ratings for all sources are converted to quantitative data for the purposes of analysis and comparison of results in this research.

Chapter 5 follows the discussion on research methodology and described the analysis and results of this research.

## **5 RESULTS**

Chapter 4 outlines and justifies the research design applied to the generation of the IPMM and the secondary data for project evaluation for international development organisations. The three phases of the research design are explained, and details are provided on the methodological choices made to conduct the research. This chapter is dedicated to presenting the results of the research. It is structured to address the first research question: to understand relative procurement performance of international development organisations from the perspective of DE and VFM. Secondly, to address the second research question: to review a sample of 1,920 project evaluations to examine how existing project evaluation assessment methodology measures DE and VFM in leading international development organisations.

In response to the third research question, analysis was conducted on the correlation and relationship between procurement performance measurement and project performance evaluation in international development organisations. Finally, to address the fourth research question, this research provides further analysis on the procurement performance measurement, project evaluation criteria and cost data to understand how these processes contribute towards DE and VFM in international development projects. The chapter follows the link between procurement, DE and VFM in international development projects to understand the influence of procurement on project success and effective development outcomes.

### **5.1 PROCUREMENT MATURITY RESULTS**

#### *5.1.1 Introduction*

Chapter 5 applies the research design outlined in the previous chapter for analysis of the IPMM procurement maturity results for the selected international development organisations and three private sector organisations for comparison and benchmarking.

Raschke and Ingraham (2010) noted that there is often an expectation that as maturity increases there is an impact on procurement performance and that procurement maturity has positive effect on process outcomes. Aulia and Isvara (2021) argued that procurement maturity is an indicator for procurement performance and that the concept of procurement maturity is correlated with improvement of the procurement process.

The IPMM scores are compared against each of the international development organisation's key procurement criteria for both original and weighted IPMM results. The original IPMM assessment results are modified by the VFM factors to develop the weighted IPMM results and these enable analysis and comparison of IPMM results to explore the relative influence of the procurement process towards VFM. The VFM factors are developed from the consultation with the procurement expert panel, which prioritises and ranks the procurement criteria according to their relative influence and contribution towards VFM in the procurement process. The research provides analysis and comparison of the results by organisation with the procurement maturity results from three private sector organisations to benchmark maturity results with the private sector.

In addition, analysis is provided to examine how each of the four procurement phases and criteria, shown in Figure 5, individually contribute towards the procurement maturity and VFM. In this way, the more mature components and criteria of the procurement process are identified and categorised to be incorporated and potentially promote best practices for procurement and VFM in international development organisations.

#### *5.1.2 Integrated procurement maturity model questionnaire results*

As discussed in Chapter 4, the IPMM was sent to ten international development organisations, including six MDBs, two UN agencies, two bilateral development organisations and three private sector organisations. The IPMM questionnaire was also

sent to three comparable private sector organisations in the oil and gas industry to provide benchmark information on maturity for comparative purposes.

The results of the total questionnaire ratings of each organisation for the 12 procurement criteria of the IPMM are therefore represented in Table 25, and shown in Figure 12.

The total questionnaire ratings are consolidated for each organisation, and the results and rank order is represented in Table 25. The results in Table 25 and Figure 12, show the original IPMM results, and provide a comparison of the individual procurement criteria and the total of the results for each of the international organisations.

The totals, percentage and rank order shown in Table 25, are the results of the participants' responses for each question for the four procurement phases and 12 key criteria that make up the IPMM questionnaire Appendix 1. The responses to the questions are based on a Likert, or multiple response, four-point scale and the results are the summary of the total responses for all participants for each organisation. For example, as discussed in Chapter 3, and illustrated in Figure 5, the procurement strategy phase comprises of three key criteria: 1) the procurement strategy criterion, 2) the supplier qualification and eligibility criterion, and 3) the sustainable procurement criterion. In the same way, the procurement organisation phase, regulatory framework phase and contract management phase are each comprised of three of the 12 key criteria represented in Table 25 and Figure 12. The numbers represent the average score for all respondents by criteria and the numbers in bold represent the average score by phase.

The overall results are also shown in terms of rank order to provide a comparison of the IPMM results between organisations and follows procurement maturity assessment models and benchmarking techniques used in the leading PMAs (Schuh et al., 2017).

TABLE 25. OVERALL IPMM QUESTIONNAIRE TOTAL SCORES

Phase Criteria	MULTILATERAL DEVELOPMENT BANKS						UN AGENCIES		BILATERALS	
	WBG	ADB	AFDB	EBRD	IADB	ISDB	• UNDP	UNICEF	• DFID	JICA
<b>Procurement Strategy</b>	<b>53.7</b>	<b>43.5</b>	<b>54.1</b>	<b>36</b>	<b>35</b>	<b>34</b>	<b>45.5</b>	<b>51.7</b>	<b>52</b>	<b>45</b>
<i>Procurement Strategy</i>	22.7	20	22.5	16	15.5	19	18	21.7	20	16
<i>Supplier Qualification &amp; Eligibility</i>	12.1	10	12.6	8	8	7	11.5	13.3	12	7
<i>Sustainable Procurement</i>	18.9	13.5	19	12	11.5	8	16	16.7	20	22
<b>Procurement Organisation</b>	<b>35.6</b>	<b>43</b>	<b>39.4</b>	<b>28</b>	<b>37.5</b>	<b>31</b>	<b>39.5</b>	<b>43</b>	<b>49</b>	<b>42</b>
<i>Procurement Organisation</i>	15.6	19	15.9	13	15.5	14	14	17	17	18
<i>Customer Management</i>	13.2	12	13.6	9	13.5	11	14	15	19	15
<i>Procurement Systems</i>	6.8	12	9.9	6	8.5	6	11.5	11	13	9
<b>Regulatory Framework</b>	<b>58</b>	<b>55</b>	<b>55.6</b>	<b>43</b>	<b>51</b>	<b>44</b>	<b>48</b>	<b>54.3</b>	<b>55</b>	<b>51</b>
<i>Policy &amp; Procedures</i>	22.4	20.5	20.4	16	19.5	18	17.5	20	18	18
<i>Procurement Governance</i>	28.7	30.5	28.6	23	26	20	25	27.7	31	27
<i>Procurement Risk</i>	6.9	4	6.6	4	5.5	6	5.5	6.6	6	6
<b>Contract Management</b>	<b>31.2</b>	<b>31.5</b>	<b>29.7</b>	<b>21</b>	<b>29</b>	<b>17</b>	<b>25.5</b>	<b>29.7</b>	<b>33</b>	<b>25</b>
<i>Contract Performance</i>	7.7	7.5	10	7	8	5	7.5	8.8	9	7
<i>Contract Administration</i>	14.1	16.5	12.3	10	12.5	8	11.5	14.5	13	12
<i>Training &amp; Certification</i>	9.4	7.5	7.4	4	8.5	4	6.5	6.3	11	6
<b>Total</b>	<b>178.5</b>	<b>173</b>	<b>178.8</b>	<b>128</b>	<b>152.5</b>	<b>126</b>	<b>158.5</b>	<b>178.7</b>	<b>189</b>	<b>163</b>
%	75.6	73.3	75.7	54.2	64.6	53.3	67.2	75.7	80	69
Rank	4	5	2	9	8	10	7	3	1	6

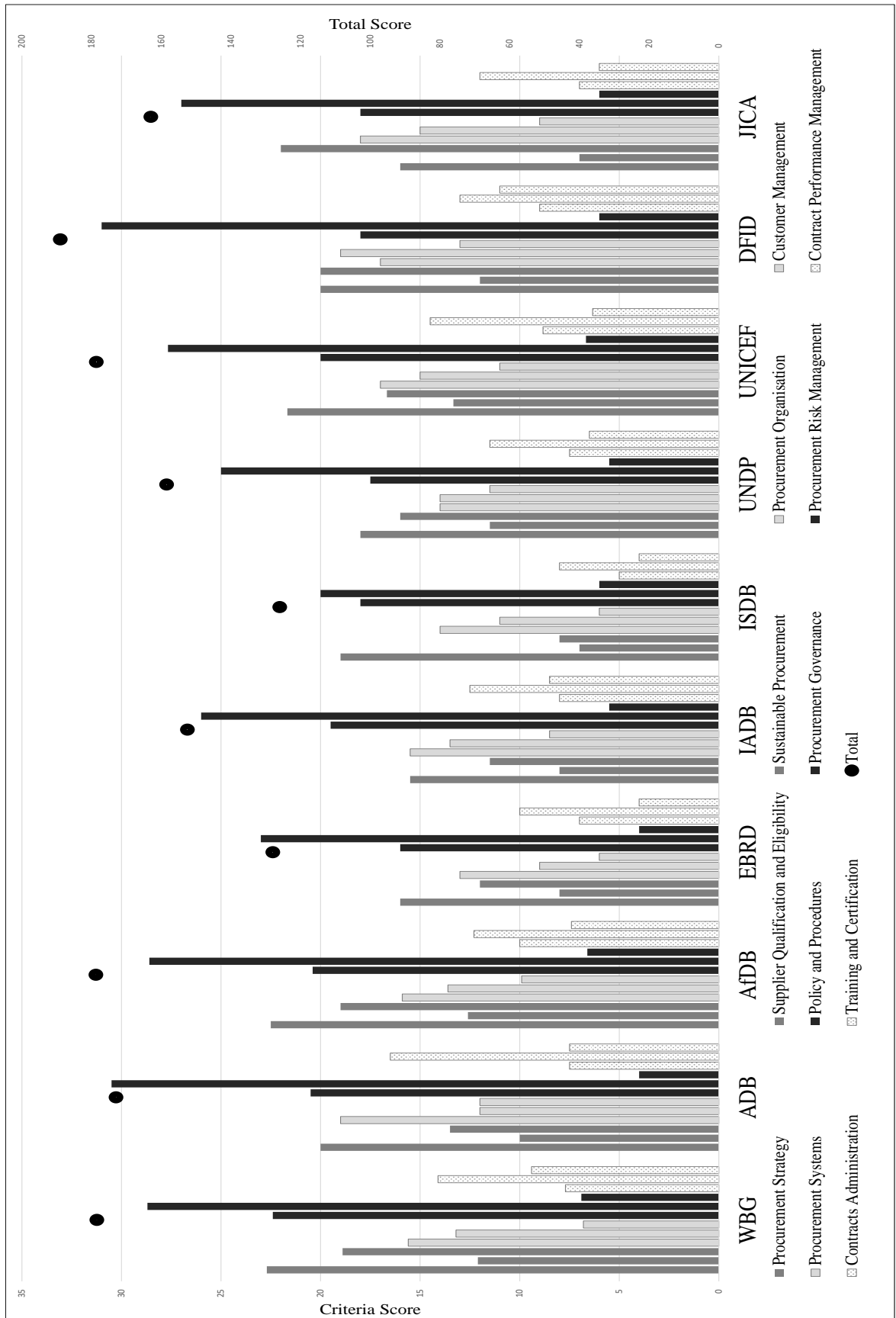


FIGURE 12. OVERALL IPMM RESULTS

The five highest-ranked organisations include DFID, AFDB, UNICEF, WBG and ADB in rank order and, at the other end of the comparison, the lowest results include JICA, UNDP, IADB, EBRD and ISDB. The higher results indicate a comparatively strong level of procurement maturity for each international development organisation. DFID falls into the integrated procurement maturity level four, in Figure 4. While AFDB, UNICEF and WBG are fall into level four, they are on the cusp between integrated procurement maturity and the managed procurement maturity level three. UNDP falls under level three but very close to achieving level four. The other five international development organisations including JICA, UNDP, IADB, EBRD and ISDB fall into the managed procurement maturity level represented by level three by rank order.

It is important to note that, in common with such questionnaires, results are self-reported responses from the procurement group of the organisation. This should be taken into account when interpreting the results as response bias, observational bias and social desirability bias are likely to be strong (Clements et al., 2008; Denizer et al., 2013).

The IPMM procurement phase results show that for international development organisations, the regulatory framework phase has the highest maximum result, and the procurement strategy phase has the second-highest result. The results suggest and imply that the strategic and regulatory criteria are considered as more important for procurement maturity by international development organisations (Findley, Milner, & Nielson, 2017). For the other two phases, the procurement organisation phase is third-ranked, and the contract management phase has the lowest rank by procurement phase. Figure 12 shows the overall IPMM results in chart form, and in this chart, the columns represent the results of the four procurement phases and the 12 procurement criteria for each international development organisation in this research. The black bullet points



represents the total IPMM results, and this represents the overall result of the combined criteria results for each international development organisation.

The individual IPMM results for each criterion are shown in Figure 12, and the procurement phases are grouped and represented by the same colour and shade of grey. The highest rated procurement phase is the regulatory framework phase, which is comprised of high ratings for the three individual criteria that represent the governance, risk and compliance aspects of the procurement process. The three criteria that make up the regulatory framework phase include the procurement policy and procedures criterion, procurement governance criterion and the procurement risk management criterion.

The IPMM results in Table 25 compares ratings by phase and criteria and show the breakdown by phase and criteria and that the regulatory framework appears to be in place and working efficiently for the MDBs, UN agencies, and bilateral development organisations. There are several possible explanations for this prominence of the regulatory framework in the IPMM results. For example, the relative importance of procurement governance and regulation in development is due to safeguard against the perceived potential for corruption in many developing countries (Findley et al., 2017).

The emphasis on procurement governance and regulation is reflected by the emphasis on the regulatory framework criteria of the procurement process, including policy and procedures, governance and procurement risk criteria by international development organisations. The procurement strategy phase has the second highest result, and indicates the importance of procurement strategy and planning across international development organisations and will be examined later in this research (Mensah, 2016).

### *5.1.3 Statistical distribution for IPMM results*

The frequency distribution for the IPMM results provides statistical analysis for each participating organisation, the four procurement phases and 12 criteria of the IPMM.

The mean results are used to rank and compare procurement maturity results.

Both the original scores and the mean results provide an opportunity to compare the IPMM results of the international development organisations. The mean results and statistical distribution frequency data are shown in Table 26, and are used to explore and perform further analysis of the comparative data and results in this research.

As discussed in Section 4.3.4, the comparison and analysis of the organisations, phases and criteria are performed using the means, medians and standard deviations. It is acknowledged that, while the statistical analysis of IPMM results provides helpful and useful comparative information, consideration is required in the interpretation of results due to the low frequency of responses by some of the organisations in this research.

TABLE 26. STATISTICAL DISTRIBUTION DATA FOR IPMM RESULTS

<i>Statistic</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	Skewness	Kurtosis	<i>p</i> -value	Rank
WBG	3.025	3	0.802	-0.724	0.239	0.000	4
ADB	2.932	3	0.958	-0.337	-0.103	0.000	5
AFDB	3.030	3	0.751	-0.749	0.381	0.000	2
EBRD	2.169	2	0.791	0.117	-0.551	0.000	9
IADB	2.585	3	0.981	-0.020	-1.010	0.000	8
ISDB	2.136	2	0.937	0.502	-0.536	0.000	10
UNDP	2.686	3	0.698	-0.468	0.102	0.000	7
UNICEF	3.028	3.16	0.503	-0.842	0.836	0.002	3
DFID	3.203	3	0.805	-1.213	1.735	0.000	1
JICA	2.763	3	1.135	-0.538	-1.100	0.000	6

The mean results in Table 26, show same rank order as the original results, in Table 25.

As discussed, DFID, AFDB, UNICEF, WBG and ADB have the top five mean IPMM results. The results indicate that DFID, AFDB, UNICEF, WBG and ADB, have comparatively strong procurement processes with the highest IPMM mean results. These are DFID ( $M = 3.203$ ,  $SD = 0.805$ ), AFDB ( $M = 3.030$ ,  $SD = 0.751$ ), UNICEF ( $M = 3.028$ ,  $SD = 0.503$ ), WBG ( $M = 3.025$ ,  $SD = 0.802$ ) and ADB ( $M = 2.932$ ,  $SD = 0.058$ ).

At the other end of the scale, IADB, EBRD and ISDB, have comparatively weak procurement processes with the low IPMM mean results, IADB ( $M = 2.585$ ,  $SD = 0.981$ ), EBRD ( $M = 2.169$ ,  $SD = 0.791$ ), and ISDB with ( $M = 2.136$ ,  $SD = 0.937$ ).

Most organisations achieved a median result ( $Mdn = 3$ ), but the lowest-ranked organisations (EBRD and ISDB) have a median result ( $Mdn = 2$ ), which reflects the lower mean results for these two organisations. The other exception is for UNICEF, which has a differently expressed median result ( $Mdn = 3.16$ ), resulting from the calculation of the combination and consolidation of six UNICEF participants.

The mean and median results indicate that each organisation has different strengths and weaknesses across the procurement processes, phases and the procurement criteria.

The standard deviations show a similar spread of values around the mean with values from 0.75 – 1.00 across most organisations. The largest standard deviation is shown for JICA ( $M = 2.763$ ,  $SD = 1.13$ ), and this indicates JICA scores more evenly across all the criteria of the IPMM. JICA, therefore demonstrates that it maintains a procurement process with consistent maturity across all of the key procurement criteria. UNICEF has the lowest standard deviation of ( $M = 3.028$ ,  $SD = 0.503$ ) and shows a flatter more consistent pattern of procurement criteria scores across the range of responses, as shown in Figure 12. The flat pattern suggests that UNICEF has a consistently high procurement maturity level by comparison with other organisations for its procurement processes.

DFID, UNICEF and WBG show the greatest negative skewness, indicating clusters of high scored procurement criteria and ranking first, third and fourth respectively. IADB, EBRD and ISDB have larger positive skewness indicating more frequent lower-rated criteria, and their resultant ranking is reflected as eighth, ninth and tenth respectively.

This indicates a lower relative capacity and maturity of the procurement process for these comparatively lower ranked organisations. DFID has the highest kurtosis result, and

UNICEF has the second-highest kurtosis due to the concentration and consistency of results around and above the mean, reflected in their relatively high-rank order of first and fourth for procurement maturity, respectively. In practice the leptokurtic kurtosis results suggest that DFID and UNICEF have high performing procurement processes and therefore, rank relatively high across all of the procurement criteria (Sharp et al.,2019).

Tests of normality for both Kolmogorov-Smirnov and Shapiro-Wilks, for all organisation results, are significant at the  $p < 0.05$  level, respectively, as shown in Table 26. The results indicate that the distribution of IPMM scores for all the organisations does not follow a normal distribution. The Kruskal-Wallis test was performed as a non-parametric cross check, due to the use of Likert-based and multiple response ratings in the IPMM questionnaire, and shows a significant result at the  $p < 0.05$  level, and confirms the results of both the Kolmogorov-Smirnov and Shapiro-Wilks tests.

Given the low frequency ( $n$ ) of participants in the test in comparison of results from some of the organisations, a significant result is fully expected. Despite the limitation of frequency of participants, the results represent participant responses from each of the international development organisations as discussed in Section 4.3.4. The interpretation of the IPMM results is presented by analysis of the frequency distribution and means. Together, these results suggest a significant difference between the IPMM procurement capacity and [maturity](#) of the different international development organisations. More specifically, EBRD, IADB and ISDB have a different and lower procurement capacity and maturity profile reflected in the  $M$  and  $SD$  results. An important finding is that there were differences between organisations, procurement phases and criteria of the IPMM results, and these will be explored further in this chapter and discussed in Chapter 6.

#### *5.1.4 Weighted IPMM questionnaire results*

The next phase of the inquiry was a validation step of the results using feedback from a procurement expert panel to rank the procurement criteria by percentage according to relative influence of each of the criteria on VFM in the procurement process.

Following the input from the expert panel to validate and prioritise the criteria of the IPMM, Table 27 displays the weighted IPMM results and represents the adjusted IPMM results. As discussed in Section 4.3.6, the results reflect the influence of VFM on the procurement process, as determined by the VFM factors developed in consultation with procurement expert panel, shown in Table 21. Following methods used in the resources sector for weighting evaluation criteria in the procurement process (Banda, 2019; Mimović & Krstić, 2016), the weighted IPMM results are calculated by multiplying the original IPMM results with the VFM factors, as shown in Table 27. As discussed in Section 4.3.6, similar weighting techniques are used to prioritise criteria for evaluation in several industries, such as oil and gas, construction, medical equipment and in the automotive industry (Stilger & Van Raaij, 2017; Toutouchian et al., 2018).

In this research the original IPMM results and criteria are multiplied by the VFM factors as a sensitivity analysis and the original (unweighted) IPMM results are compared with the weighted IPMM results. The weighted results effectively recalculate the respondents scores to provide a new set of results from the perspective of Value for Money (VFM). This new set of weighted results, therefore, provides an indication of organisations that apply procurement processes and criteria that have the greatest impact towards VFM. To seek answers to the first research question and understand the maturity of leading international development organisations from the perspective of VFM, the sensitivity analysis introduces a technique to assess procurement maturity and influence of VFM.

The overall weighted IPMM results are illustrated in Figure 13, and the columns represent the four procurement phases and 12 procurement criteria, and the horizontal line represents the total IPMM results for each international development organisation. Organisations that have a relative increase in overall weighted IPMM results from the original IPMM results, indicates that they have a greater emphasis on the procurement criteria that have a positive influence on VFM, within the procurement process. Conversely, organisations that have a decrease in overall weighted IPMM results, place a lower priority for criteria that influence VFM within the procurement process. Additional gap analysis is performed to understand the comparison between the procurement phases and the relative influence of each phase towards VFM. In this example, the findings demonstrate a useful method for procurement function to identify opportunities for the improvement to achieve VFM in the procurement process. The weighting method can potentially also be adapted to explore other areas for improvement in the procurement process. For example, factors can be adapted to identify, prioritise and weight for other phenomena, such as, corruption. In this way, further research may utilise corruption factors, rather than the VFM factors used in this research. In the example of corruption factors, the results can then be used to identify and prioritise opportunities for reduction of corruption in the procurement of goods and services for development projects.

TABLE 27. OVERALL WEIGHTED IPMM RESULTS

Phase <i>Criteria (wt)</i>	MULTILATERAL DEVELOPMENT BANKS						UN AGENCIES		BILATERALS	
	WBG	ADB	AFDB	EBRD	IADB	ISDB	• UNDP	UNICEF	• DFID	JICA
<b>Procurement Strategy</b>	<b>75.5</b>	<b>63.6</b>	<b>75.7</b>	<b>51.9</b>	<b>50.3</b>	<b>54.3</b>	<b>62.8</b>	<b>73.2</b>	<b>70.3</b>	<b>57.1</b>
<i>Procurement Strategy</i>	48.6	42.8	48.2	34.2	33.2	40.7	38.5	46.4	42.8	34.2
<i>Supplier Qualification &amp; Eligibility</i>	14.4	11.9	15	9.5	9.5	8.3	13.7	15.9	14.3	8.3
<i>Sustainable Procurement</i>	12.5	8.9	12.5	7.9	7.6	5.3	10.6	11	13.2	14.5
<b>Procurement Organisation</b>	<b>30.4</b>	<b>36.7</b>	<b>34.4</b>	<b>24.1</b>	<b>31.7</b>	<b>26.5</b>	<b>32.6</b>	<b>36.1</b>	<b>40.4</b>	<b>35.7</b>
<i>Procurement Organisation</i>	16.1	19.6	16.4	13.4	16	14.4	14.4	17.5	17.5	18.5
<i>Customer Management</i>	9.4	8.5	9.7	6.4	9.6	7.8	9.9	10.7	13.5	10.7
<i>Procurement Systems</i>	4.9	8.6	8.3	4.3	6.1	4.3	8.3	7.9	9.4	6.5
<b>Regulatory Framework</b>	<b>63.3</b>	<b>61</b>	<b>59.4</b>	<b>47.3</b>	<b>55.9</b>	<b>47.9</b>	<b>52.3</b>	<b>59.1</b>	<b>59.7</b>	<b>55.4</b>
<i>Policy &amp; Procedures</i>	28.2	25.8	24.6	20.2	24.6	22.7	22.1	25.2	22.7	22.7
<i>Procurement Governance</i>	30.7	32.6	30.6	24.6	27.8	21.4	26.8	29.6	33.2	28.9
<i>Procurement Risk</i>	4.4	2.6	4.2	2.6	3.5	3.8	3.5	4.3	3.8	3.8
<b>Contract Management</b>	<b>25.1</b>	<b>24.7</b>	<b>25.2</b>	<b>17.6</b>	<b>23.7</b>	<b>13.9</b>	<b>21</b>	<b>24.3</b>	<b>27.3</b>	<b>20.3</b>
<i>Contract Performance</i>	9.5	9.2	12.3	8.6	9.8	6.2	9.2	10.9	11.1	8.6
<i>Contract Administration</i>	8.2	9.6	7.1	5.8	7.2	4.6	6.7	8.4	7.5	7
<i>Training &amp; Certification</i>	7.4	5.9	5.8	3.2	6.7	3.2	5.1	5	8.7	4.7
<b>Total</b>	<b>194.3</b>	<b>186</b>	<b>194.7</b>	<b>140.7</b>	<b>161.6</b>	<b>142.6</b>	<b>168.8</b>	<b>192.7</b>	<b>197.6</b>	<b>169</b>
%	77.4	74.1	77.6	56	64.4	56.8	67.2	76.7	78.7	67.1
Rank	3	5	2	10	8	9	6	4	1	7

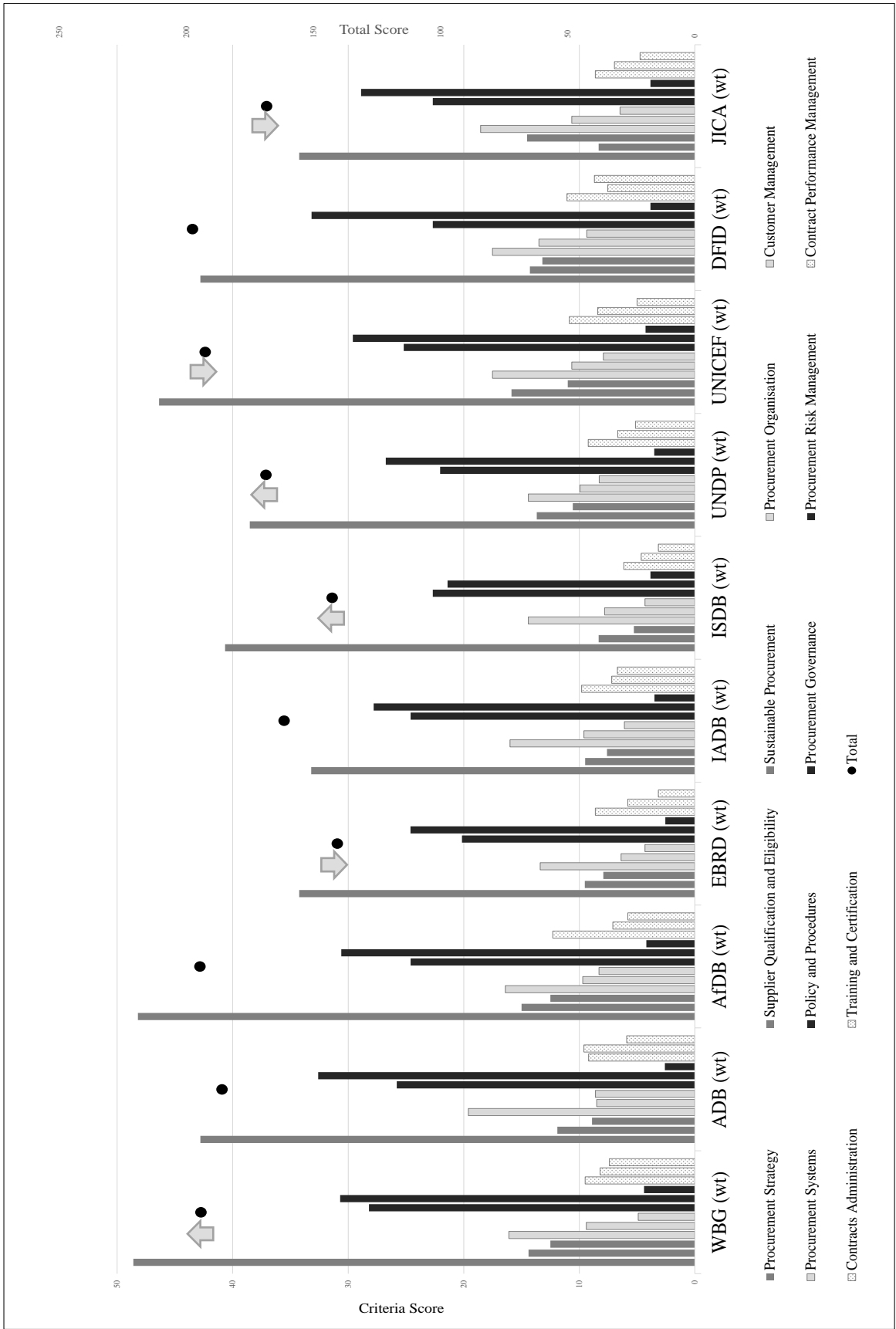


FIGURE 13. OVERALL WEIGHTED IPMM RESULTS



For the weighted IPMM results, DFID, AFDB, WBG and UNICEF all fall into the integrated procurement maturity level four, as shown in Figure 4. The other international development organisations fall into the managed procurement maturity level three.

For the individual organisations, DFID and AFDB again emerge as the first and second-ranked organisations respectively, and WBG changes by dropping to fourth from third-ranked overall. UNICEF dropping in rank from third to fourth, as shown in Table 27.

At the lower end of the rank order, there are other changes in results, with UNDP rising from seventh to sixth ranked, JICA dropping from sixth to seventh, EBRD changes from ninth to tenth, followed by ISDB changing from tenth to ninth in rank order. This suggests that WBG maintains a procurement process that gives priority to criteria that promote and influence VFM and hence rises by one rank. Conversely, it suggests that UNICEF does not prioritise the criteria for VFM and hence UNICEF drops by one rank.

There are changes in rank order for weighted IPMM results in all cases except for DFID, AFDB, ADB and IADB, which remain at the same rank, and this illustrates the influence of the VFM factors after multiplication with the original IPMM results. For weighted IPMM results, DFID shows the highest weighted IPMM result and rank and this reflects that DFID has a robust procurement process with an additional emphasis on the criteria that influence and contribute towards VFM. In summary, the weighted IPMM results, shown in Table 27, show the international development organisations' results by rank order of DFID, AFDB, WBG, UNICEF, ADB, UNDP, JICA, IADB, ISDB and EBRD.

A resulting change in overall rank order indicates the comparative procurement capacity and maturity to promote VFM by the different international development organisations.

The weighted IPMM results were also reviewed by the procurement phase and criteria and show that the procurement strategy phase and regulatory framework phase have the

highest relative weighted IPMM results by phase. To remind the reader, the overall weighted IPMM results reflect the procurement phases and criteria that have the greatest influence for VFM in the procurement process. The changes in rank order from the original IPMM results to the weighted IPMM results indicate organisations that prioritise procurement processes that potentially have a greater impact on VFM. For example, DFID scores highly for the phases that include procurement strategy, procurement governance and procurement risk criteria, all of which have high VFM factors. DFID was the first ranked for unweighted IPMM results and, due to strong maturity ratings for criteria with high VFM factors, was also first ranked for the weighted IPMM results.

The consistent first rank indicates that DFID' procurement process is at the integrated maturity level and DFID prioritises the phases and criteria that promote VFM in projects including strategy planning procurement criteria and procurement governance criteria.

#### *5.1.5 Statistical distribution for the weighted IPMM results*

The frequency distribution for the IPMM results provides statistical analysis for each participating organisation, the four procurement phases and 12 criteria of the IPMM.

In a similar manner with the statistical results for unweighted IPMM results, shown in Table 26, the statistical weighted IPMM statistical distribution results are used to rank the organisations in descending order by mean as shown in Table 28.

The mean, median, standard deviation and frequency data from the IPMM are used for the review and comparison of the international development organisations and private sector organisations for this research.

TABLE 28. STATISTICAL DISTRIBUTION DATA FOR WEIGHTED IPMM RESULTS

<i>Statistic</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	Skewness	Kurtosis	<i>p</i> -value	Rank
WBG	3.292	2.990	1.845	1.225	1.295	.000	3
ADB	3.154	2.840	1.789	1.300	2.114	.000	5
AFDB	3.298	3.190	1.827	1.322	1.681	.001	2
EBRD	2.384	2.060	1.478	1.086	0.848	.003	10
IADB	2.740	2.380	1.555	1.207	2.145	.000	8
ISDB	2.418	2.140	1.790	1.482	1.946	.000	9
UNDP	2.860	2.380	1.502	1.427	2.103	.008	6
UNICEF	3.265	3.075	1.592	1.146	1.224	.020	4
DFID	3.349	3.090	1.558	1.013	1.334	.027	1
JICA	2.855	2.640	1.737	1.693	3.963	.007	7

As discussed in the previous section, the mean results all increase between 0.1 and 0.3 and the changes in the average rating adjust the rank order as discussed in the previous section. The median results also change both positively and negatively, thereby indicating that the VFM factors have an effect on the skewness, as shown in Table 28. The changes in rank order for weighted IPMM results are reflected by the increase or decrease of the mean result after multiplication by the VFM factors. As discussed, an increase in the mean and rank indicates that the procurement capacity and maturity for the organisation, procurement phase and criteria have a greater influence and impact towards achieving VFM in the procurement process when adjusted for VFM factors. The standard deviations show a similar spread of values around the mean with values (1.478 to 1.845) across the organisations. The largest standard deviation is shown with WBG ( $SD = 1.845$ ), this suggests WBG scores more evenly across all the criteria of the weighted IPMM so has a relatively stable procurement maturity. At the other end of the spectrum of results, EBRD has the lowest standard deviation ( $SD = 1.478$ ), and reflects a low, flatter, yet more consistent pattern of IPMM results across the range of responses for procurement criteria and this is reflected in Figure 13.

JICA, ISDB and UNDP have positive skewness results and the rank order for these organisations was seventh and ninth and sixth, respectively. The positive skewness indicates a concentration of ratings below the median and reflects the relatively low rank. JICA, IADB and ADB have the highest positive kurtosis results are all  $> 1.0$ , which indicates the concentration and consistency of the results around and above the mean. The rank order for these organisations was seventh and eighth and fourth, respectively. In practice, the leptokurtic kurtosis weighted IPMM results is not an indication that JICA, IADB and ADB have comparatively high procurement maturity but rather reflects the concentration or consistent cluster of similar ratings against the IPMM criteria. Tests of normality for both Kolmogorov-Smirnov and Shapiro-Wilks have a significant result, at the  $p < 0.05$  level, for all organisations, shown in Table 28. The Kruskal-Wallis test shows a significant variation in IPMM results  $H(9) = 38.74, p = .000$ , and shows a significant result at the  $p < 0.05$  level, and confirms the results of both the Kolmogorov-Smirnov and Shapiro-Wilks tests. As discussed in Section 4.3.4, the results represent participant responses as collected from each of the international development organisations. The significant result is not surprising and can be largely attributed to the small sample size. Similar to the unweighted IPMM results, the interpretation of the weighted results was presented from analysis of the frequency distribution and means. The next section provides an analysis of the difference between the two data collection activities and explains the results. Analysis is provided on the change in rank order from between unweighted and weighted IPMM results, reflecting the relative procurement maturity of the organisations towards achieving VFM. In addition, the gap analysis provides the change in results for the procurement phases and criteria and indicates the relative influence of process criteria towards VFM in the procurement process.

### 5.1.6 Comparison between original and weighted IPMM scores

The IPMM questionnaire results provide a comparative score and rank order for procurement capacity and maturity of the international development organisations across the four procurement phases and 12 criteria. The comparison of the original IPMM results with the weighted IPMM results shown in Table 29, is intended to illustrate the relative priority and influence of the VFM factors on the procurement processes.

In the original IPMM results, the procurement phases and criteria are considered of equal importance and do not have any relative scale or distinction by weight. The weighted IPMM results are calculated by multiplication of the original results by the VFM factors, as discussed and shown in Table 28. As discussed in Section 4.3.6, after multiplication, the results for the four procurement phases and the 12 criteria were adjusted for their relative importance, or influence, for VFM in the procurement process and towards project outcomes. The net differences and changes in rank for all the selected international development organisations are shown in Table 29 below.

TABLE 29. GAP ANALYSIS BETWEEN ORIGINAL AND WEIGHTED IPMM RESULTS

PHASE	WBG	ADB	AFDB	EBRD	IADB	ISDB	UNDP	UNICEF	DFID	JICA	Gap
Procurement Strategy	21.8	20.1	21.4	15.9	15.3	20.3	17.3	21.5	18.3	12.1	184
Procurement Organisation	(5.2)	(5.2)	(5)	(3.9)	(5.8)	(4.5)	(6.9)	(6.9)	(8.6)	(6.3)	(58.3)
Regulatory Framework	5.3	5	3.8	4.3	4.9	3.9	4.3	4.8	4.7	4.4	45.4
Contract Management	(6.1)	(6.8)	(4.5)	(3.4)	(5.3)	(3.1)	(4.5)	(5.4)	(5.7)	(4.7)	(49.5)
<i>Total Gap</i>	<i>15.8</i>	<i>13.1</i>	<i>15.7</i>	<i>12.9</i>	<i>9.1</i>	<i>16.6</i>	<i>10.2</i>	<i>14</i>	<i>8.7</i>	<i>5.5</i>	<i>-</i>
Rank Change	4 to 3	5	2	9 to 10	8	10 to 9	7 to 6	3 to 4	1	6 to 7	-

The gap analysis for weighted IPMM results for procurement phases and criteria is shown in Table 29, and provides another perspective on the influence of the procurement phases and criteria towards VFM across the international development organisations. The gap analysis shows that ISDB has the largest change from original unweighted IPMM results to weighted IPMM results of an increase of 16.6 followed by WBG with 15.8 and then AFDB with 15.7. The results indicate the total net difference between the original and weighted IPMM results and differences in results and rank for the organisations and the procurement phases. The net increase for ISDB and WBG increases the rank position from tenth to ninth and fourth to the third rank respectively while AFDB does not change in rank position. The increase in rank is the result of the multiplication and alignment of the procurement criteria with positive VFM factors that are considered to have a greater influence on VFM in the procurement process. In the case of DFID, which has a smaller net increase, and is first ranked for the original IPMM results, there is no change from original IPMM results to the weighted IPMM results. The smaller net difference indicates that DFID has good procurement capacity and maturity and also exhibits and emphasises for procurement criteria that contribute towards VFM. Correspondingly, the change of overall ranking for WBG from fourth rank, for the original IPMM results, to the third rank for weighted IPMM results, indicates similar sound emphasis for procurement criteria that contribute towards VFM. Other changes include increases for UNDP from seventh to sixth, UNICEF decreases from third to fourth and JICA decreases from sixth to seventh by ranked position. The procurement phase with the greatest influence on VFM, derived from the procurement expert panel feedback, is the procurement strategy phase, as indicated in Table 21. The procurement strategy phase is comprised of the procurement strategy criterion, followed by the procurement qualifications and eligibility criterion, and

sustainability criterion. The VFM factors differ for each of the procurement criteria; some are positive, and others are negative. The multiplication by the factors, therefore, increases or decreases the resulting weighted IPMM results, accordingly. The other procurement phases include the procurement organisation, the regulatory framework and the contract management phases, which have lower calculated VFM factors.

Procurement criteria that have negative factors are not necessarily an indication that these procurement criteria are not important, but rather, that they have less impact on VFM and contribution through the procurement process towards project outcomes. The IPMM results for individual organisations and for procurement phases and criteria, indicate the relative procurement maturity, capacity and maturity. The weighted IPMM results indicate the contribution and influence of the procurement process towards VFM. The following section explores the comparison between the three different types of organisation; MDBs, UN Agencies and bilateral development organisations.

#### *5.1.7 Comparison of the three types of international organisations*

Table 30 shows the statistical frequency distributions for original and weighted results for international development organisations grouped by type of organisation, which includes MDBs, UN Agencies, bilateral development and private sector organisations. The mean scores for unweighted and weighted IPMM results for different types of organisations, including private sector benchmarks, shows that the private sector has the highest original and weighted IPMM results. This is followed by the bilateral development organisations, UN agencies and then the MDBs. The unweighted IPMM results and weighted VFM results follow the same rank order for all organisation types.

TABLE 30. FREQUENCY DISTRIBUTION FOR IPMM RESULTS BY ORGANISATION TYPE

<i>Statistic</i>	<i>M</i>	<i>SD</i>	Skewness	Kurtosis	Rank
<u>MDB</u>					
IPMM Original Results	2.745	0.936	-0.287	-0.846	4
IPMM Results (wt.)	2.977	1.764	1.261	1.725	(4)
<u>UN Agency</u>					
IPMM Original Results	2.857	0.628	-0.758	0.543	3
IPMM Results (wt.)	3.063	1.554	1.249	1.443	(3)
<u>Bilateral Development</u>					
IPMM Original Results	2.983	1.004	-0.893	-0.187	2
IPMM Results (wt.)	3.103	1.662	1.301	2.370	(2)
<u>Private Sector</u>					
IPMM original Results	3.042	0.763	-0.737	0.289	1
IPMM Results (wt.)	3.254	1.675	1.181	1.320	(1)

The mean, median and standard deviations indicate that after multiplication of the original IPMM results by the VFM factors, the procurement maturity profile of the weighted IPMM results becomes more aligned with less variation by organisation type. The change in statistical distribution for the weighted IPMM results suggests that the four organisation types have procurement process characteristics and have a focus on procurement criteria that contribute favourably towards VFM. However, each of the organisation's procurement processes have individual differences and peculiarities. The comparison between UN agencies and bilateral development organisations, which may reflect the similar procurement organisational structure and functional approach in both these organisation types. The statistically significant result between the MDBs and the private sector organisations indicates a difference in the procurement approach and orientation of these two types of organisations. The difference may be explained by the observation that the operational procurement group of the MDBs primarily have an oversight role and prioritise governance, transparency and adherence to the regulatory



framework. On the other hand, private sector organisations have a greater emphasis on procurement strategy, innovation and customer management, technology and training criteria (Arlbjørn & Freytag, 2012; Tadelis, 2012).

The mean results and rank order indicate that for the unweighted IPMM results, there are differences of procurement maturity between all the four types of organisation as shown in Table 30. For the weighted IPMM results, the mean results, positive kurtosis and positive skewness implies that there is a common approach and alignment to VFM in the procurement processes of all four organisation types. The positive procurement maturity results and rank order of the private sector suggests that there are procurement best practices beneficial for international development organisations (Manyathi et al., 2021). The following section explores private sector procurement results in more detail.

#### *5.1.8 Comparison of IPMM results; private sector and development organisations*

Thus far, comparisons have been made between MBDs, UN agencies and bilateral organisations. This section is dedicated to broadening the comparison in IPMM results for international development organisations in Table 30 to private sector organisations in the oil and gas industry. The oil and gas industry has several similarities to international development. These include organisational structure, geographical distribution, similar scale and scope of infrastructure projects, and the requirement to implement multiple projects of different types in remote and fragile countries (Hassan, 2013; Kaiser, 2022). For benchmarking and comparison purposes, the IPMM results for three private sector organisations were collated and compared with the development organisations.

Table 31 shows the results for three private sector organisations in the oil and gas industry for both weighted and unweighted IPMM results and the breakdown by procurement phase.

TABLE 31. IPMM RESULTS FOR PRIVATE SECTOR (OIL AND GAS) ORGANISATIONS

Organisation IPMM	Shell		Woodside		Chevron	
	Original	(Weighted)	Original	(Weighted)	Original	(Weighted)
<b><u>PHASE</u></b>						
Procurement Strategy	60	81.6	56	77.2	47	67
Procurement Organisation	43	35.1	42.5	35.3	39	32.3
Regulatory Framework	57	61.3	56	60.7	48	51
Contract Management	29	24.4	30	24.4	31	25.7
Total Score	189	202.4	184.5	197.6	165	176
<i>M</i>	<b>3.20</b>	<b>3.43</b>	3.12	3.35	2.80	2.98
Rank by <i>M</i>	1	1	2	2	6	6

Shell and Woodside have the highest results for original IPMM results, and are first and second-ranked, and Chevron is ranked sixth for original IPMM results when placed in comparison with the results in Table 25. In the case of the weighted IPMM results, the results are the same and Shell is ranked in first place, Woodside ranked second, and Chevron ranked sixth compared to the results in Table 27. Shell and Woodside are both at the integrated procurement maturity level four, and Chevron falls in the managed procurement maturity level three, shown in Figure 4, and the results for all three organisations are the same for both unweighted and weighted IPMM results.

The private sector organisations also score highly by the procurement phases, particularly for the procurement strategy phase, and comparatively well for all the other phases across the four phases and 12 criteria. Compared with the international development organisations, the private sector organisations score highly for the procurement strategy and procurement organisation phases and the regulatory framework and contract management phases. The key difference is that private sector organisations

emphasise the strategic procurement criteria, such as customer management and sustainability and invest more in systems and technology (Manyathi et al., 2021). In contrast, the results suggest that the international development organisations have a greater emphasis on the governance and compliance-oriented requirements of the international development procurement process (Patrucco et al., 2019; Tadelis, 2012).

#### *5.1.9 Comparison of procurement groups in development organisations*

Now that the original and weighted IPMM results and a comparison of private sector benchmarks have been investigated, an analysis of procurement maturity by procurement group will be provided to give insight into organisation structure. Procurement departments of the MDBs have a distinction between what operational; procurement-oriented towards development objectives, and corporate procurement; mainly responsible for the direct or internal procurement requirements of the organisation or institution. The operational procurement departments of the MDBs perform oversight and governance functions for the borrowing government agencies, and Implementing Partners (IPs), that are delegated to perform the direct procurement activities (Pallas & Wood, 2009). Table 32 shows the comparison between operational and corporate procurement groups for MDBs. The purpose of comparing the procurement groups is to explore the similarities or differences in capacity and maturity of the procurement groups in general and between the operational procurement and corporate procurement groups within the MDBs.

TABLE 32. FREQUENCY DISTRIBUTION FOR IPMM RESULTS BY PROCUREMENT GROUP

<i>Statistic</i>	<i>M</i>	<i>SD</i>	Skewness	Kurtosis	Rank
<u>All Procurement Groups</u>					
IPMM Original Results	2.920	0.838	-0.817	0.259	1
IPMM Results (wt.)	3.083	1.606	1.274	1.930	(1)
<u>Operational Procurement</u>					
IPMM Original Results	2.654	0.931	-0.143	-0.907	3
IPMM Results (wt.)	2.915	1.810	1.259	1.518	(3)
<u>Corporate Procurement</u>					
IPMM Original	2.881	0.929	-0.521	-0.582	2
IPMM Results (wt.)	3.070	1.692	1.302	2.221	(2)

The MDBs are distinctive in that they have an operations procurement group, which provides oversight for procurement performed by appointed implementing partners of the Low-Income Countries (LICs) and Middle-Income Countries (MICs). MDBs also have dedicated corporate procurement groups, which manage direct internal procurement requirements for the organisation and for WBG which includes the procurement of consulting firms and services for operations in developing countries. These organisational structures and functions make MDBs of particular interest because of the challenges involved with the management and oversight of third parties that are responsible for the procurement process (Ahsan, 2012). In the MDBs, the operational procurement groups manage an oversight function over the appointed Implementing Partners (IPs) engaged either directly by the MDB or by the recipient countries; operational procurement groups rarely manage procurement activities directly. By comparison, the UN agencies and bilateral development groups do not have this separation and distinction between operational and corporate procurement and they manage all procurement required by the organisation (Asian Development Bank, 2022; Department for International Development (DFID), 2018; World Bank Group, 2016).

In Table 21, the mean original and weighted IPMM results for the different procurement groups show that the procurement group, ‘All Procurement Groups’ (represented by both the UN agencies and bilateral organisations) have comparatively better performing procurement processes. The next best performing group is the Corporate Procurement group, followed by the Operational Procurement group of the MDBs. The weighted IPMM results follow the same rank order for procurement maturity as the original IPMM results. The operational procurement group of the MDBs has a fundamental difference in that its principal function is one of oversight and governance of procurement activity managed directly by borrowing countries or by their IPs on behalf of the developing country. The higher results for All Procurement Groups and Corporate Procurement suggest that practical experience may have advantages for capacity and maturity from experience gained by engagement in direct procurement activity (Gladilina, 2017).

Secondly, the lower comparative mean results for the MDBs may be an indication of the challenges with identifying and managing Implementing Partners (IPs) with equivalent required procurement capacity and maturity in developing countries (Gulrajani, 2016).

The mean, median and standard deviation results suggest that procurement groups have variable strengths and weaknesses across the procurement phases and criteria for their respective procurement processes. This finding points to a need for a more consistent approach to procurement processes. In addition the finding supports the argument made by Gladilina (2017) on the value of gaining direct practical procurement experience and expertise for improving competency of procurement professionals (Gladilina, 2017).

The results support and confirm the earlier observation that there are advantages where procurement staff have direct practical experience executing and managing the procurement function for development projects (Gladilina, 2017). Secondly, the results may also reflect the challenge for the MDBs, with oversight of the host country

Implementing Partners (IPs). The IPs manage procurement activity for the donors with oversight from operational procurement group (Ahsan & Paul, 2018). Now that we have reviewed the organisational characteristics and maturity of procurement organisations, we will explore the procurement phases and criteria of the procurement process itself.

#### 5.1.10 Analysis of the four procurement phases

The previous analysis considered procurement functions of the MDBs. This section broadens the analysis to investigate frequency distributions of the four procurement phases. Table 33 shows the original IPMM and weighted IPMM results and statistical distribution for the four procurement phases.

TABLE 33. FREQUENCY DISTRIBUTION FOR IPMM RESULTS BY PROCUREMENT PHASE

<i>Statistic</i>	<i>M</i>	<i>SD</i>	Skewness	Kurtosis	Rank
<u>Procurement Strategy</u>					
IPMM Original Results	2.676	0.926	-0.296	-0.762	3
IPMM Results (wt.)	3.779	2.430	-0.762	0.738	(1)
<u>Procurement Organisation</u>					
IPMM Original Results	2.773	0.939	-0.345	-0.786	2
IPMM Results (wt.)	2.343	1.009	0.234	-0.731	(3)
<u>Regulatory Framework</u>					
IPMM Original	3.087	0.754	-0.642	-0.088	1
IPMM Results (wt.)	3.376	1.075	-0.244	-0.643	(2)
<u>Contract Management</u>					
IPMM Original Results	2.557	0.968	-0.179	-0.954	4
IPMM Results (wt.)	2.092	1.095	0.799	-0.021	(4)

The first ranked procurement phase for original results is the procurement regulatory framework phase, followed by the procurement organisation phase, the procurement strategy phase and finally, the contract management phase. International development organisations have an emphasis on strong governance and compliance procurement procedures; this is a response to operating in developing countries and the prevalence of weak infrastructure and perceived corruption (Pallas & Wood, 2009). The procurement

strategy phase is ranked third and the contract management phase is ranked fourth in the original IPMM results. The low rank for contract management is in common with many public sector organisations, and the contract management phase is often recognised as an area for improvement across many organisations (Aulia & Isvara, 2021; Rendon, 2010). The weighted IPMM results change the rank order, and the procurement strategy phase ranks first followed by the procurement regulatory framework phase, the procurement organisation phase and the contract management phase. The implications of the weighted IPMM results are that the procurement strategy phase and procurement regulatory framework phase have a greater influence on VFM than the other phases. Therefore, international development organisations with strong procurement capacity and maturity in strategic and regulatory phases indicates that they have processes that support VFM. The results show that there are significant differences in procurement maturity by phase for the international development organisations for the original and weighted IPMM results. The implication of the finding highlights the importance of the procurement phases on procurement maturity and VFM in international development organisations. This result indicates that the international development organisation have variations in IPMM scores and that there is a statistically significant difference between procurement phases of the procurement processes of the international development organisations. The comparatively high weighted mean result of the procurement strategy phase and the regulatory framework implies that there is some alignment between the critical strategic procurement phases, the governance phases and VFM in the procurement process. The consistently lower result for the contract management phase, in comparison to the other three procurement phases of the procurement process, and the low rank indicates a reduced prioritisation towards the contract management phase. The results confirms the

observation that the international development organisations prioritise the procurement governance criteria of the regulatory framework. However, the implication of the results is that international development organisations should prioritise the procurement strategy and procurement organisation phases to improve effective project success and VFM. Now that we have reviewed the four procurement phases, the following section provides some analysis of the results by the 12 procurement criteria.

#### *5.1.11 Analysis of the 12 procurement criteria*

The final stage of scrutinising the IPMM data is to analyse the distribution of all 12 procurement criteria to determine the relative maturity level of each criterion within the procurement process of the organisation concerned. Table 34 shows the frequency distribution for the 12 procurement criteria for unweighted and weighted IPMM results. For the unweighted IPMM results, the first ranked criterion is represented by the procurement organisation criterion followed by the policy and procedures criterion, the procurement governance criterion and finally the risk management criterion. Notably, the highest-ranked procurement criteria include structural and procedural criteria that are included in the regulatory framework phase and procurement organisation phase. These results confirm and support the findings of the previous section. International development organisations and public sector organisations have an emphasis on the regulatory framework procurement criteria, such as the governance, policy and risk management criteria. This is considered to be due to public sector sensitivity to risk and the peculiar challenges and risks of operating and disbursing funds in developing countries (Matthew et al., 2013; Patrucco et al., 2019).



TABLE 34. FREQUENCY DISTRIBUTION IPMM RESULTS BY PROCUREMENT CRITERIA

Statistic	<i>M</i>		<i>Mdn</i>		<i>SD</i>		Rank by <i>M</i>	
	Original (Weighted)	Original (Weighted)	Original (Weighted)	Original (Weighted)	Original (Weighted)	Original (Weighted)	Original (Weighted)	Original (Weighted)
Procurement Strategy	2.776	5.941	3	6.42	1.037	2.219	5	1
Supplier Eligibility	2.577	3.066	3	3.57	0.818	0.974	9	6
Sustainable Procurement	2.625	1.732	3	1.78	0.852	0.562	8	10
Procurement Organisation	3.214	3.310	3	3.09	0.728	0.750	1	3
Customer Management	2.681	1.903	3	2.13	0.992	0.704	6	8
Procurement Systems	2.336	1.682	3	1.76	0.882	0.636	11	11
Procurement Procedures	3.251	4.096	3.33	4.2	0.735	0.926	2	2
Procurement Governance	3.026	3.238	3	3.1	0.772	0.826	3	5
Procurement Risk	2.865	1.834	3	1.92	0.649	0.416	4	9
Contract Management	2.637	3.243	3	3.69	0.914	1.124	7	4
Contract Administration	2.567	1.489	3	1.74	0.959	0.556	10	12
Training & Certification	2.462	1.945	2	1.58	1.048	0.828	12	7

The lowest-ranked criteria for the original IPMM results include the training and certifications criterion, procurement systems criterion and the contract administration criterion. The low rank of procurement systems and technology in international development reflects the lagging, but emerging, uptake of new technology in the public sector and international development as compared to quicker uptake in the private sector (Walker & Rowlinson, 2007). The higher ranked organisations, including the private sector, have better maturity results for the training and certification criteria. The results provide an interesting insight into the long-term approach to skills and resources in the procurement function, and supports current knowledge on the important link between competence, procurement maturity and performance (Aulia & Isvara, 2022).

In the weighted IPMM results, the procurement strategy criterion is first ranked, the procurement policy and procedures criterion second, and the procurement organisation criterion is ranked third. The change in mean result and rank order reflects the distinctive role and impact of the different procurement criteria, and the importance of decisive early-phase procurement strategic planning and procurement governance. All the procurement criteria have different characteristics and impacts towards the procurement process. Lindstrom (2014) argued that the high-performing procurement phases and criteria, contribute more productively towards project performance (Lindstrom, 2014). The results suggest that the combination of effective procurement strategy and governance, and the associated criteria, are key drivers for procurement maturity and which contribute towards successful project performance and DE. The findings imply that the combination of mature procurement processes contribute to improving project performance, DE and VFM. Section 5.1 concludes the analysis of the IPMM data. The next sections of this thesis are dedicated to analysis of the project evaluation results.

## **5.2 PROJECT EVALUATION RESULTS**

Section 5.1 describes the research results for procurement maturity and the following section provides analysis for the project evaluation results and the second project delivery mechanism of project management as described in the theoretical framework illustrated in Figure 6. The project evaluation results are analysed by individual organisation, the six regions of operation, and the 11 sectors, as defined by the international development organisations. The results are further broken down to understand how the project evaluation criteria measure project performance and contribute to the overall project evaluation results. Finally, the concept of cost forecasting and estimation is introduced to examine the difference between the project cost estimates and measurement of actual project costs, for the five organisations where data are available (Bayram & Al-Jibouri, 2018). The

aims and objectives of this research seek to understand the concepts of DE and VFM in international development projects. The research also examines how this is currently measured within the project evaluation criteria and provide further analysis from cost data included in project evaluation reports.

### *5.2.1 Introduction*

As detailed in Section 4.3.1, project evaluation results were obtained from ten international development organisations. The data were taken from a sample of ( $n = 1,920$  projects) from the ten organisations, across six regions and 11 different sectors, for 2015, 2016 and 2017, with data available at time of collection, for this research.

The ten international development organisations follow the OECD evaluation methodology, or variation of the methodology, as described in Section 3.10. Project evaluation results, that follow the OECD methodology in this research, are expressed as an overall project evaluation rating comprised of ratings of four evaluation criteria, namely: relevance, effectiveness, efficiency and sustainability (OECD, 2019). The international development organisations' ratings are based on different rating scales, with a description of the result either on a four, five or six-point scale. The evaluation rating scale includes a description according to the particular scale. For example, the descriptions for the six-point scale range from highly satisfactory to highly unsatisfactory, as shown in Table 35 and Figure 14. Due to the slightly different scales used in the project evaluations by the different organisations, and to enable better comparison of the results, the scales are normalised, and the revised scale and the calculation is shown in Table 23.

All the ten development organisations disclose the overall ratings and the breakdown of the component criteria evaluation results, except UNICEF and DFID. The evaluation results for the component evaluation criteria, for UNICEF and DFID, were not available

in the respective secondary data sources at time of data collection for this research. In addressing Research Questions 2 and for the analysis to address Question 3, this research provides statistical analysis of the overall project evaluation results, for the sample of ( $n = 1,920$ ) projects, by organisation, region and sector.

In addition, five of the ten international development organisations provided data on cost estimation and actual costs for each project. In-keeping with Research Questions 2, 3 and 4. This research reviews the relative performance of project management performance as shown in the theoretical framework in Section 3.13 and Figure 6. As mentioned in Section 3.4, Baccarini and Love (2014), later supported by Love et al. (2018), defined cost performance as the difference between final cost and the Cost Estimate (CE).

In support of these questions, the research also pursues analysis of VFM and explores the concept of cost performance for international development projects, by the comparison of the variation between project cost estimation and actual costs on project completion.

The project evaluation results are presented both by the percentage distribution of evaluation findings and are also presented as mean results. The mean results provide a comparative measurement of project performance for each organisation in the form of a normalised mean value of the evaluation results. As discussed in Section 5.2.1, the normalised results are used in this research for comparison of the project evaluation results between the selected international development organisations.

As discussed, the project evaluation performance results for each international development organisation in this research is comprised of the four criteria of: 1) relevance, 2) effectiveness, 3) efficiency and 4) sustainability. The international development organisations derive their overall project evaluation rating from the combined ratings of the four component criteria used for analysis in this research.

The relative percentage distribution of the ratings shows how each organisation has rated project evaluation results across the different rating descriptions from Highly Satisfactory to Highly Unsatisfactory; the results are shown in Table 35 and Figure 14. Table 35 shows the total percentage distribution for the sample of projects, published and available for the research, between 2015 and 2017 ( $n = 1,920$ ).

TABLE 35. TOTAL PERCENTAGE DISTRIBUTION OF PROJECT EVALUATION RATINGS

Ratings (%)	WBG	ADB	AFDB	EBRD	IADB	ISDB	UNDP	UNICEF	DFID	JICA
Highly Satisfactory	3	7	7	17	0	5	21	8	16	41
Satisfactory	29	77	71	20	38	61	51	59	75*	41
Partially Satisfactory	45	14	15	44	43	32	24	32	9	13
Partially Unsatisfactory	13	-	2	-	14	-	4	-	-	-
Unsatisfactory	9	2	5	19	5	2	0	1	0	5
Highly Unsatisfactory	1	-	0	-	0	-	0	-	-	-
Total projects	702	145	235	36	21	51	155	271	37	267
Percentage above Satisfactory	32	84	77	37	38	66	71	67	91	82
Rank	10	2	4	9	8	7	5	6	1	3

\* DFID follows a five-point scale and ratings A++ and A+ are treated as Highly Satisfactory as shown in Figure 14.

Table 35 and Figure 14 show the relative percentage distribution for all the organisations across the different rating scales. Table 35 shows WBG's projects from 2015 to 2017 based on the OECD (2019) criteria. The results show that 3% and 29% of its projects were ranked as highly satisfactory and satisfactory, while 9% and 1% of its projects were ranked as unsatisfactory and highly unsatisfactory, respectively. Therefore, this shows that 32% of the WBG's projects achieved satisfactory or above, which is less than the other organisations and is ranked tenth overall in the percentage distribution of project evaluation ratings in Table 35.

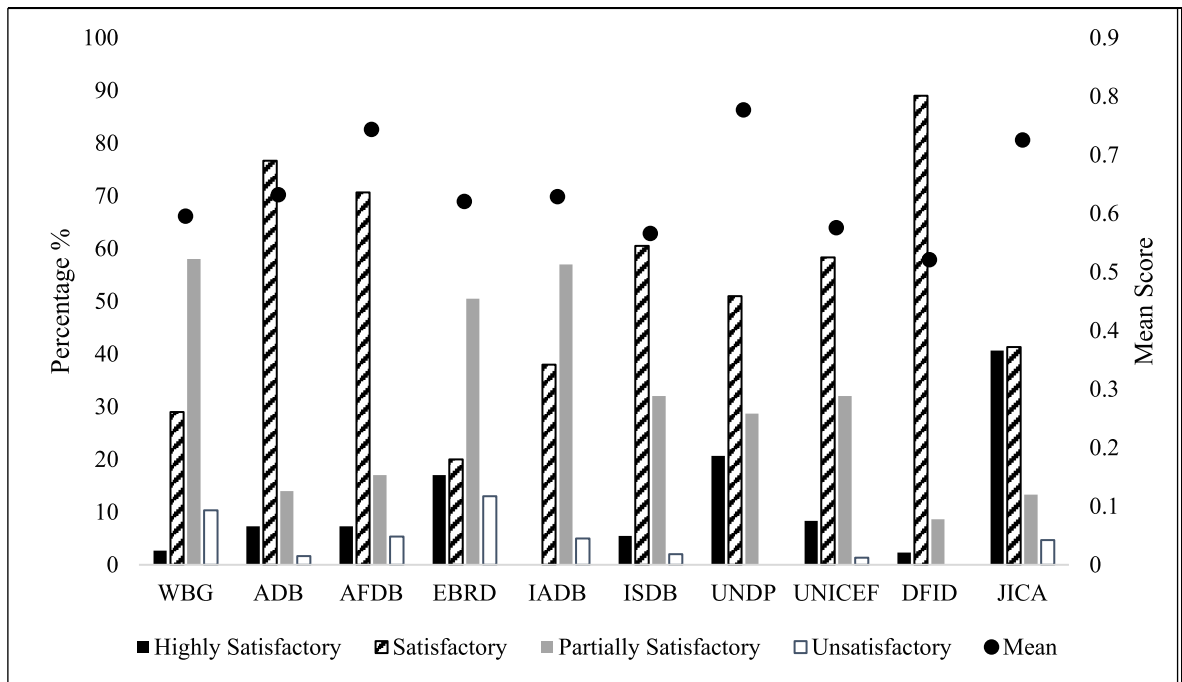


FIGURE 14. TOTAL PERCENTAGE DISTRIBUTION OF PROJECT EVALUATION RATINGS

Bulman et al (2016) suggested that project evaluation results are consistent between development organisations and therefore, the results are unexpected (Bulman, Kolkma, & Kraay, 2017). However, the results may also be an indication that there is variability in the definition of project success and the interpretation of the evaluation criteria between different organisations (Ika et al., 2010).

To illustrate this observation further, the first ranked organisation is DFID, with 16% rated highly satisfactory, 75% satisfactory and 9% partially satisfactory. The results can also be expressed as 91% satisfactory and above, and 9% below satisfactory.

Notably, despite the finding that JICA has 41% of its project evaluations rated as highly satisfactory, the largest proportion of highly satisfactory rating for all the development organisations, it is ranked third overall when the percentages are combined.

In terms of percentage ratings in Table 35, DFID, ADB, JICA and AFDB emerge as the four highest-ranked organisations and have evaluation ratings at 91%, 84%, 82% and 77% above satisfactory, respectively. At the other end of the spectrum, WBG is ranked

tenth with ratings of 32% above satisfactory with 68% below satisfactory, EBRD is ninth with 37% above satisfactory and 63% below satisfactory, and IADB ranked eighth with 38% above satisfactory and 62% below satisfactory, respectively.

There are substantial variations in the percentage distribution patterns between different organisations and variation may be explained by widely different approaches to the definition and interpretation of each development organisation's evaluation results. It is possible that this is a methodological issue, as differences can be the result of the different rating scales and due to the different standards used in the application of the ratings in the project evaluation process. The variations highlight an important issue with the analysis of the different rating scale of the evaluation data from the ten organisations. A major problem with comparing disparate data sets was overcome using normalisation equations to reduce the potential distortion effect of the different rating scales, as shown in Table 23. Comparison of statistical results of two different normalisation equations is discussed in the next section and shown in Appendix 3.

#### *5.2.2 Overall project evaluation results by the total mean score*

For further analysis and comparison of the international development organisations for all three years, 2015 to 2017, the total overall mean results are calculated using the total normalised scores and divided by the total number of projects evaluated.

The total mean result provides a representation of the sample of 1,920 projects as a single value. The total overall mean scores are tabulated in Table 36 and illustrated in Figure 15 with the statistical distribution analysis shown in Table 37.

TABLE 36. PROJECT EVALUATION RATINGS TOTAL MEAN RESULTS

Year	WBG	ADB	AFDB	EBRD	IADB	ISDB	UNDP	UNICEF	DFID	JICA
2015	.589	.596	.747	NA	.628	.604	.742	.517	.522	.755
2016	.576	.673	.748	.694	NA	.527	.796	.590	.571	.704
2017	.621	.626	.735	.547	NA	NA	.792	.621	.479	.716
Overall <i>M</i> Result	.595	.635	.740	.617	.628	.569	.777	.576	.527	.725
Total	702	145	235	36	21	51	155	271	37	267
Rank by <i>M</i>	7	4	2	6	5	9	1	8	10	3

The total overall mean score results are arranged to give a rank order for all project evaluation ratings and a total overall score for all years combined, as shown in Table 36.

The total mean result of project evaluations is a broad measure of project performance for each organisation for all three years, regions and sectors combined. For example, the WBG projects achieved a mean result for project evaluation ratings of .595 for the 702 total evaluated projects between 2015 and 2017, and results indicate that WBG is ranked seventh across all organisations, as shown in Table 36. The overall mean results are useful for the comparison of project performance for international development organisations that follow the OECD project evaluation methodology. As discussed in Section 5.2.2, there are substantial differences in project evaluation rating scales.

The normalised mean results provide a basis for the comparison of the project evaluation results of the international development organisations, and the calculation used to develop the normalised results is shown in Table 23.

The mean results as shown in Table 36, indicate that UNDP has the highest performance overall compared to the other international development organisations, for self-evaluated results over the three-year study period. AFDB has the second-highest result, followed by JICA. The three organisations of UNDP, AFDB, and JICA represent each of the three types of international development organisations reviewed in this study. At the other end



of the scale, the lowest results are closely grouped, and include UNICEF, ranked seventh, ISDB, ranked eighth and finally, DFID which is ranked tenth for the mean project evaluation results.

TABLE 37. FREQUENCY DISTRIBUTION DATA FOR PROJECT EVALUATION RESULTS

<i>Statistic</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	<i>Skewness</i>	<i>Kurtosis</i>	<i>p-value</i>	<i>Rank</i>
WBG	.595	.600	.201	-0.697	0.300	.000	7
ADB	.635	.667	.168	-0.837	2.788	.000	4
AFDB	.740	.800	.173	-1.880	3.570	.000	2
EBRD	.617	.600	.272	0.519	-0.180	.000	6
IADB	.628	.600	.171	0.827	0.375	.002	5
ISDB	.569	.667	.203	-1.303	0.252	.000	9
UNDP	.777	.800	.155	-0.310	-0.246	.000	1
UNICEF	.576	.667	.208	-0.010	-0.082	.000	8
DFID	.527	.500	.142	1.000	3.250	.000	10
JICA	.725	.667	.278	-0.851	0.202	.000	3

As discussed in Section 5.2.1, each organisation has different approaches to project evaluation rating methodology, and many of the project evaluations are performed by the project managers or Task Team Leaders (TTL) responsible for the project. The self-evaluation potentially introduces a positive, and sometimes negative, bias in the evaluation process (Denizer et al., 2013). In addition, as Ika et al. (2010) noted, the evaluation procedures, data collection methods, and interpretation of the ratings differ markedly between organisations. There are notable differences between the percentage distribution shown in Table 35, and the mean project evaluation results are shown in Table 36 and Table 37. The percentages of ratings that fall above satisfactory serves as a useful perspective for the project evaluation results and, also suggests that there are differences in the way that each organisation rates the project evaluation results. As discussed in Section 4.2.1, the organisations use three different ratings scales to evaluate the OECD project evaluation methodology. The project evaluation results reveal some of

the distortions that these different scales can introduce, in particular, the potential differences that an odd number rating scale can present when compared to even number rating scales (Cummins & Gullone, 2000). This is evident in these results for DFID, which is the only organisation to use a five-point rating scale.

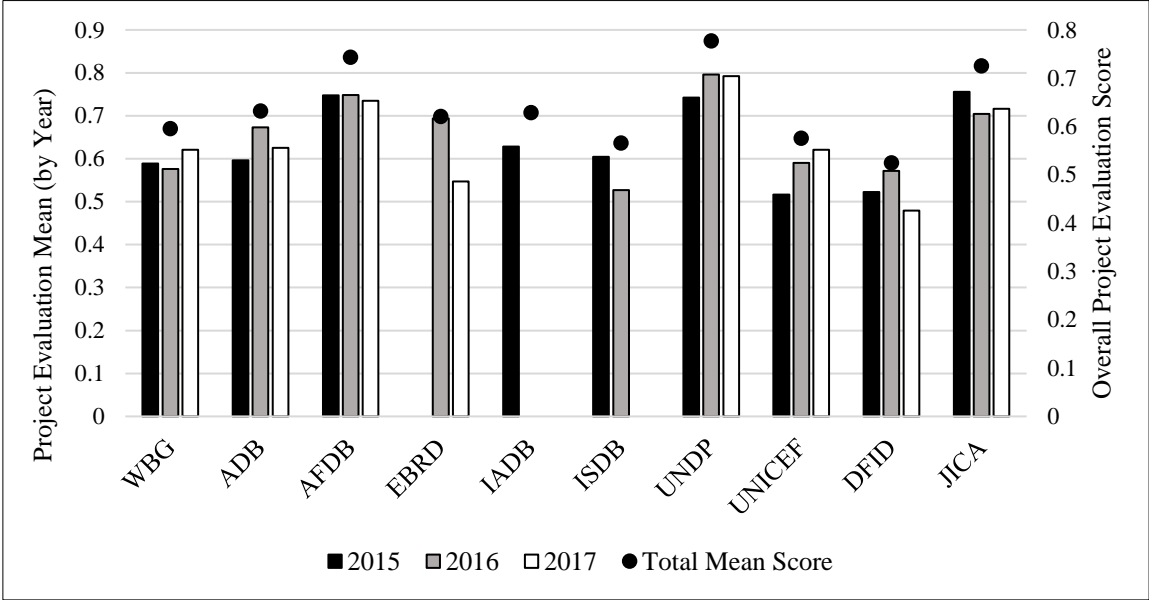


FIGURE 15. PROJECT EVALUATION RATINGS TOTAL MEAN RESULTS

As discussed in Section 5.2.2, the difference for DFID from highest rank for percentage distribution, shown in Table 35, to lowest rank for total mean score shown in Table 36 and Table 37, is a result of the different description and treatment of the five-point scale. DFID rates 91% of projects as satisfactory, or meets expectations, which is the mid-point of the five-point scale, as this is shown in Figure 14. The normalised rating scale has the effect of reducing the mean result and ranking placement for DFID from first to tenth rank. The change in rank for DFID is a direct result of the normalisation and treatment of the five-point scale and the difficulty of comparing different scales.

As discussed in Section 4.5, an alternative treatment of the normalised DFID five-point rating scale was calculated as a sensitivity analysis to understand possible different interpretations of the five-point scale. For example, if A++ and A+ are combined and

treated as highly satisfactory, and A treated as satisfactory, the mean result is recalculated to give a mean result of .693. Alternatively, if A++ is treated as highly satisfactory, and A+ and A combined and treated as satisfactory, the mean result would then be recalculated to give a mean result of .648 on the normalised scale.

In both scenarios of the sensitivity analysis, after recalculation of the mean results, DFID would be ranked fourth in Table 37 for comparative project evaluation results.

It is noted that the alternative treatment of the ratings gives different results and rank order for DFID. However, despite the effect of different rating scales on the results, for purposes of consistency, the calculation for normalisation of the rating scales, shown in Table 23, is maintained and utilised for further comparison in this research.

This is unfortunate and should be taken into consideration for the DFID results, and further demonstrates the challenges and difficulty in the comparative analysis where different scales are used in project evaluation rating methods. Due to the extreme difference in the results for DFID after normalisation of the five-point scale it is important to apply caution before drawing conclusions. Nevertheless, it is useful to be able to measure and benchmark project evaluation results with other organisations that use the OECD project evaluation methodology and the analysis in this research includes the mean results for DFID. This finding potentially makes a strong case for the adoption of uniform rating scales and standards for project evaluations in international development where project performance comparison and benchmarking is required.

Tests of normality for both Kolmogorov-Smirnov and Shapiro-Wilks are significant at the  $p < 0.05$  level for the sample ( $n = 1,920$ ) organisation's project evaluation results, as tabulated and shown in Table 37, and therefore, shows that the distribution for project evaluation results by the selected international development organisations does not

follow a normal distribution. The results indicate that each organisation has different strengths and weaknesses by the year, region and sector of project activity. Further analysis of project performance is performed in the following sections of the research.

A one-way ANOVA test was conducted to compare the project evaluation results between international development organisations. The results show that there is a significant difference between organisations at the  $p < 0.05$  level and the distribution does not follow a normal distribution  $F(9, 1910) = 27.19, p = .000$ . The ANOVA results confirm that international development organisations have variable project evaluation performance across the sample of project evaluation results ( $n = 1,920$ ).

The Levene's test was similarly conducted to test the variability of the means and median of the project evaluation results and indicates a significant result,  $F(9, 1910) = 15.13, p = .000$ , based on the mean and,  $F(9, 1910) = 13.11, p = .000$ , based on the median scores, which confirms the results of the other statistical tests.

These results show that there is a statistically significant difference between the means and the medians for project evaluation results for overall project evaluation results. The results confirm the one-way ANOVA results indicating that international development organisations have a range of significantly different project performance results across the 1,920 projects reviewed. This is an important finding of the research because it provides evidence that the different project evaluation methods of development organisations make it difficult for performance to be evenly judged. The findings also suggest that inconsistency in measurement makes it challenging to gauge lessons learned and identify best practices from particularly successful organisations.

The Bonferroni multiple comparison tests show that there appears to be specific patterns and two groupings to emerge from the data based on the statistical distribution results.

WBG, ADB, EBRD, IAD, ISDB, UNICEF and DFID form one group and UNDP, AFDB and JICA form the other. The  $p$ -values for comparison within this first group are not significant, at the  $p > 0.05$  level, indicating that the null hypothesis is accepted and there are similarities of statistical results within the first group of seven of organisations. UNDP, AFDB and JICA that form the second group, do not show a statistically significant result, in relation to the first group of seven organisations. The second group of three organisations, including UNDP, AFDB and JICA have the highest overall mean results and corresponding high-rank order of first, second and third-ranked. A possible explanation for the difference between the three high-performing organisations is that they have high-performing projects across the years, regions and sectors and may include best practices. Further analysis will examine the results in the following sections.

The Kruskal-Wallis test shows a significant variation in project evaluation results  $H(9) = 295.41, p = .000$ , and the null hypothesis is not accepted. This confirms the other statistical test results that there are significant statistical differences in the project evaluation results of the international development organisations.

### *5.2.3 Analysis of project evaluation by type of organisation*

To understand if there is a difference between the types of organisations, such as MDB, UN Agency and bilateral development organisation statistical analysis was conducted by organisation type. The frequency distribution for the project evaluation results by type of organisation shown in Table 38, indicates that the UN Agencies have the highest overall mean (indicating that these organisation rate themselves highest when it comes to project evaluation), followed by the MDBs and then the bilateral development agencies. As shown in Table 38, the MDBs have a combination of the highest-ranking mean scores, for example, ADB and AFDB are second and fourth, and WBG and ISDB have the lowest-ranked mean scores and are ranked at seventh and tenth. The mean ranking of the

UN Agencies, UNDP and UNICEF, are ranked first and eighth, and the bilateral development organisations, JICA and DFID, are ranked at third and tenth. The mean results are therefore comprised of a combination of different results by organisation type.

TABLE 38. STATISTICAL DISTRIBUTION PROJECT EVALUATION BY ORGANISATION TYPE

Project Evaluation	Organisations	Projects	<i>M</i>	<i>Mdn</i>	<i>SD</i>	Rank
Overall						
MDB	6	1,190	.628	.600	.202	<b>2</b>
UN Agency	2	426	.677	.670	.213	<b>1</b>
Bilateral	2	304	.623	.670	.273	<b>3</b>

Tests of normality for both Kolmogorov-Smirnov and Shapiro-Wilks for the types of organisations are statistically significant, at the  $p < 0.05$  level, for the project evaluation results shown in Table 38. The results indicate that the distribution for project evaluation results, by types of the international development organisation, does not follow a normal distribution. The tests suggest that the three types of international development organisations have variable project performance by organisation type across the total of the 1,920 project evaluation results.

A one-way ANOVA test was conducted to compare the three types of international development organisations' project evaluation results. The results show a significant difference between the types of organisations at the  $p < 0.05$  level and the distribution does not follow a normal distribution,  $F(2, 1917) = 13.77, p = .000$ . The ANOVA test indicates that the types of international development organisations have variable project evaluation performance across the sample of project evaluation results ( $n = 1,920$ ).

The Levene's test was conducted to test the variability of the means and median of the project evaluation results for the types of the international development organisation and indicates significant result at the  $p < 0.05$  level with,  $F(2, 1917) = 23.74, p = .000$ , based on the mean and,  $F(2, 1917) = 16.98, p = .000$ , based on the median. The results support

the one-way ANOVA results and indicate the international development organisations have a range of significantly different project performance results by organisation type across the projects reviewed.

The Bonferroni multiple comparison tests for different types of organisations results, shown in Table 39, indicates that the MDBs have a statistically significant result at the  $p < 0.05$  level in comparison with the bilateral agencies and a non-statistically significant result at the  $p > 0.05$  level in comparison with the UN agencies.

TABLE 39. BONFERRONI STATISTICAL TEST PROJECT EVALUATION BY ORGANISATION

<i>Bonferroni statistic test</i>			
Organisation	MDB	UN Agency	Bilateral
MDB		.251	.000
UN Agency	.251		.005
Bilateral Development	.000	.005	

Results in Table 39 suggest that the MDBs have a slightly greater alignment with UN agencies than bilateral development groups. The Bonferroni multiple comparison test for UN Agencies and bilateral development organisations show a statistically significant result at the  $p < 0.05$  level. The project evaluation ratings, across types of organisations, have a large degree of variation between each type, although the results suggest that MDBs and UN Agencies have greater alignment with each other for project performance results. The three types of organisations strive to align their evaluation processes through continuous collaboration and engagement. The collaboration and results may be explained by the similar size and approach of the MDB and UN Agencies to project evaluations (Morra-Imas et al., 2009; Savedoff, Levine, & Birdsall, 2005).

The Kruskal-Wallis test shows that there is a significant variation in project evaluation results  $H(2) = 32.41, p = .000$ , and the null hypothesis is not accepted and supports the

statistical observations that there are significant differences in the project evaluation results of the types of international development organisations.

The statistical analysis indicates an important finding that there are significant differences in the project evaluation results for the international development organisations. There are several possible explanations for the differences in project evaluation results. As discussed in Section 3.10, project evaluation reports are the principal measure for project success and DE at the project level for international development projects. Nevertheless, potential explanations include the observation that international development organisations use different rating scales, that there is variability and inconsistency in the description, input and interpretation of the ratings, and that there is a tendency for a positive and negative bias in self-evaluated project evaluations and reporting (Clements et al., 2008; Denizer et al., 2013).

#### *5.2.4 Analysis of project evaluation results by region*

The previous section considered project evaluation results based on organisation type (MDB, UN Agency or Bilateral Development) and it was found that the UN Agencies have higher performance ratings as group, followed by the MDBs and bilateral development organisations. The results of the individual international development organisations is very different and therefore the results provide an interesting general summary of the project evaluation process followed by the organisation types.

This section considers the data on project evaluation from different perspectives by conducting analysis based on region. The analysis provides additional perspective in response to Question 2 on the project evaluation methodology of project performance in different geographical regions of operation. For analysis of the project evaluation results by region, the normalised mean results for each region are compared and illustrated in Table 40 and Figure 16.



TABLE 40. TOTAL MEAN RESULTS FOR PROJECT EVALUATION RATINGS BY REGION

Region	WBG	ADB	AFDB	EBRD	IADB	ISDB	UNDP	UNICEF	DFID	JICA
Africa (AFR)	.551		.754			.571	<b>.763*</b>	.512	.509	.650
East Asia & Pacific (EAP)	.670	.660				.566	.764	.645	.625	<b>.796</b>
Europe & Central Asia (ECA)	.644	.650		.620		.583	<b>.808</b>	.675	.604	.481
Latin America (LCR)	.560				.628		<b>.759</b>	.552	.500	.587
Middle East (MNA)	.542					.472	.780	.598	.604	<b>.833</b>
South Asia (SAR)	.629	.585				.599		.642	.500	<b>.797</b>
Global	.600		<b>.655</b>					.516		
TOTAL	.601	.633	.704	.620	.628	.558	<b>.775</b>	.596	.560	.679
Rank by <i>M</i>	7	4	2	6	5	10	1	8	9	3

\*The values in bold, Table 40, indicate the highest mean project evaluation results by region

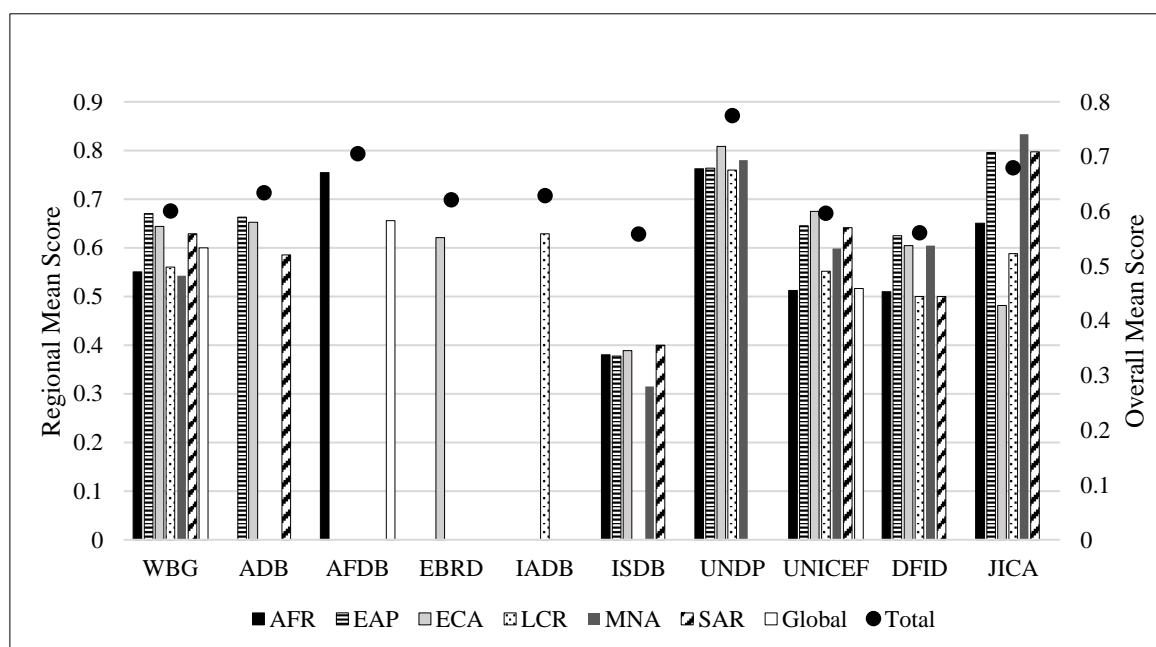


FIGURE 16. TOTAL MEAN PROJECT EVALUATION RESULTS BY REGION

The rank order by region is very similar to the overall mean results table as indicated in Table 36, Table 37 and Figure 15, with one exception where DFID and ISDB change rank order. DFID changes from tenth to ninth rank and ISDB changes from ninth to tenth rank and is this is most likely as ISDB does not operate in LCR and data not available. UNDP has the highest mean project evaluation result overall by region and has the highest mean for the Africa (AFR), Europe and Central Asia (ECA) and Latin America

(LCR) regions, as presented in Table 40. The results show a mean of all project evaluations collated for the organisations, in this research, and provide an indication of overall project performance over the period and regions. For specific regional results, JICA has the highest mean result and rank for the East Asia and Pacific region (EAP), the Middle East (MNA) region and the South Asia (SAR) region.

The regional development banks such as ADB, AFDB, EBRD, IADB and ISDB have a smaller range of countries due to their regional focus of activity and relatively smaller size. WBG, UN Agencies and the bilateral organisations, have a more global presence, organisational size, scale and assigned resources to cover a global presence. WBG, AFDB and UNICEF include a category of projects as Global, and these projects are often centrally managed, and the category covers a global range or more than one region.

The UN agencies and bilateral development organisations have similar global coverage and presence as WBG, although projects are generally smaller in size and scale.

There is some variation of project evaluation ratings between different regions for each organisation, suggesting that there are either regional opportunities and challenges or operational reasons for different project performance results. For example, ADB has wide geographical distribution, including East Asia and the Pacific (EAP), and manages projects within the member countries in ECA and South Asia (SAR). The AFDB results include project evaluation reports for two regions, including Africa (AFR) and Global projects. IADB includes projects evaluation results in the Latin America (LCR) region only, and EBRD includes projects for Europe and Central Asia (ECA). The variation of regional results will be discussed further in the research.

### 5.2.5 Statistical distribution of project evaluation results by region

The statistical frequency distribution for each region provides an analysis of the total project evaluation results, as shown in Table 41, and shows the consolidated project evaluation data and Corruption Index (CI) indicators for all ten organisations by region.

TABLE 41. PROJECT EVALUATION RESULTS AND CORRUPTION INDEX BY REGION

Project Evaluation (PE) Corruption Index (CI)	PE <i>M</i>	PE <i>SD</i>	CI <i>M</i>	CI <i>SD</i>	PE Rank	CI Rank
East Asia & Pacific (EAP)	.706	.209	35.55	8.52	<b>1</b>	<b>2</b>
Europe & Central Asia (ECA)	.658	.214	38.51	7.64	<b>2</b>	<b>1</b>
South Asia (SAR)	.649	.221	33.53	6.86	<b>3</b>	<b>5</b>
Africa (AFR)	.633	.214	33.28	10.50	<b>4</b>	<b>6</b>
Latin America (LCR)	.594	.223	34.81	10.38	<b>5</b>	<b>4</b>
Middle East & N. Africa (MNA)	.577	.223	33.09	12.99	<b>6</b>	<b>7</b>
GLOBAL	.570	.214	34.96	9.77	<b>7</b>	<b>3</b>

The region with the highest mean rating results and rank is East Asia and Pacific (EAP) followed by Europe and Central Asia (ECA), South Asia (SAR) and then Africa (AFR).

The analysis of project evaluation results suggests that these four regions show better project performance from the results of the international development organisations.

The lowest-ranked region, or region with the lowest self-reported project evaluations, is the Global category contributed by WBG, AFDB and UNICEF, in Table 40. The low results for the Global category may be due to the relative difficulty, or complexity in managing global-type projects, particularly when the projects are centrally managed across a large geographical region (Anantatmula & Thomas, 2010).

Tests of normality for both Kolmogorov-Smirnov significance results are significant, at the  $p < 0.05$  level, for the project evaluation results by region, shown in Table 41, and the results show that the distribution for project evaluation results by region does not follow a normal distribution. The Kolmogorov- Smirnov results for global project evaluations are non-significant, at the  $p > 0.05$  level, and follow a normal distribution. The significant statistical results suggest that there is variable comparable project

performance by comparison of the regions, and the non-significant result for the Global category, suggests that there is greater alignment in project performance for global projects. As was the case with mean rankings, there are several possible explanations for the low mean result and rank order for the Global category, which may include the challenges in management of global projects administered from centrally located offices.

A one-way ANOVA test was conducted to compare the international development organisations' project evaluation results by region. The results show a significant difference between regions, at the  $p < 0.05$  level, and the distribution does not follow a normal distribution,  $F(6, 1913) = 9.97, p = .000$ . The ANOVA test indicates that the international development organisations have variable project evaluation performance across the ( $n = 1,920$ ) project evaluation results by regions.

The Levene's test was conducted to test the variability of the means and median of the project evaluation results for the international development organisation by region and indicates a non-significant result at the  $p > 0.05$  level with,  $F(6, 1913) = 0.379, p = .892$ , based on the mean. The Levene's test indicates that there is statistical alignment across regions for the project evaluation results by means and median.

As discussed in Section 3.7, Transparency International (TI) produces an annual Corruption Index (CI) which scores each country on a scale from 0 – 100 with a high score indicating high transparency and low levels of corruption (Transparency International, 2021). To understand the relationship between the corruption index and the project evaluation ratings, a Pearson's correlation test was conducted between the country corruption index and the overall project evaluation results. The test was performed to understand the correlation between the overall project performance and the TI corruption index and whether corruption may be a factor in project performance.

There were 1,786 projects in the research sample against which the country corruption index could be identified from the Transparency International Corruption Index data.

Field (2017) and Kent State University (2013) specify that, for the Pearson correlation coefficient, the values are interpreted as follows:

- 0 - No correlation
- - 0.2 to 0 / 0 to 0.2 – very weak negative/positive correlation
- - 0.4 to - 0.2/0.2 to 0.4 – weak negative/positive correlation
- - 0.6 to - 0.4/0.4 to 0.6 – moderate negative/positive correlation
- - 0.8 to - 0.6/0.6 to 0.8 – strong negative/positive correlation
- -1 to - 0.8/0.8 to 1 – very strong negative/positive correlation
- -1/1 – perfectly negative/positive correlation

(Field, 2017; Kent State University, 2013)

The Pearson’s correlation result shows a statistically significant very weak correlation of  $r(1786) = .115, p = .000$ . This indicates that there is a very weak relationship between the Corruption Index (CI) and the 1,786 project evaluation results which are mapped to CI. To provide more resolution of project evaluation and corruption in different regions, the Pearson’s correlation test was conducted between project evaluation results and the CI for each region. The correlation results for the six regions are shown in Table 42 below.

TABLE 42. CORRELATION OF PROJECT EVALUATION AND CORRUPTION BY REGION

Region	Pearson’s correlation test results
East Asia and Pacific (EAP)	$r(389) = .009, p = .855$
Europe and Central Asia (ECA)	$r(215) = .117, p = .087$
South Asia (SAR)	$r(215) = .117, p = .087$
Africa (AFR)	$r(642) = .128, p = .001$
Latin America (LCR)	$r(231) = .064, p = .334$
Middle East and North Africa (MNA)	$r(207) = .128, p = .085$

The results show that all regions have a very weak positive correlation between the project evaluation results and TI corruption indices. However, there are differences between the regions. The strongest correlations are for MNA and AFR and this implies that for these two regions, there is a greater relationship between project performance and the level of corruption as measured by the TI index. MNA and AFR regions have the lowest mean result for the TI index at 33.09 and 33.28 respectively. EAP and ECA, on the other hand have the highest means by region for the TI index of 38.5 and 35.54 respectively. The mean results provide only an approximate indication of the overall levels of corruption by region. However, despite the very weak correlation results, this implies there is a relationship between corruption and project performance by country. The slightly stronger correlation results for MNA and AFR, which have both low project evaluation results and low CI (more corruption), implies that projects are more impacted, or influenced, by corruption for these two regions than for the other four regions.

The descriptive statistics shown in Table 41 indicates the relative performance of the project evaluation and corruption index results and provides an indication of the relative project performance in comparison with corruption index by country and region. The difference in project performance by country and region may be due to different internal or external regional management, political, economic or corruption factors. The weak correlation between the regional project evaluation results and the TI corruption index is a new finding and further research is recommended to further explore and understand the implications. The project evaluation and CI results are further discussed in Chapter 6.

#### *5.2.6 Analysis of project evaluation results by sector*

International development organisations classify and orientate projects according to 11 sectors shown in Table 43 and Table 44 To understand how the international development organisations perform by sector, the project evaluation results are therefore

analysed by the 11 sectors in this research and Table 43 shows the total mean project evaluation results and rank order for each organisation. The results by sector follow the same rank order as the overall mean results, shown in Table 36 and Table 37.

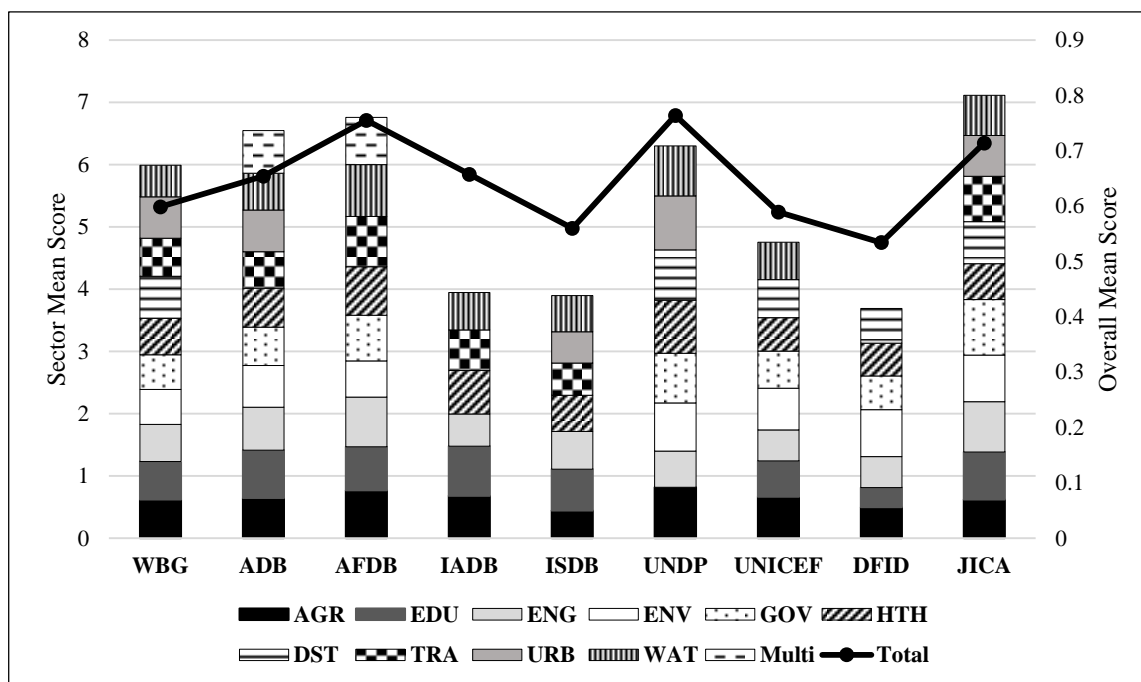
TABLE 43. TOTAL MEAN RESULTS FOR PROJECT EVALUATION RATINGS BY SECTOR

Sector	WBG	ADB	AFDB	EBRD	IADB	ISDB	UNDP	UNICEF	DFID	JICA
Agriculture	.623	.648	.768		.680	.444	<b>.816**</b>	.666	.500	.624
Education	.611	.768	.700		<b>.800</b>	.666		.574	.312	.761
Energy	.593	.688	.523		.514	.604	.583	.499	.500	<b>.806</b>
Environment	.563	.666	.618				<b>.773</b>	.666	.750	.752
Government	.557	.615	.730				.798	.596	.542	<b>.888</b>
Health	.585	.629	.777		.700	.583	<b>.850</b>	.540	.528	.574
Disaster Preparedness	.673						<b>.811</b>	.607	.558	.677
Transportation	.615	.586	<b>.810</b>		.650	.518				.727
Urban Infrastructure	.664	.666				.499	<b>.866</b>			.655
Water	.504	.591	<b>.832</b>		.600	.583	.800	.603		.644
Multisector*	.682	<b>.758</b>								
<b>TOTAL</b>	<b>.598</b>	<b>.654</b>	<b>.737</b>		<b>.657</b>	<b>.559</b>	<b>.764</b>	<b>.589</b>	<b>.534</b>	<b>.714</b>
Rank by Mean	6	5	2		4	8	1	7	9	3

\* Multi-sector refers projects that include more than one of the ten sectors

\*\* The values in bold, Table 43, indicate the highest mean project evaluation results by sector

As with the regional results, UNDP has the highest overall mean project evaluation results by sector. UNDP has the highest sector results for the agriculture, environment, government, health, disaster preparedness and urban infrastructure sectors, as shown in Figure 17. Similarly, AFDB is ranked second overall and has the highest results for transportation, water management and multi-sector projects. Secondary data by sector were not available for EBRD, and therefore, EBRD results are not shown in Table 43.



\* EBRD did not provide project evaluation results by sector

FIGURE 17. PROJECT EVALUATION MEAN RESULTS BY SECTOR

None of the international development organisations have data for all 11 of the sectors.

WBG and JICA have data for ten of the 11 sectors, with the only exclusion of the Multi-Sector Category. ADB also includes data for ten of the 11 sectors with the only exception of the Disaster Preparedness sector.

Figure 17 shows the project evaluation results for the individual sectors for the international development organisations, and the line represents the overall mean results by sector. The shaded cells in the table indicate that UNDP has the highest results for the eight project evaluation sectors represented by the secondary data as disclosed. The rank order remains the same as the overall project evaluation results in Table 37, and the project evaluation results by region in Table 40, taking into consideration that EBRD is not included as there is no data were available for EBRD by sector for this research.

### 5.2.7 Statistical distribution of project evaluation results by sector

The statistical distribution of the project evaluation results for all the international development organisations by sector is shown in the following Table 44.



TABLE 44. STATISTICAL DISTRIBUTION OF PROJECT EVALUATION RESULTS BY SECTOR

Project Evaluation Sector	<i>M</i>	<i>N</i>	<i>Mdn</i>	<i>SD</i>	Rank
Multisector (Multi)	.729	45	.800	.149	<b>1</b>
Environment (ENV)	.690	146	.667	.229	<b>2</b>
Agriculture (AGR)	.666	204	.667	.199	<b>3</b>
Energy (ENG)	.660	146	.667	.222	<b>4</b>
Transportation (TRA)	.657	197	.667	.207	<b>5</b>
Education (EDU)	.657	178	.667	.224	<b>6</b>
Disaster Preparedness (DST)	.643	121	.667	.186	<b>7</b>
Urban Infrastructure (URB)	.627	84	.600	.216	<b>8</b>
Water & Sanitation (WAT)	.624	194	.667	.239	<b>9</b>
Health (HTH)	.617	331	.667	.224	<b>10</b>
Governance (GOV)	.586	160	.667	.199	<b>11</b>

The best performing sector for overall statistical results is the Multisector category, followed by the environment and agriculture sectors and the lowest performing and ranked sectors, are represented by the government and health sectors.

Tests of normality for both Kolmogorov-Smirnov and Shapiro-Wilks for project evaluation results by sector are statistically significant, at the  $p < 0.05$  level and indicate that the distribution for project evaluation results by sector does not follow a normal distribution. These results provide evidence to suggest that international development organisations have variable project performance by sector and that some sectors have more successful projects and others do not perform as well or are less successful.

A one-way ANOVA test was conducted to compare the project evaluation results by sector and the results,  $F(6, 1913) = 9.97, p = .000$ , show that there is a significant difference between project evaluation performance by sector, at the  $p < 0.05$  level.

The Levene's test was conducted to test the variability of the mean and median results for project evaluation by sector and gives a statistically significant result, at the  $p < 0.05$  level, by mean and a statistically significant result, at the  $p < 0.05$  level, by the median.

These results indicate a statistically significant variation between the means and the medians for project evaluation results by sector.

These results suggest that the high-performing top three sectors have different project performance characteristics with the lowest performing, 11th ranked governance sector.

The differences in project performance by sector may be due to many factors such as the nature, technology, conditions and environment of the project concerned. For example, the water and sanitation, transportation, energy and urban infrastructure sectors often include complex construction design, technical scope and performance requirement (Karaca et al., 2020; Love et al., 2018; Tayefeh et al., 2020). The public health sector also includes many types of projects from supply of equipment and medicines to training and infrastructure. Santos et al. (2020) argued that measurement of success factors in the health sector has lagged behind the some of the other sectors (Santos et al., 2020). Kim et al. (2018) recorded several causes of delays and cost increases in hospital project in Vietnam and while these have similarities to causes in other industries, there was little consensus amongst the respondents to the survey (Kim et al. 2018; Vu et al., 2020). The education industry similarly includes many project types including construction of schools and infrastructure, procurement of books and materials, training and governance (Famiyeh et al., 2017). Asiedu et al. (2017) observed that project success was measured by cost, schedule and quality in education construction projects in Ghana and that 28% of projects completed either met, or were below the cost estimates (Asiedu et al., 2017). The results suggest that the high-performing sectors share sector-specific attributes and project performance characteristics, and these can provide examples of best practices for improved project performance in international development organisations. As such, multi-sector sector, environment, and agriculture sectors show better project evaluation results, in this research, and may offer valuable lessons for the improvement of project outcomes in sectors like health and governance that showed lower results.

### 5.2.8 Project evaluation ratings by region and sector

The project evaluation results for any given region may experience different results for each of the sectors and this is further explored in this section. The project evaluation results illustrated in Table 45 and Figure 18, shows the relative percentage of the highly successful and successful project ratings cross tabulated for each sector by region. This is helpful as it shows, for example, that the Africa region has 73.80% of projects successful and above for water and sanitation projects whereas it has 41.38% successful and above for the health sector. The best performing sector for all regions is highlighted in bold.

TABLE 45. PROJECT EVALUATION RATINGS BY REGION AND SECTOR

Sector (%)	AGR	EDU	ENG	ENV	HTH	GOV	DST	TRA	URB	WAT	Multi
AFR	60.22	48.61	54.90	53.33	41.38	48.00	50.00	58.44	46.66	<b>73.80*</b>	85.29
EAP	68.96	80.95	65.71	<b>73.44</b>	<b>77.77</b>	77.27	<b>75.00</b>	66.66	56.52	64.28	88.88
ECA	42.11	72.22	<b>75.00</b>	61.54	61.66	53.85	60.00	<b>75.00</b>	40.00	33.33	<b>100.00</b>
LCR	38.71	68.75	21.43	37.77	40.00	42.85	63.16	43.75	18.75	31.82	
MNA	0.00	37.50	36.36	71.43	38.88	37.50	38.88	50.00	40.00	33.33	
SAR	<b>69.23</b>	<b>85.00</b>	68.42	66.66	52.77	55.55	68.75	59.37	<b>70.00</b>	65.00	
Global		50.00			33.33	<b>100.00</b>					
Total (n)	204	178	146	224	331	160	121	197	84	194	45

\*The values in bold, Table 45, indicate the highest percentage project evaluation results by sector (%)

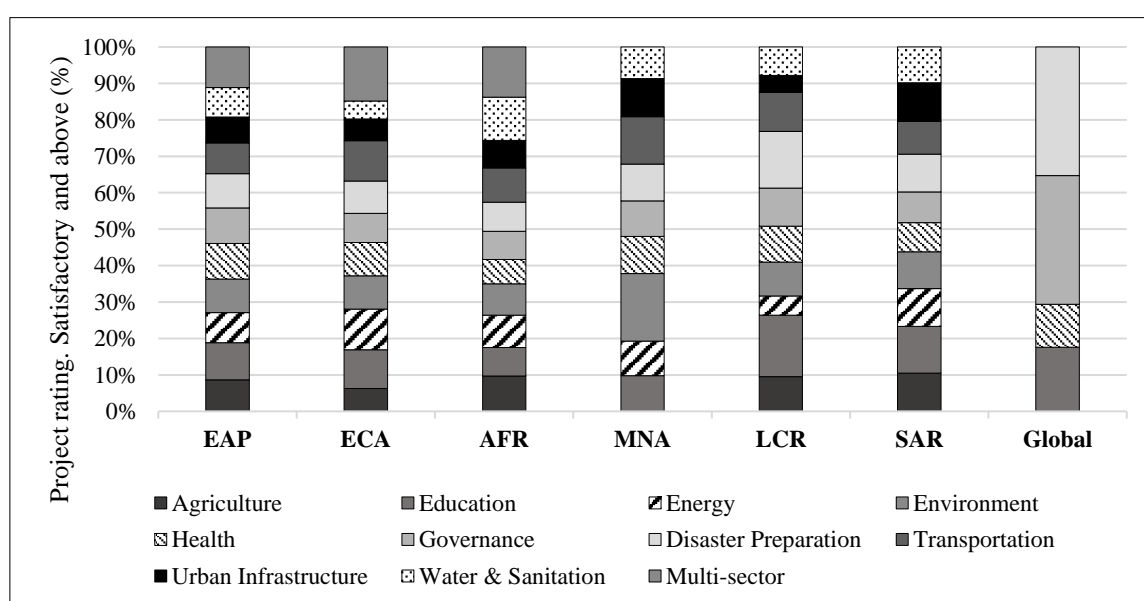


FIGURE 18. TOTAL PROJECT EVALUATION RATINGS BY REGION AND SECTOR

There are a number of findings of interest. For example, EAP has the highest percentage of successful projects for three sectors of environment, health and disaster preparedness. ECA has the highest percentage of successful projects for the energy, transportation and multisector sectors. In addition, SAR has the highest percentage of successful projects for the agriculture, education and urban infrastructure sectors. Table 41 shows that EAP, ECA and SAR are the highest performing three regions based on the project evaluation results in this research. AFR has the highest percentage of successful projects for the water and sanitation sector and the lowest for the health sector. The multisector category performs comparatively well across all regions, and this may be due to centralised project management and increased quality of oversight applied to centralised projects (Desouza & Evaristo, 2006). LCR has a lower percentage of successful projects in five sectors including the energy, environment, transportation, urban infrastructure and water and sanitation sectors. Asian Development Bank (ADB), in the 2019 annual evaluation report, classified the energy, transportation, water sanitation sectors as infrastructure-type sectors as they include substantial construction and infrastructure requirements.

The World Bank Group (WBG), in the 2020 annual evaluation report, noted that projects face particular difficulties in Fragile, Conflict and Violence (FCV) affected countries such as MNA and AFR. The Japanese International Co-operation Agency (JICA), in the annual evaluation report 2017, observed that a majority of successful project were located in the EAP and SAR regions. These projects included infrastructure-type sectors such as transportation, water sanitation, urban infrastructure and environment sectors.

The JICA results and the JICA annual report in 2020 suggest that JICA has project management strengths and demonstrates best practices in the infrastructure-type sectors.

As discussed, the cross tabulation in Table 45 helps indicate the differences in project performance for sectors for each of the regions. There are several possible reasons for such differences in project performance and offers an opportunity for further research. For example, EAP and ECA, have consistently high percentage of successfully rated projects across all sectors. The AFR region has a low percentage of successful projects across sectors with the exception of high ratings for the agriculture, water and sanitation and multisector sectors (World Bank Group, 2020). The results support the WBG's observation that projects face additional challenges in FCV countries across all sectors (McKechnie, 2017). Secondly, the results support the ADB categorization between infrastructure and non-infrastructure projects and differences in performance. The ADB and JICA acknowledge the limitations of the project evaluation indicators in their annual reports and commit to continually improve and adapt the evaluation methodology, process and criteria (Asian Development Bank, 2019; JICA, 2020). Research findings support current knowledge in evaluation reports regarding the high project performance for EAP, ECA and SAR and in the energy, transportation and water sanitation sectors (Asian Development Bank, 2019; World Bank Group, 2020). Santos et al. (2020) argued that health projects have lagged behind other sectors and observed that urgency and lack of adequate scope and specification contribute to poor performance. Asiedu et al. (2021) argued that management of an effective procurement process includes attention to strategic planning, risk management, efficiency, supplier evaluation and relationship management for projects in infrastructure projects. In a similar way, the 2020 WBG results and performance report noted that maintenance of a clear focus on project quality, cost performance and effective delivery to beneficiaries contributed to positive outputs and outcomes as observed for the agriculture and environmental sectors in the report.

### 5.2.9 *Project evaluation criteria ratings*

As discussed in Section 3.3, the overall project evaluation results, in this research, are comprised of four criteria of relevance, effectiveness, efficiency and sustainability, based on OECD evaluation methodology. As discussed, the coherence criteria was introduced in 2019 and the organisations did not measure the impact criterion (OECD, 2019).

The following section reviews the mean project evaluation results by each of the project evaluation criteria. To provide further analysis in answer to Question 2, to understand how project evaluation measures DE and VFM, this section reviews the relative performance of the evaluation criteria and how they contribute the overall project success, DE and VFM.

Data for the evaluation criteria was available, as published secondary data, for eight of the ten international development organisations for analysis in this research.

The results follow the same rank order for the overall project evaluation results, shown in Table 37, the project evaluation results by region, shown in Table 40, and the project evaluation results by sector, shown in Table 43.

The OECD definitions of the evaluation criteria, as discussed in Section 3.10, vary between organisations, for example, there is no consensus on the definitions of the evaluation criteria, such as the sustainability criterion (Khagram et al., 2009). As a result, different interpretations of the criteria have implications for reliability and consistency for reporting and comparison of project evaluation results (Bayiley & Teklu, 2016).

For example, OECD, 2019 defines the relevance criteria as “the extent to which the intervention objectives and design respond to beneficiaries’, global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change” (OECD, 2019 pp. 5-12). As the author of evaluation reports are often also the project managers or Task Team Leaders (TTLs), it is unlikely that they will give

negative ratings on the relevance of the design and objectives for their own projects (de Montclos, 2012). Secondly, in many cases the organisation only uses two rating options of highly relevant and relevant leaving no possibility of more critical or negative ratings and hence the ratings for relevance are consequently skewed higher than the other criteria (OECD, 2019). Effectiveness refers to measuring the extent to which an aid activity attains its objectives and, once again, is a broad definition. However, the effectiveness criterion tends to incorporate more detail in project evaluation reports, including limited information on expenditure and costs. Efficiency is defined as the measure of the qualitative and quantitative outputs in relation to the inputs and includes more data and discussion on cost performance and expenditure (Chianca, 2008). The efficiency evaluation criterion often includes Economic Internal Rate of Return (EIRR) as a measure of VFM, but this measure is not always appropriate for many smaller projects (Escadafal, 2014).

The efficiency evaluation criterion is often rated inconsistently, has minimal detail on cost performance and frequently has the lowest rating and results of the four criteria in this research data (Clements et al., 2008). The equivalent definition of effectiveness and efficiency, in the private sector and industry, is specific to the “iron triangle” Critical Success Factors (CSF) of project cost, schedule and quality as discussed in Section 3.5 (Atkinson, 1999). By comparison, while the cost is discussed in the narrative of project evaluation reports in international development organisations, cost and Value for Money (VFM) do not appear to contribute significantly to the overall project evaluation ratings and results (Sundqvist et al., 2014).

Sustainability concerns the measure of whether benefits of an activity are likely to continue after donor funding has been withdrawn and, like the relevance criterion, sustainability is subjective and difficult to measure (Chianca, 2008). WBG and EBRD do

not disclose results for the sustainability criterion in their respective secondary data, and this criterion is not included in the analysis their evaluation criteria in this research.

Table 46 breaks down the project evaluation results into the four evaluation criteria to provide the evaluation results and contribution for each criteria towards the overall project evaluation rating and measurement of DE. In addition, the analysis of the criteria provides further detail on how individual criteria measure cost performance and VFM. Table 46 and Figure 19 shows the different mean results for the evaluation criteria, including the mean evaluation results for the four criteria, included in this research and the overall project evaluation mean results. The evaluation criteria are combined to derive the overall evaluation score. However, as discussed in Section 3.10, the combination of the criteria to calculate the overall project evaluation result, differs for each international development organisation (Clements et al., 2008; Lloyd et al., 2014). The broad definition of the OECD evaluation criteria, inconsistent application of rating scales and different consolidation methods, contributes to challenges in the comparability and accuracy of the overall project evaluation results (Clements, 2020). Results of this analysis suggest that the OECD evaluation methodology would benefit from improvements in definition and consistency for more meaningful and comparable project evaluation results across international development organisations (Chianca, 2008).

TABLE 46. PROJECT EVALUATION CRITERIA RATINGS

Results <i>M</i>	WBG	ADB	AFDB	EBRD	IADB	ISDB	UNDP	JICA
Relevance	.636	.748	.846	.741	<b>.777*</b>	.727	.675	.974
Effectiveness	.542	.616	.686	.604	.508	.553	.782	<b>.846</b>
Efficiency	.455	.591	.671	.466	.579	.477	<b>.749</b>	.546
Sustainability		.641	.694		.540	.517	<b>.819</b>	.704
Rank by <i>M</i>	7	4	2	6	5	8	1	3

\*The values in bold, Table 46, indicate the highest mean project evaluation criteria



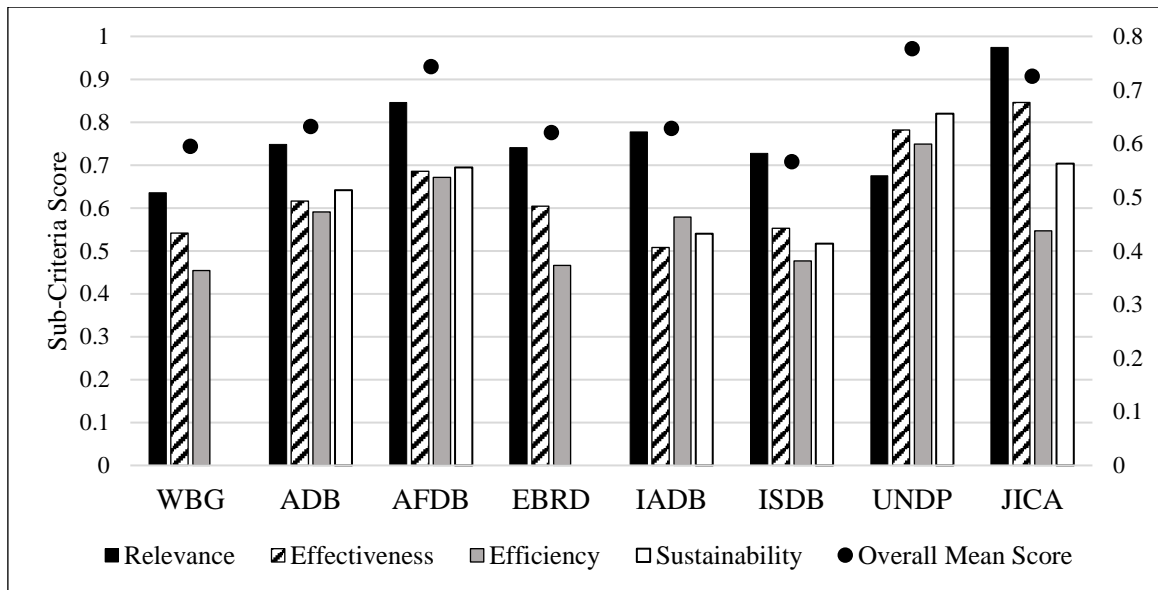


FIGURE 19. TOTAL PROJECT EVALUATION CRITERIA RATINGS

As discussed in Section 3.5, project performance is measured in the private sector and industry by several key criteria including; leadership, policy, processes, performance matrices and resources and design, planning, implementation and completion (Khang & Moe, 2008; Mir & Pinnington, 2014).

The measurement of project performance in the public sector, private sector and international development, have common factors that include similar indicators that measure relevance, effectiveness, efficiency and sustainability, but involve differences in definition and detail (Jugdev & Müller, 2005). The private sector has a greater emphasis on measuring specific aspects, such as leadership and resources, design, planning, implementation, performance and completion metrics (Cooke-Davies, 2002). By comparison, the definitions for the evaluation criteria in the OECD methodology, are more difficult to interpret and evaluate, compared to the objective and tangible nature of the iron triangle indicators and criteria (Frefer et al., 2018). For example, macro indicators in the education sector may include broad metrics, such as, the numbers of school children attending school country-wide after schools are built. An example in the health sector might include regional reduction in mortality from malaria, after the

delivery of bed nets, in a given country or region (Howes et al., 2011; Mosley, 1986).

The macro indicators are essential for assessment of the concepts and goals of aid effectiveness, but do not measure direct project implementation, quality, schedule and how the money was spent in the delivery of specific projects (Easterly & Pfitze, 2008). Project-specific and objective micro-level indicators serve distinctly different purposes to the more subjective nature of the macro-level indicators, but they are equally important and need to be measured and reported (Denizer et al., 2013; Lim & Mohamed, 1999).

#### *5.2.10 Cost estimation and actual costs*

Five of the ten international development organisations have publicly available secondary data on project cost estimations and the completed actual project costs in the published project evaluation reports. Some organisations have vast differences in consistency of reporting and measurement of the cost estimates and actual completed project costs. There is also a considerable difference in the variation between actual project costs above and below the Cost Estimate (CE) between each of the selected international development organisations, reviewed in this research. A set of thresholds, or bands, relative to the CE, is established to formulate a method to illustrate for the degree of variation of actual cost from CE between the organisations in the research. Table 47 shows these five thresholds developed to measure number of project that have overruns and underruns and include intervals between 0 to 10%, 10% to 25%, 25% to 50%, 50% to 100% and more than 100% above and below the CE. The thresholds are intended to help understand and show the frequency of actual project costs above and below the CE (Bayram & Al-Jibouri, 2018; Klakegg & Lichtenberg, 2016). Table 47 and Figure 20 show the relative frequency and percentage for projects that either meet the CE, or above and below (overrun or underrun), by the comparison of the actual project costs with the project CE for each of the different thresholds (Ahsan & Gunawan, 2010).

Unfortunately, these data are not available for all ten of the organisations included in previous analyses. As such, analyses have been provided for the five available data sets.

TABLE 47. PERCENTAGE VARIATION OF ACTUAL COSTS FROM COST ESTIMATE

Variation from Cost Estimate	WBG		ADB		ISDB		UNDP		DFID	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
0 to 50%	55	7.8%	3	2.1%	2	4%	16	10.7%	1	2.7%
50 to 75%	84	11.9%	8	5.5%	2	4%	11	7.3%	6	16.2%
75 to 90%	96	13.7%	20	13.8%	6	12%	30	20%	4	10.8%
10% below CE	186	26.5%	61	42.1%	19	38%	37	24.7%	8	21.6%
Cost Estimate	45	6.4%	35	24.1%	5	10%	12	8%	6	16.2%
10% above CE	78	11.1%	14	9.7%	4	8%	11	7.3%	2	5.4%
110 to 125%	51	7.3%	2	1.4%	4	8%	9	6%	1	2.7%
125 to 150%	44	6.23%	2	1.4%	4	8%	8	5.3%	2	5.4%
150 to 200%	32	4.6%	0	0%	2	4%	8	5.3%	6	16.2%
> 200%	31	4.4%	0	0%	2	4%	8	5.3%	1	2.7%
<b>Total</b>	<b>702</b>	<b>100</b>	<b>145</b>	<b>100</b>	<b>50</b>	<b>100</b>	<b>150</b>	<b>100</b>	<b>37</b>	<b>100</b>

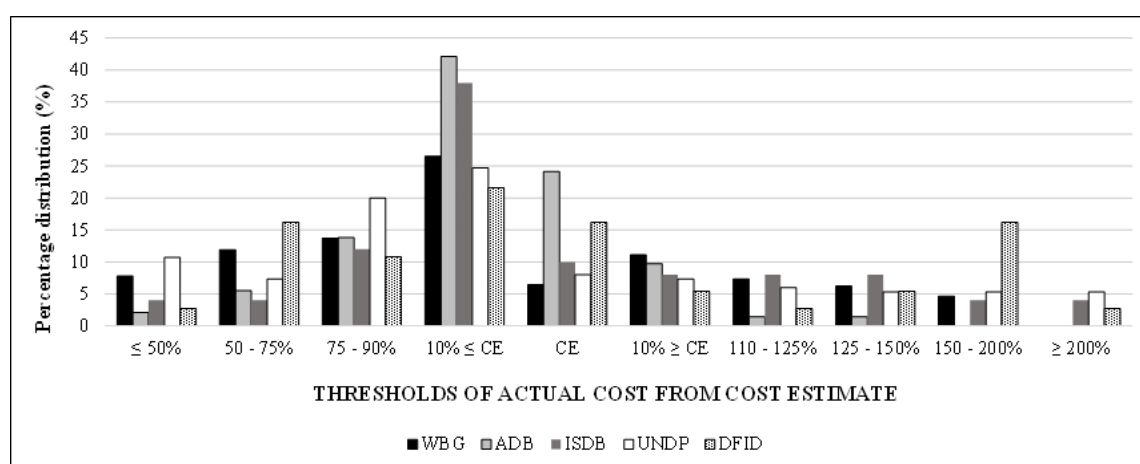


FIGURE 20. PERCENTAGE COST VARIATION BY THRESHOLD

The results in Table 47 show that all international development organisations have more project underruns than overruns and approximately 13% of projects on average, have actual costs that meet the CE precisely. The reason for such a precise match for complex project management activities is an interesting observation for further analysis. The percentage of projects with actual costs below the CE, an underrun, is 60% for WBG, 62.72% for ADB, 57.38% for ISDB, 62.27% for UNDP and 51.35% for DFID.

Close alignment, or actual costs that exactly meet the CE, is often regarded as the result of good project management performance, although it may also be due to other factors. These factors may include different and less-stringent measurement and poor monitoring and evaluation practices for the monitoring of costs, that apply to some development projects (Glassman et al., 2013). Frequent cost underrun and delay in projects, is thought to be caused by factors such as poor cost estimates, currency fluctuation, scope changes, design changes and sometimes due to the effect of competitive bidding (Ahsan & Gunawan, 2010). The results of this research, shown in Figure 20, show a spike, higher trend, for the 90% to 100% threshold, which is a threshold that falls narrowly below the CE. Many organisations, particularly in the private sector, classify actual costs, that fall within  $\pm 10\%$  of the CE, as a positive indicator for cost variation in projects and record cost results in this particular threshold as savings (Odeck, 2004). In a similar way, Bertisen and Davis (2008) observed that approximately half of projects fell outside of the feasibility error of  $\pm 15\%$  of the CE, defined as the optimal amount of error for cost estimates in project engineering in the resources industry. For project actual costs that fall outside of the CE, Andrić et al. (2019) reviewed infrastructure development projects in Asia and observed that 60% of projects experienced overrun and 40% experienced underrun. The authors attributed cost variation to design changes, material costs, currency fluctuation, consultant cost and unexpected geological conditions among others. The results in Table 47, show that 40% to 50% of actual costs for the international development organisations fall in the threshold of 10% above and below the CE. While this indicates that there is a high proportion of projects that have cost variations outside of the 10% threshold from the CE, and may appear as far from ideal, it is not unusual in projects in the public sector and industry (Flyvbjerg et al., 2018; Odeck, 2004).

Baccarini (2004) observed that contingencies for project estimates are frequently calculated within 10% of the CE, and this also indicates that this threshold is a common, if not a desirable, target for the variation of actual costs from the CE (Baccarini, 2004; Baccarini & Love, 2014). Figure 20 shows cost variation by the difference of actual costs from CE by different thresholds for the five international development organisations.

As discussed in Section 3.5, cost, schedule and quality are the recognised Critical Success Factors (CSFs) serve as indicators for project efficiency and effectiveness in the industry and the private sector (Vrchota et al., 2020). Both quality and schedule influence costs, and therefore, the analysis of cost performance; a concept term used by Love et al., 2018, serves as useful perspective for measuring cost and VFM.

The thresholds, or bands, allow the possibility of comparison of the organisations' project management cost performance through the comparison of accuracy of the cost estimation process and the ability to align actual costs to meet the CE on project completion. The benefit of the analysis helps to indicate the effectiveness of the cost estimation process and the adherence of the actual costs with the original budget and Cost Estimate (CE).

As discussed earlier in this section and indicated in Table 47, there is a greater frequency of project cost underruns for the international development projects evaluated in this research. Ahsan and Gunawan (2010) noted that high incidence of cost underruns is thought to be caused by several factors, that include poor cost estimates, currency depreciation, scope changes, and government approvals (Andrić et al., 2019).

The variation between CE and actual cost is indicated by the division of the actual cost by the CE and this provides a relative and proportionate measurement for cost variation.

Table 48 shows the mean cost variation of actual costs divided by the CE by organisation

and provides an overall perspective of overrun against underrun (Love et al., 2018). For example, where the actual cost is equal to the CE, the mean cost variation value is 1.0.

TABLE 48. STATISTICAL DISTRIBUTION FOR ACTUAL COSTS FROM COST ESTIMATE

<i>Statistic</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	Skewness	Kurtosis
WBG	1.021	.978	.557	6.008	74.158
ADB	.941	.985	.138	-2.828	16.435
ISDB	1.113	.982	.615	4.247	19.834
UNDP	1.107	.952	.992	5.170	31.816
DFID	1.055	.991	.425	.661	.364

The mean results are all above the median, with the exception of ADB, which suggests a greater value of project overruns; the positive skewness suggests that there is a greater frequency of project underruns. ADB, as the exception, has a mean result under the median and is negatively skewed and the results suggest that ADB has more project underruns and projects that meet the Cost Estimate (CE). The range of means above and below the CE is from 5.9% below the CE for ADB to 10.7% above the CE for UNDP. The standard deviation for ADB is lower than for the other organisations and this also supports the observation ADB has a relatively large number of projects close to the CE. UNDP and ISDB follow a similar profile to WBG for means, median, and skewness although UNDP has a larger standard deviation which may reflect the greater incidence of both overruns and underruns. DFID has a flatter distribution and low kurtosis which suggests that actual costs for DFID are more evenly distributed above and below CE. The difference in the means and standard deviations and the lower frequency of cost overruns and underruns for cost variation suggests that ADB has a different approach to cost performance. The approach aligns with Andrić et al. (2019) findings that actual costs close to the CE and low standard deviations indicate positive cost performance. This is an important observation and suggests that project managers and TTLs from

different organisations follow different approaches regarding cost estimation and the management of actual project costs. The implication of findings is that a low number of overruns and underruns and high frequency close to the CE, indicates positive cost performance with more accurate cost estimation and effective monitoring of actual costs.

To test the association of actual project costs from the CE, a Chi-Squared test of independence ( $\chi^2$ ) was conducted. In performing this test it was assumed that the optimal objective, or target, for actual costs be equal to CE, or within the ( $\pm 10\%$  threshold) of the CE as a measurement of project cost performance. The Chi-Squared test of independence ( $\chi^2$ ) was applied in this research to understand the association between actual costs that fall within or outside the optimal target threshold. The Chi-Squared ( $\chi^2$ ) test of independence results were  $\chi^2(4, 1094) = 53.94, p = .000$ . The results show a statistically significant association between the variables for the different international development organisations. The results therefore suggest that the five international development organisations have alignment with the cost performance target across the different thresholds. Four of the organisations have between 40% to 50% of actual costs within the ( $\pm 10\%$  threshold) and ADB has 76% of actual costs within this threshold.

Andrić et al. (2019) observed that there was correlation between project size and cost variation, most notably cost overruns, in East Asia and South Asia but no such correlation in the Central Asia region as distinguished in the research. Welde and Odeck (2017) observed that cost overruns were correlated with smaller projects and hence it appears observations differ for various studies on project size (Love et al., 2018).

The results in this research show that 627 projects had underruns in the sample as compared to 324 overruns and 103 projects where the actual costs exactly met the CE. Projects where actual cost meet the CE, may be due positive cost performance or other

possible reasons potentially related to contract structure, method of payment and disbursement of funds, such as lump sum or fixed price payment terms (Vu et al., 2020).

To examine the relationship between project size and cost performance in this research, a Pearson's correlation test was performed on project size or value and cost variation. In addition, a Pearson's correlation test was conducted for overruns and underruns against project size by value to understand whether contract size influences cost variation. The results of the test was  $r(1084) = .184, p = 000$  and this indicates a very weak positive correlation for project size by value against cost variation. The very weak positive correlation results partly supports current knowledge and the observations by Andrić et al. (2019) that there are slightly more cost overruns for larger projects. To pursue the analysis for project size further a Pearson's correlation test was conducted between overruns, underruns and project size by value of the ( $n = 1,084$ ) projects in this research. The result of the Pearson's correlation test for overruns was  $r(324) = .151, p = .006$ . The result of the Pearson's correlation test for underruns was  $r(627) = .030, p = .448$ .

The Pearson's correlation results show a weak positive correlation between overruns and project size by value and the result was significant. For underruns, however, the results show almost no correlation between underruns and project size and the result was not significant. There is a slight correlation between overrun and project size which suggests that larger projects may be more complex and difficult to control costs. Underruns, on the other hand, were not correlated with project size and occur more randomly and may have many different causes in relation to project size and value (see Table 49).

As discussed in Section 3.10 and Section 5.2.10, each organisation follows individual modifications and adaptations to the OECD-DAC project evaluation methodology, and these differences may also include their standards and approach to the rating process.



The same variation applies at the project manager and TTL level, and introduces distortions and potential bias in the results, and these distortions emerge, particularly in this research, when results are compared alongside each other (Denizer et al., 2013). Despite the absence of cost data capture, international development organisations do use cost estimation for the quantification, budget preparation and description of development projects as part of the complex assessment and approval process (Khang & Moe, 2008). Cost Estimates (CEs) are also used in practice at all stages of the procurement process, project management, monitoring and measurement of success for projects and programs (Flyvbjerg et al., 2018; Long, Ogunlana, Quang, & Lam, 2004).

As discussed, the OECD project evaluation criteria of effectiveness and efficiency include cost data, however, the estimation and documentation of the actual costs, included within these two evaluation criteria, are often inconsistent and inaccurate (Clements, 2020; Khagram et al., 2009). Many international development projects have durations that cover several years, and the project schedules frequently face delays resulting in poor completion results and cost underruns. In international development projects, minor underruns are considered as a positive achievement and as a saving against the CE. Ahsan and Gunawan (2010) noted that larger cost underruns are often associated with schedule delays and can also signify project procurement issues, government approval issues, inaccurate CEs, or failure to complete specific parts, or all of the scope. It is also notable that, if the cost estimate is not built from accurate analysis and good market intelligence, then the comparison with actual costs has little meaning as a measurement regarding success or failure of projects (Doloi, 2013). There is inconsistency in the cost estimation and the measurement of actual costs in project evaluations of the different international development organisations as tabulated in Table 47, in this research. Secondly, the analysis also reflects the project management

performance regarding monitoring cost, schedule and quality during the contract management and implementation of the project. The cost variation, accuracy of the CE, and the project management skills and expertise to deliver projects successfully, are important component factors leading to the results shown in Table 47 and Table 48. The causes of project overrun and underrun of cost and schedule are often not recorded in detail by the five international development organisations in project evaluation reports and the causes and impact are often difficult to determine from the narrative. However, Ahsan and Gunawan (2010) identified several causes of cost overrun, underrun and schedule changes in ADB project evaluation reports. Similarly, as observed in this research sample, ADB included more consistent and detailed cost performance data. There are often multiple causes of cost overrun and underrun for each project and the description of the main causes in the sample are tabulated and shown in Table 49.

TABLE 49. CAUSES OF COST OVERRUN AND UNDERRUN IN ADB PROJECTS

Description	Frequency Overrun	Frequency Underrun
Cost escalation/decrease (Materials/Labour)	14	35
Design/Scope change	7	38
Currency depreciation/appreciation	6	19
Unknown geological/hydrological	2	
Natural Disaster/Civil unrest		2
Procurement efficiency/delays	3	10
Government loan issues/delays	5	19
Competency of contractor/sub-contractor	5	
<b>TOTAL</b>	<b>42</b>	<b>123</b>

In this research there were 18 ADB project overruns, 92 ADB project underruns and 35 ADB projects that meet the CE and some projects had more than one cause indicated. In this research sample, the causes for project cost overruns and underruns broadly fall into eight different categories as ADB is indicated in Table 47. The descriptions for overruns and underruns often have more than one explanation and authors such as Vu et al. (2020) identified and ranked 31 causes of overrun and underrun (Kim et al., 2018).

The most common cause of overruns are cost increases, scope changes and currency appreciation. The most common causes of underruns are from design and scope reduction or cancellation of project scope and secondly, from cost decreases and efficiencies for materials, equipment and labor followed by currency fluctuation. This aligns with Ahsan and Gunawan's (2010) work who observed a high frequency of project delays, which have cost underruns in international development projects.

Frefer et al. (2018) argued that cost, schedule and quality are important critical success factors for measuring project success and must be properly captured. An important observation and conclusion from this research, is that the development of reliable and realistic CEs, and the accurate measurement of actual project costs, are essential for the understanding of how development funds are spent and where the money to ensure that existing projects are effective and future projects are properly planned, estimated and funded (Easterly & Pfitze, 2008; Frefer et al., 2018).

The previous section has provided analysis of the project evaluation results for the international development organisations by year, region and sector. The individual evaluation criteria are further analysed to understand their relative performance and how they contribute to the overall project evaluation results. Finally, to understand how cost parameters are evaluated for projects, further analysis is conducted on cost variation (actual projects costs against CE), for projects where the secondary data are available. The following section examines the relationship between procurement and project management as the mechanisms of project delivery for international development. As discussed in Section 3.13, the combination of effective maturity and performance of both processes is required for project success leading to DE and VFM (Ika et al., 2010).

### 5.3 CORRELATION OF THE PROJECT EVALUATION MEAN RESULTS

To understand the relationship between the procurement and project management processes, in answer to Research Question 3, it is necessary to examine project evaluation criteria for reference to the procurement process maturity and performance. As mentioned in Section 3.9, project evaluation reports lack detail on the procurement process and refer to process compliance rather than provide details of procurement maturity and performance. There is, therefore, no direct relationship between the measurement of procurement and project performance evaluation in the OECD project evaluation methodology (Ahsan & Paul, 2018; Lamhauge et al., 2012; OECD, 2019). Aulia and Isvara (2021) argued that procurement performance is a combination of efficiency and effectiveness criteria defined on the basis of a cost to value equation. The cost to value equation combines cost with other value factors such as quality, output and Total Cost of Ownership (TCO) (Axelsson, 2005). The cost estimation process and the active monitoring actual costs in projects are important responsibilities of the procurement function required for the measurement of value (Mimović & Krstić, 2016). Love et al. (2018) define and apply the term cost performance as the difference between final project cost and approved project budget or Cost Estimate (CE). The authors suggest that there is a link between procurement, cost variation, cost performance and Value for Money (VFM). The research findings indicate that there is indeed a link between procurement, cost variation and project performance. However, while the project evaluation reports do indeed contain limited reference to all of these factors, they do not measure their impact and contribution to the overall project performance, DE and VFM. The following section examines the correlation and relationship between the individual project evaluation criteria and overall project evaluation results.

Table 50 summarises the mean results for the overall project evaluation results and individual evaluation criteria results for the ten international development organisations; it is a collation and summary of mean data from Section 5.2.

TABLE 50. SUMMARY OF PROCUREMENT AND PROJECT EVALUATION RESULTS

<i>M Results</i>	WBG	ADB	AFDB	EBRD	IADB	ISDB	UNDP	UNICEF	DFID	JICA
IPMM	3.025	2.932	3.039	2.169	2.585	2.136	2.686	3.028	3.203	2.763
Project Evaluation	.595	.635	.740	.617	.628	.567	.777	.576	.527	.725
Relevance	.636	.749	.846	.741	.741	.727	.675			.974
Effectiveness	.542	.617	.686	.604	.508	.553	.782			.846
Efficiency	.455	.591	.671	.466	.579	.477	.749			.546
Sustainability		.641	.694		.540	.517	.820			.703

Eight of the international development organisations publish secondary data for the evaluation criteria of relevance, effectiveness, efficiency and sustainability, and these include WBG, ADB, AFDB, EBRD, IADB, ISDB, UNDP and JICA. As discussed in Section 5.2.9, the evaluation criteria are used to determine the overall evaluation results following the OECD evaluation methodology; the mean results for the project evaluation criteria are shown in Section 5.2.9, tabulated in Table 46 and illustrated in Figure 19.

Five organisations provide publically-available data on the cost estimates and the actual cost data for all projects, and these include WBG, ADB, ISDB, UNDP and DFID. The cost estimation and actual costs are discussed in Section 5.2.10, tabulated in Table 47.

A Pearson's correlation test was conducted with the evaluation criteria to understand how each individual criterion contributes to the overall project evaluation results. The overall project evaluation results comprise four evaluation criteria, as discussed and defined in Section 3.10, and the test was conducted to explore further the contribution of the evaluation criteria to the project evaluation results as shown in Table 51.

TABLE 51. CORRELATION OF PROJECT EVALUATION RESULTS AND COST DATA

<i>Correlation</i>	Project Evaluation	Relevance	Effectiveness	Efficiency	Sustainability
Project Evaluation	1				
Relevance	.428**	1			
Effectiveness	.709**	.477**	1		
Efficiency	.646**	.289**	.415**	1	
Sustainability	.519**	.149**	.315**	.061	1

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

The Pearson's correlation coefficient for evaluation criteria shows a positive correlation, at the < 0.01 level of significance (2-tailed), for both effectiveness and efficiency and moderate correlation with the sustainability criteria and the overall mean project evaluation results. The positive and moderate correlation with the overall project evaluation results indicate that effectiveness, effectiveness and sustainability make a substantial contribution to the overall project evaluation results. The OECD (2019) definition of the efficiency evaluation criterion, discussed in Section 3.3, is “the extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way”, and includes consideration for both cost and schedule. The association between the efficiency criterion and the Critical Success Factors (CSFs) such as cost performance, as discussed in the previous sections, implies that it contains characteristics in common with procurement maturity. The definition of effectiveness is “the extent to which the intervention achieved, or is expected to achieve, its objectives, and its results”. While contract management is a phase of the procurement process, the delivery of the objectives described by the effectiveness criterion is largely implemented by the project managers and TTLs. The sustainability evaluation criterion, defined as “the extent to which the net benefits of the intervention continue or are likely to continue”, takes a long-term perspective of

project success in common with good procurement practices and performance (Mena et al., 2014; OECD, 2019 pp. 5-12). The results in Table 51, indicate differences in the relative contribution of individual evaluation criteria towards overall project evaluation results and that the efficiency criterion, in particular, incorporates measures of cost. However, as discussed in Section 5.2.10, the estimation and documentation of the actual costs, included within the evaluation criteria, are often inconsistent and inaccurate (Clements, 2020; Picciotto, 2020). Of the four project evaluation criteria, the efficiency evaluation criterion is described in terms of cost by Chianca (2008) as a measure of the outputs that use least-cost resources to achieve the desired objectives (Chianca, 2008). The statistically significant moderate positive Pearson's correlation for all the evaluation criteria reflects the relationship between the evaluation criteria and their respective contribution towards the overall project evaluation results. The results indicate that the effectiveness and efficiency two criteria provide a greater contribution to the overall project evaluation results than the other criteria respectively as follows:  $r = .709, p = .000, N = 1610$  and  $r = .646, p = .000, N = 1610$ . The relevance evaluation criterion shows a high correlation with effectiveness and a lower correlation with the other evaluation criteria. As discussed in Section 5.2.9, the relevance evaluation criterion is consistently and uniformly rated as relevant in the project evaluations with little variability or flexibility of the ratings applied to project performance and the overall project evaluation results. A better definition of the relevance evaluation criterion for the OECD evaluation methodology would provide more depth and value for the project evaluation process. As it is applied, the relevance criterion does not add a great deal of comparative value to the evaluation results in this research (Hák et al., 2016).

As discussed the OECD-DAC project evaluation methodology does not consistently define cost performance in the project evaluation reporting, or within the individual

project evaluation criteria (Khang & Moe, 2008). The finding suggests that there is an opportunity for further research and to drill down and explore the analysis of the project evaluation criteria to understand cost performance and VFM in development projects. To further explore Research Question 3, and possible link and connection between the effectiveness, efficiency evaluation criteria and overall project evaluation results, a regression analysis was conducted and this is described in the following section.

#### **5.4 REGRESSION ANALYSIS OF THE PROJECT EVALUATION MEAN RESULTS**

The regression analysis herein uses the overall project evaluation results as the dependent variable and the project evaluation criteria and cost variation results as independent variables, or predictors, as shown in Table 52.

The overall project evaluation mean results shown in Table 50 provide a summary measurement by mean of the project evaluations in the data set ( $n = 1,920$ ) and represent total project success for each organisation and organisation type. The overall project evaluation results are derived from the OECD's (2019) four component evaluation criteria of relevance, effectiveness, efficiency and sustainability. Table 52 shows the results of a regression analysis conducted to review the relative influence of the four evaluation criteria as predictors towards developing the overall project evaluation results. The overall mean project evaluation results are represented as the dependent variable, and the evaluation criteria result as the independent variables, or predictors, in the regression analysis and the regression results are shown in Table 52.



TABLE 52. REGRESSION ANALYSIS FOR OVERALL PROJECT EVALUATION AND CRITERIA

<i>Variable</i>	<i>R</i>	<i>Rsq</i>	<i>AdjRsq</i>	<i>Std Error</i>
Relevance	.428	.184	.183	.19781
Effectiveness	.709	.503	.502	.15439
Efficiency	.646	.417	.416	.16726
Sustainability	.519	.269	.269	.18528

*Note: the dependent variable is the overall project evaluation mean results (Table 52).*

The strong regression *R* results for the effectiveness (.709) and efficiency (.646) criteria, and moderate results for the relevance (.428) and sustainability (.519) evaluation criteria reflect the relative importance of the effectiveness and efficiency evaluation criteria as predictors of the overall project evaluation results. The relevance evaluation criterion shows the lowest regression variable (.428) of the evaluation criteria which indicates that it contributes less to the overall project evaluation results. This finding supports and confirms the earlier discussion in the literature review and the correlation results shown in Table 51 (Clements et al., 2008). A multiple *R* of 1 indicates a perfect linear relationship, while a multiple *R* of 0 indicates that there is no relationship between the dependent variable and the variable (Saunders et al., 2016, p. 547)

The *R-sq* result indicates that 50.3% and 41.7% of the variance of overall mean project evaluation results can be predicted from both the effectiveness and efficiency evaluation criteria, respectively. The *R-sq* result for the sustainability evaluation criterion is 26.9%, and for the relevance evaluation criterion, 18.4% indicates the corresponding variance in overall mean evaluation results by the variables of sustainability and relevance. The strong significance values shown in Table 52, supports the finding that the effectiveness and efficiency evaluation criteria have a statistically significant relationship as predictors of the overall mean project evaluation results. The summary of mean results in Table 50 and the correlation results in Table 51, also confirm that the effectiveness and efficiency evaluation criteria have significant correlation with the overall project evaluation results.

As a reminder of the OECD definitions, the definition of effectiveness, as discussed in Section 3.10, is “the extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups”. The definition of efficiency is defined as “the extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.” ( OECD, 2019 pp. 5-12).

As discussed in Section 3.9, the efficiency evaluation criterion includes cost and often Economic Internal Rate of Return (IRR) indicators, and therefore include measurement of spending activity, in the hierarchy of performance measurement, shown in Figure 2.

As discussed in Section 3.5, the cost variation results are often used as indicators of project cost performance and VFM, in the private sector and industry, as part of the iron triangle including cost, schedule and quality (Atkinson, 1999; Zidane and Olsson, 2017).

However, for international development projects, the OECD project evaluation criteria do not adequately capture cost-related indices, and these are often inconsistently measured and not well defined (Chianca, 2008; Clements, 2020).

The analysis of cost variation, as a measure of cost performance in this research, suggests that there is a continuous thread, or link, that connects the effectiveness and efficiency project evaluation criteria with Value for Money (VFM) (Love et al., 2018).

The previous section describes the analysis of the regression between the overall project evaluation results and evaluation criteria in the research. The findings indicate that the effectiveness and efficiency criteria make the largest contribution towards the overall project evaluation result and rating. The research suggests that there is a common link between cost performance and the evaluation criteria and that development of the evaluation criteria can provide a meaningful way to measure cost performance and VFM.

## 5.5 SUMMARY

Chapter 5 provides the research results and details the data analysis and the statistical approach to addressing the research questions, as illustrated in Figure 10 and Figure 11. The chapter provides the results of the procurement maturity IPMM questionnaire results in Section 5.1, to address the first research question to compare the procurement maturity of leading international development organisations, from the perspective of DE and VFM, and to provide comparative analysis of the original IPMM with weighted IPMM results for ten international development organisations. Additional comparative analysis is performed of the IPMM results for types of the international development organisation and includes benchmark results from three private sector organisations and shown in Table 30 and Table 31. The findings indicate that DFID, AFDB and UNICEF have the best comparative procurement capacity and maturity. Secondly, the results indicate that the private sector has better procurement capacity and maturity than the development organisations. The international development organisations are ranked by procurement maturity in order of the bilateral organisations first, followed by the UN Agencies and then the MDBs. The procurement maturity results for all organisations, are shown in rank order in Table 25, and maturity results for the weighted IPMM results, in Table 27. As discussed in Section 5.1.9, the MDBs are organised into two functional procurement groups, including the operational and corporate procurement groups, and the comparative maturity results for the different group are shown in Table 32. The corporate procurement group is responsible for the management of internal corporate procurement activity and requirements. The operational procurement group provides oversight, governance and capacity building for procurement activity performed by borrowing, or recipient countries, and their Implementing Partners (IPs).

The IPMM results are further analysed by the four procurement phases and 12 procurement criteria to understand the comparative capacity, maturity and influence of VFM in the procurement process and the results are shown in Table 33 and Table 34.

The results from the additional analysis suggest that organisations that combine direct procurement activity with oversight functions, such as the bilateral organisations and UN Agencies, have improved procurement capacity and maturity. Secondly, for the analysis of the procurement phases and criteria, the procurement strategy and regulatory framework phases and the associated criteria make a relatively strong contribution to the procurement process. The results suggest that procurement strategy and the regulatory framework phases and associated criteria have a more substantial contribution and influence on procurement maturity and VFM. The chapter goes on to provide the project evaluation results to address the second research question, which seeks to understand how existing project evaluation assessment methodology measures project performance, DE and VFM. The research performs a comparative analysis of a sample of ( $n = 1,920$ ) project evaluations by organisation, region and sector for the same ten international development organisations and discussed in Section 5.2. The research provides analysis of cost variation results for five development organisations and the implications of the relationship and link with cost performance, procurement and VFM. Cost performance is explored in this research as a critical but somewhat overlooked measurement of Development Effectiveness (DE), which has the potential to connect the procurement process with project performance, DE and Value for Money (VFM).

To address the third research question to examine the relationship between procurement maturity measurement and project evaluation, the project evaluation criteria and cost variation is examined to understand how these two different processes are measured.

Project evaluation methodology and the evaluation criteria do not measure either the

procurement maturity, performance or cost performance indicators in detail. The results suggest that improved evaluation definitions and methods should strengthen the visibility and measurement of procurement, cost and VFM parameters.

To address the fourth research question on whether the maturity and performance of the procurement and project management processes contribute towards DE and VFM in development projects. The project evaluation results are broken down to indicate the relative contribution that the individual project evaluation criteria make towards the overall project evaluation results. The results suggest that the effectiveness and efficiency criteria make the greatest contribution to the overall project evaluation results and that the efficiency criterion can be improved to better measure DE and VFM.

Finally, a regression analysis was conducted to examine the different evaluation criteria as predictors of the project evaluation performance results, as shown in Table 51. The different definitions and attributes of the individual evaluation criteria provide distinct contributions to the overall project evaluation results. For example, the efficiency and effectiveness evaluation criteria contain cost-related data and provide a potential for improving the project evaluation methodology for the measurement of DE and VFM. The findings and implications of the results are discussed in the following Chapter 6.

## **6 DISCUSSION**

This chapter provides a reflection on how the research results address the research questions and demonstrate whether research findings support existing knowledge or new knowledge has emerged about procurement and project performance, DE and VFM.

### **6.1 INTRODUCTION**

The research questions and aims and objectives of this research, specified in Section 3.14 and Section 3.15, seek to provide an analysis of the measurement of procurement maturity and project performance in development projects. Secondly, the research reviews the relationship and influence of the two key project delivery mechanisms to Development Effectiveness (DE) and Value for Money (VFM). The literature review and research design chapters, in Chapter 3, and Chapter 4, include discussion and analysis of the measurement and evaluation of procurement maturity and project performance in international development projects. The results in Chapter 5 provide further analysis on the measurement of the two processes, the relationship between them and explores their mutual contribution towards DE and VFM in international development projects.

Biscaye et al. (2017) depicted the hierarchy of performance measurement shown in Figure 2, which indicates that there are four levels of measurement of development including the effectiveness, outcomes, outputs and spend activity levels. The different levels of the hierarchy of performance measurement in Figure 2, suggests that these four levels can be described in two groups of measurement of effectiveness (Denizer et al., 2013). The top two levels of the hierarchy, shown in Figure 2, are defined as the effectiveness and outcome levels and both have corresponding evaluation reports taken from the region, country and thematic perspectives of development (World Bank Group IEG, 2016). The second two levels, at the bottom of the hierarchy, are represented by the more fundamental project level measurements of outputs and spend activity evaluated by

the project evaluation reports. As mentioned in Section 3.10, there are challenges of measurement at the top levels, and development metrics can be influenced by other interdependent factors and development activities. These include economics, political factors, parallel private sector initiatives and contributions towards goals from the efforts of competing or collaborating development organisations (Howes et al., 2011).

Examples of such goals and objectives set at the top level of measurement include the Sustainable Development Goals (SDGs), listed in Section 3.8. For example, SDG 1 aims to end poverty in all its forms everywhere, the SDG 2 goal is to end hunger, achieve food security, improve nutrition and promote sustainable agriculture, and SDG 3 and is intended to ensure healthy lives and promote well-being for all (United Nations, 2015, p. 43). The other 17 SDGs have similarly expansive and broad descriptions (Uitto, 2021).

The achievement of development goals and objectives would most certainly be interpreted as effective when such broadly expressed goals are accomplished. The metrics to measure the aspirations of the SDGs have 169 targets and 229 individual indices (Janoušková et al., 2018). However, the metrics as currently expressed, are broadly defined and do not address the stakeholders' questions on how money is spent. Importantly the majority of the metrics do not indicate or measure whether development funds are used effectively at the project level (Atkisson, 2016; Hák et al., 2016).

Many development indicators are measured with a top-down, outcomes, strategy with the assumption that a "rising tide lifts all boats", while others take a more bottom-up approach (Elalfy 2021). Project level performance indicators, for example, include indicators for outputs and spend activity, taken from a more bottom-up approach and measurement of development outcomes can be also evaluated from individual contributions from multiple project outputs (Yuan et al., 2009). The hierarchy of performance measurement was the subject of considerable debate at fora, such as the

Paris Declaration of Aid Effectiveness in 2005 and Busan 2011, which set commitments for improvement of the measurement of DE (Hansen & Tarp, 2000; Wood et al., 2011). Israel (2006) noted Edison's famous quote, credited to the invention of the light bulb, that "Invention is 1% inspiration and 99% perspiration" (Israel, 2006). This describes and contrasts the difference between the design, definition and creation of products and services, with the challenges of the implementation, production process and delivery. In this analogy, as applied to this research, the inspiration may be considered to represent the planning, development and design of stages of procurement and project management processes, and the perspiration is represented by the project management and delivery components in development projects (Denizer et al., 2013; Dixon et al., 2013).

The theoretical framework, shown in Figure 6, was developed for this research to represent the planning and delivery mechanisms of procurement and project management (the inspiration and perspiration) that lead to DE and VFM in international development.

The first two research questions seek to understand and address procurement capacity, maturity and project performance at the project level for ( $n = 1,920$ ) projects in the selected international development organisations included in this research.

Procurement capacity and maturity is primarily measured through surveys or questionnaires, such as PMAs, or by internal KPIs specific to the industry and organisation concerned. The PMA models, such as the OECD-MAPS, used by many of the international development organisations are specifically intended to assess the capacity of government departments, to perform procurement in developing countries (OECD, 2010; World Bank Group, 2016). Furthermore, PMA models, applied in public procurement organisations, do not always include criteria to measure the efficiency of process, technology and systems, client management and risk factors of the procurement process (Raymond, 2008). In the example of the OECD project performance evaluations



used by international development organisations, the evaluation does not include details on the procurement process, and often consist merely of broad statements affirming that procurement procedures have been followed (World Bank Group IEG, 2016). The literature review in Section 3.12, critically evaluates articles and models of leading PMAs used in government and industry, and the OECD- MAPS models used by international development organisations as listed in Table 11, Table 12 and Table 13. The review of existing PMA models in the literature review led to the development of a new Integrated Procurement Maturity Model (IPMM), which is more aligned to the procurement process and criteria of international development organisations. As a result, the IPMM discussed in Section 4.3, is adapted to assess and compare the procurement maturity and capacity for the international development organisations in this research.

To assess and understand project performance, project evaluations in international development, are regularly measured using OECD-DAC development evaluation methodology including the ten organisations reviewed in this research (OECD-DAC, 2012). The research reviews secondary data from a sample of project evaluation results ( $n = 1,920$ ) from ten selected international development organisations between 2015 to 2017. The research provides analysis and comparison of the overall project evaluation ratings results and examines the data to provide further analysis on the four component evaluation criteria that derive the overall project performance ratings results. The literature review in, Section 3.5, discusses and makes comparison with private sector project performance indicators, such as cost, schedule and quality which are standard and proven measures of project performance (Atkinson, 1999; Zidane & Olsson, 2017).

To further examine and compare how expenditure and cost is evaluated in development, measurement of cost variation is examined to measure the difference of actual costs from the cost estimates for projects, as discussed in Section 5.2.10 and shown in Table 47.

The relationship between procurement maturity, project evaluation results, DE and VFM, is assessed to address the third and fourth research question by performing correlation and regression analysis. The correlation and regression tests are conducted between the overall project evaluation, criteria and cost performance results. The following sections are dedicated to addressing the four research questions in light of current knowledge.

## **6.2 PROCUREMENT CAPACITY AND MATURITY ASSESSMENT**

The first research question seeks to address the comparative procurement maturity of leading international development organisations from the perspective of DE and VFM. To explore the research question the commentary in Section 3.12, examines the measurement of procurement capacity and maturity in international development organisations and three private sector benchmark organisations. Common existing Procurement Maturity Assessment (PMA) models, used in the public and private sectors, are discussed in comparison with current procurement capacity and maturity measurement methods used for international development organisations. Procurement maturity assessment is typically carried out using PMAs, and the literature review in Section 3.12, discusses different PMA models used in the private and public sectors. As discussed in Section 3.12, there is no existing effective standardised PMA model for the assessment procurement maturity for the international development organisations themselves. Many of the popular PMA models include criteria that are more applicable to private sector supply chain and warehousing functions. Many PMA criteria contained in the more common models are difficult to adapt and apply to the procurement processes of international development organisations (American Productivity and Quality Center APQC, 2018; Estampe et al., 2013; Schuh et al., 2017; Schweiger, 2014).

The procurement capacity assessment models listed in Table 11 are designed to assess government agency procurement capacity and are not originally intended to measure the internal procurement capacity of the development organisations themselves (OECD, 2010; World Bank Group, 2016). The literature review finds that current project evaluation reports include basic statements on procurement compliance but have limited detail on procurement performance (Raymond, 2008; World Bank Group IEG, 2016). Therefore, for the purposes of the assessment and comparison of procurement capacity, maturity and VFM of the international development organisations, a new Integrated Procurement Maturity Model (IPMM) was developed in this research. The IPMM maturity levels and procurement process diagrams are shown in Figure 4 and Figure 5. The procurement processes of international development organisations in this research, are organised into four primary procurement phases and 12 criteria, as discussed in Section 4.3.5. The four procurement phases and 12 criteria have been identified in this research, as they are included in the majority of the PMA models, and the relevant references for these criteria are discussed in Section 4.3.5, and tabulated in Table 20. The first research question seeks to compare the procurement maturity and performance of international development organisations from the perspective of DE and VFM. In order to address cost and value parameters, VFM factors are introduced to calculate the weighted IPMM results. The VFM factors are developed from feedback from the panel of procurement experts, and used to calculate the weighted IPMM results. The weighted IPMM results reflect the relative influence of VFM on the procurement processes of the international development organisations, as discussed in Section 4.3.6. The original and weighted IPMM results are compared and the rank order for the ten international development organisations shown in Table 25 and Table 27 and the differences are shown in the gap analysis of the IPMM results shown in Table 29.

In addition, the frequency distribution for the original and weighted IPMM results are summarised in Table 26 and Table 28 for further statistical analysis in this research.

The original IPMM results, shown in Table 25 and statistical analysis for original IPMM results in Table 26, indicate that DFID, AFDB and UNICEF are the three top-ranked development organisations for procurement maturity. These organisations represent all three of the organisation types evaluated in this research. Gulrajani (2016) suggested that donors have a preference for multi-lateral organisations, such as the MDBs and UN Agencies, as they are less politicised and fragmented than bilateral development organisations (Brown, 2012). The findings, therefore, do not completely align with some of the assumptions of the current knowledge that claims multi-lateral channels are more efficient than the bilateral channels (Gulrajani, 2016). Despite this apparent preference, Biscaye et al. (2017) argued that there are no clear conclusions on which type of organisation has the most effective development outcomes. The mixed nature of the results suggest that, at least for procurement capacity and maturity requirements, donors should carefully consider how to channel development funds through multilateral or bilateral channels. This decision should to make maximum use of the relative advantages of the organisations and safeguard disbursement of funds (Gulrajani, 2016; Neumayer, 2003). Clements et al. (2008) argued there is inconsistency and positive bias in the approach and standards of project evaluation between different development agencies.

The weighted IPMM results show that DFID and AFDB maintain first and second rank as with the original IPMM results. However, WBG moves from fourth to third rank, and UNICEF moves from third to fourth rank. The changes in rank order for the weighted IPMM results, in this example, suggests that the WBG procurement process makes a comparatively greater contribution to VFM than the UNICEF procurement process.

The weighted IPMM results introduce a new perspective from this research for assessment and prioritisation of VFM, and other factors, in procurement for development projects. VFM is a key principle for development organisations (Barr & Christie, 2015) and the findings offer opportunities to identify best practices for VFM in the procurement process. For example, the results suggest that an emphasis on the early procurement strategy phase combined with the requirements of the regulatory framework phase promotes improved procurement maturity and VFM. More specifically, Table 34 shows that the top six procurement criteria promoting VFM are: procurement strategy, procurement procedures, procurement organisation, contract management, procurement governance and supplier eligibility criteria. This is an important finding that indicates the importance of the fundamentals of the procurement strategy and organisation criteria in combination with the regulatory framework criteria to achieve VFM in procurement. The approach to weighting is adapted from practices used in industries such as Life Cycle Impact Assessment (LCIA) for prioritisation, evaluation and decision analysis, as discussed in Section 4.3.6 (ISO 14044, 2022, Noh & Lee, 2003; Itsubo et al., 2018). Pizzol et al. (2017) noted that while weighting techniques are relevant for decision-making they can be subjective and suggested that weights are developed transparently and robustly to mitigate potential uncertainties and inconsistencies. Nevertheless, the weighting approach is a useful technique to explore priorities and differences in procurement maturity assessment questionnaires such as the IPMM. There are other potential applications for the use of weighting techniques and sensitivity analyses that can be used to explore different aspects of procurement in addition to VFM, such as the influence of corruption and other factors on procurement and project performance.

Another new key finding, as shown in Table 30, in this research indicates that the private sector organisations show greater procurement maturity and contribute more towards

VFM compared to the MDB, UN agencies and bilateral development organisations. The upstream oil and gas industry is selected for comparison as it has similarities in size, geographical distribution and procurement characteristics with international development organisations, as discussed in Section 5.1.8. The results show that Shell and Woodside are ranked first and second, and Chevron is ranked sixth the IPMM assessment results. The findings from the private sector results confirm current knowledge on the benefits of comparison of procurement between public and private sectors and suggest that there are potential advantages in the adoption of private sector procurement practices (Patrucco et al., 2019; Arlbjørn & Freytag, 2012). The advantages of the private sector approach to procurement confirm the research observation that stronger maturity for the procurement strategy and regulatory framework phases and criteria are important to achieve procurement outcomes and VFM (Cooke-Davies, 2002). In addition, the private sector organisation results demonstrate strong maturity across all of the procurement criteria. Many of these criteria include attributes such as innovation, technology, strategic design and training (Arlbjørn & Freytag, 2012; Tadelis, 2012).

The IPMM results in Table 32, shows a peculiar distinction for procurement activity in the MDBs where procurement is separated into operations and corporate procurement groups. Corporate procurement manages procurement activity for internal requirements of the organisation. Operations procurement, on the other hand, provides an oversight function for procurement activity delegated to the borrowing countries and their Implementing Partners (IPs). The results for the two group are compared and the IPMM results in Table 32, indicate that the corporate procurement group results on aggregate, shows a slightly higher procurement capacity and maturity than the operational procurement group. This observation may be explained by the procurement oversight role delegated to borrowing countries and their Implementing Partners (IPs) as applied by

the operational procurement group of the MDBs. In addition, the MDBs have the additional challenges of capacity development and monitoring of the procurement activity and performance of the borrowing countries and their IPs. Gladilina (2017) argued that advantages gained from practical procurement activity leads to improved performance and the observation in this research supports this argument.

The results in Table 33, suggest that procurement processes with good strategic planning, included in the procurement strategy phase, combined with strong governance and compliance of the regulatory framework phase, promotes procurement maturity.

The findings align with studies on procurement maturity that suggest that procurement strategy, regulatory framework and organisational structure are important aspects that are correlated with procurement maturity and performance (Batenburg & Versendaal, 2008).

The contract management phase of procurement, which shadows the implementation stage of the project management activity cycle, is the lowest-ranked phase and is often considered as a poorly managed phase of the procurement process (Gupta, Karayil, & Rajendran, 2008). However, despite the finding that the contract management phase does not rank highly for the IPMM results, contract management is widely recognised as a key contributor to project performance (Doloi, 2013). The literature review suggests that the contract management phase is a critical area and the top performing procurement practices rate highly for this criteria in this research (Cherono & Chepwoy, 2021).

The original IPMM results for the 12 procurement criteria are shown in Table 34 and indicate that the following criteria: procurement organisation, procurement procedures, governance and procurement risk criteria are the highest-ranking procurement criteria.

The procurement technology and systems, contract administration and training criteria are ranked as the lowest procurement criteria from the original IPMM results.

Procurement technology, systems and tools, such as eProcurement, have developed

rapidly in importance in both the private and public sectors over recent years (Kim, Suresh, & Kocabasoglu-Hillmer, 2015). Development organisations have been slower to embrace technology, systems and tools but are rapidly gaining ground (Molino, 2019).

For the weighted IPMM results, the four criteria of procurement strategy criteria, procurement procedures, procurement organisation and contract management criteria are ranked highest of the 12 criteria. The high rank order for these four criteria supports the observation and current knowledge that the procurement strategy and the regulatory framework phases contribute in a substantial way towards VFM (Potage, 2017). The weighted IPMM results also align with current knowledge and confirms that contract management is an essential but overlooked phase of the procurement process that follows and supports the project management process to achieve VFM (Molino, 2019).

In summary, the results in Section 5.1, address the first research question through comparative analysis of the procurement capacity and maturity of the international development organisations, using benchmarks with private sector organisations.

The results also confirm current knowledge regarding procurement best practices in the private sector that can potentially be applied to improve procurement maturity and performance (Aulia & Isvara, 2021; Johnson et al., 2017; Patrucco et al., 2019; Tadelis, 2012). Some of the results may be unexpected given the current knowledge and the literature on the development of procurement reforms and improvements led by WBG, ADB and DFID (World Bank Group, 2016). Several authors refer to leadership in procurement reforms, in particular, by these three organisations and it may be expected that WBG, ADB and DFID would rank more highly (Boakye, 2015; Borson, 2017; Cheronon & Wangangi, 2017; Molino, 2019; Sharp et al., 2019). However, as discussed in earlier chapters, the IPMM assesses the procurement maturity of the international development organisations and not the specific procurement performance in projects.



The research introduces a new Procurement Maturity Assessment (PMA) model; the IPMM, which assesses and benchmarks procurement maturity in development organisations. The research suggests and recommends that procurement maturity is assessed for each individual project and included in the project evaluation reports.

This is the first study to apply a new procurement performance assessment model to international development organisations in comparison with private sector organisations, and hence offers a new perspective to the literature. Secondly, the IPMM incorporates a weighting technique as a sensitivity analysis to assess and prioritise VFM. The findings in Section 5.1, provide important examples of best practices in procurement maturity that can be applied to improve project performance, DE and VFM (de Araújo et al., 2017).

### **6.3 PROJECT EVALUATION PERFORMANCE ASSESSMENT METHODOLOGY**

The second research question investigates how the existing project evaluation assessment methodology measures DE and VFM in leading international development organisations. The different levels of the hierarchy of performance measurement, shown in Figure 2, highlight the different perspectives and measurement levels of DE (Biscaye et al., 2017).

The review of the OECD-DAC (2019) development evaluation methodology reveals the tendency for development organisations to emphasise the top two macro-levels of the hierarchy of performance measurement expressed as: effectiveness and outcomes.

Secondly, the review of the methodology outlines the increasing aspiration, of the international development organisations to include more recent and evolving development goals and objectives, such as sustainability, diversity and inclusion targets (OECD, 2019). Nevertheless, as discussed in Section 3.11, measurement at the bottom two levels is equally important to evaluate performance of development activity and to provide a necessary perspective for Development Effectiveness (DE) at the project level.

The Paris Declaration on Aid Effectiveness, 2005 and the Accra Agenda for Action, 2008 resolved to improve development effectiveness, management for results and set the path towards developing the concept of VFM (Wood et al., 2011). Fora such as these, championed the pursuit of the measurement of DE, with a focus at the higher levels of the hierarchy of performance measurement for development shown in Figure 2. The support for measurement of DE was partly a response to the challenges presented in these meetings in managing and measuring development at the project level for multiple projects across many regions and sectors (Gulrajani, 2014). As discussed in Chapter 3, development aid at the project level, represents approximately 50% of ODA (Ellmers, 2011). This research chiefly concentrates at the project level of development and seeks to understand the contribution of projects towards the measurement of DE and VFM.

The research provides an analysis of a sample of project evaluations ( $n = 1,920$ ) from ten selected international development organisations. The OECD-DAC project evaluation results are expressed by an overall project evaluation rating, which represents the measure of project performance towards development outputs, outcomes and DE.

As discussed in the literature review, Section 3.10, the measurement of project success, DE and VFM is often difficult to identify in project evaluation reports. This difficulty can be partly explained by the absence of specific definitions and measurement of either DE or VFM in the OECD evaluation methodology (Chianca, 2008; OECD, 2019).

The measurement of cost and value at the project evaluation level are defined and expressed as cost-effectiveness and, at first glance, appears to provide answers to the question of “where does the money go?” and give insight into VFM (Easterly & Pfutze, 2008). However, as discussed in Section 5.2.10, the analysis of the ( $n = 1,920$ ) project evaluation reports reveals that there is limited detail and several inconsistencies in the definition of cost-effectiveness in development projects (Khagram et al., 2009). As such,

there is little wonder that there is on-going debate regarding DE and VFM amongst international development organisations (Gulrajani, 2014). Cost variation is measured as the difference between the actual cost and the Cost Estimate (CE) and is referred to as cost performance in this research, as discussed in Section 3.4 (Baccarini & Love, 2014). The project evaluation methodology is comprised of six specific evaluation criteria that include: relevance, coherence, effectiveness, efficiency, impact and sustainability (Chianca, 2008). The coherence and impact evaluation criteria are not measured by the selected international development organisations at the time of data collection, and therefore, data are not available for analysis for this research. Cost-related measurement, described as cost performance, is intended to be incorporated to the efficiency criterion. However, the research finds that cost and expenditure factors are not consistently represented in efficiency criterion by the selected organisations (Clements et al., 2008).

The OECD approach, evaluation and rating methodology has been adapted by many of the international development organisations, and different rating scales ranging from a four-point, five-point and six-point scales are used, as shown in Table 22 and Figure 9. The different scales are normalised in this research, by the calculations shown in Table 23, to enable comparison and analysis of the results of the development organisations. The normalisation of the three different scales introduces minor distortions however, as the normalised scale enables comparison between the different organisations, it is used for the analysis in this research, as discussed in Section 5.2.3. To test the normalisation method, an alternative method of normalisation of the scales is calculated for comparison in this research. The comparison indicates that both methods arrive at very similar results and findings and the comparison is shown in Appendix 3 (IBM, 2020; Leung, 2011). Table 35 shows the overall project evaluation results by the frequency to indicate the comparative pattern of evaluation results of the different rating methods by percentage.

The mean project evaluation results shown in Table 37, serve as the primary measure of project success and DE at the project level for development projects in this research.

The mean results are listed in descending order by organisation in rank order from UNDP, AFDB, JICA, ADB and IADB. The rank order represents the organisations that have the largest number of high performing projects in the sample results ( $n = 1,920$ ).

The five lower-ranked organisations are listed in order from EBRD, WBG, UNICEF, ISDB and DFID include a greater number of lower performing projects in the sample.

The findings show the comparative project evaluation results as evaluated and published by of the international development organisations in their reports and electronic media.

Overall, the analysis of project evaluations has yielded new knowledge about how the international development organisations self-evaluate their project performance. In addition, this is the first time a normalised scale has been used to compare project performance and project evaluations across international development organisations.

New knowledge in the research includes the comparison and analysis of published project evaluation results by organisation, region and sector after normalisation of the different evaluation ratings scales. The different application of the project evaluation process, evaluation criteria and rating scales implies that the international development organisations do not anticipate cross comparison of their project evaluation results.

Clements (2020) observed that most Official Development Assistance (ODA) project evaluations are conducted using the OECD-DAC evaluation framework and this framework is intended to serve as the common standard for development organisations.

Brown (2020) referred to the “aid effectiveness norm”, which contains the principles of ownership, harmonisation, alignment, managing for results and mutual accountability and argued that the use of norms and principles has weakened from 2005 to the present.

Current knowledge, suggests that the project evaluation process does not perform as a

consistent standard. The research recommends improvements to develop more effective and consistent standards and definitions for the project evaluation process and criteria.

Statistical analysis, including the mean results and rank order, are also conducted between the type of organisations, including MDBs, UN Agencies and bilateral development organisations, as shown in Table 38. The grouping of the international development organisations by type, is comprised of a combination of the results of the six MDBs, two UN agencies and two bilateral development agencies. The results indicate relative project performance and show that the UN agencies are ranked first, MDBs ranked second and the bilateral development organisations ranked third. The results show that there is considerable variation in ratings between the individual development organisations and this is new knowledge in the field of project evaluation. The results indicate that larger multilateral agencies appear to perform better for project evaluation results than bilateral organisations, which aligns with current knowledge on comparative development performance (Birdsall et al., 2010; Engen & Prizzon, 2018). The project evaluation results are also reviewed by region and sector to compare the project performance for the different international development organisations by these parameters, and the results are shown in Table 40 and Table 41. The results indicate that EAP, ECA, SAR and AFR are the best-performing regions and LCR, MNA, and Global projects are the lowest performing regions for development projects in this research. The new findings of the results build and align with current knowledge that indicates that regional and cross-country influences, such as politics, governance and economics are possible factors that explain regional differences in project development performance (Han, Khan, & Zhuang, 2014). For example, the research results indicate that there is a relationship, although very weak, between corruption and the project evaluation results based on the Transparency International Corruption Index (CI). Therefore, the slightly

higher correlation results for MNA and AFR implies that project performance is negatively affected by corruption and that international organisations should consider strengthening procurement and project management processes for countries with high CI. The correlation results between project evaluation and the corruption indicators imply that project performance in different countries is influenced by corruption and this brings a new perspective to the analysis of project performance. However, as discussed in Section 5.2.8, the results are also influenced by many other possible factors including inconsistencies in the project evaluation process and further research is recommended. The mean project evaluation results are also reviewed by sector, as shown in Table 44, and the results indicate that the multisector, environment and agriculture sectors are the top performing sectors by the interpretation of the mean results. At the other end of the scale, water, health and government sectors are the lowest-performing sectors, both by comparison of mean result and rank, for the international development projects studied. There are several possible explanations for the differences in performance including variation in project complexity, different government policy, the infrastructure and the the degree of internal project management capability of the development organisations. Multisector projects are often managed centrally, which may potentially increase the quality of management and oversight applied to such projects (Desouza & Evaristo, 2006). In the example of environmental projects, Sachs (2012) remarked that there is a growing emphasis and sense of urgency and recognition on the importance of environmental projects linked to the SDGs. The phenomenon suggests and implies that the additional focus and attention increases the prioritisation and visibility for these types of projects and potentially may improve the comparative performance. Current knowledge on project performance in the construction industry indicates that project performance is correlated with larger projects and there are Critical Success Factors

(CSFs) that promote project success (Andrić et al., 2019; Moradi et al., 2020). This research supports the current knowledge and indicates that there are more overruns in large complex infrastructure projects in the transportation, urban infrastructure, water and sanitation, health and education sectors. In the health sector, Santos et al. (2020) observed that measurement of project success factors have lagged behind other sectors and suggests this may be due to complexity of project scope and specific design factors. In the education sector, Famiyeh et al. (2017) identified similar factors causing schedule delays and cost increases in large, complex school and education construction projects. It is perhaps not surprising that the size and complexity of projects can be related to challenges in cost performance and subsequently project performance, DE and VFM. Examples of good project performance in this research include the multi-sector projects, which often have many stakeholders and participants. Iyer and Banerjee (2016) argued there can be advantages in managerial efficiency for such projects, which often attract attention from senior management who demand best practices from similar projects. The research findings align with current knowledge that there are particular challenges for large complex projects. The findings also imply that there are project management practices that can both mitigate these challenges and improve project performance. The findings also align with current knowledge on project performance by sector, which shows that project management techniques are adapted and adjusted towards different sectors, and thus offers an opportunity for further research (Tayefeh et al., 2020).

As discussed in Section 5.2, the project evaluation results serve as the principal measure of project performance and DE for international development projects (OECD, 2019). The overall project evaluation results in this research are comprised of four component project evaluation criteria, as discussed in Section 3.10. Eight of the ten international

development organisations publish secondary data for the four evaluation criteria and ratings used to calculate the overall project evaluation rating for analysis in this research.

The breakdown and analysis of the project evaluation criteria for the eight international development organisations, is tabulated in Table 46 and shown in Figure 19.

As discussed in Section 5.2.9, authors such as Chianca (2008) and Denizer et al. (2013) argued that the definitions of the criteria are imprecise and can be misleading and the individual criteria are not recognised for their relative contribution towards the overall evaluation results. Zidane and Olsson (2017) noted several articles that link efficiency and effectiveness with project performance, project success, impact and value. The detailed review of the evaluation criteria, in this research, finds that the efficiency criterion incorporates limited but relevant data on project budget, cost and expenditure. The implications of these findings for the research are substantial as the efficiency and effectiveness criteria are shown to make the largest contribution towards the project evaluation results. Secondly, these two criteria, in particular the efficiency criterion, are regularly applied to record data on procurement, cost and value within the project evaluation reports. However, as the cost and expenditure description is often limited in detail in the project evaluation reports, it is difficult to determine and measure cost performance and whether VFM is achieved (Chianca, 2008; Yves-Adou, 2016).

In the private sector, the key criteria of cost, time and quality are considered the core measurement of project success, and the three criteria are termed the “iron triangle” (Atkinson, 1999). Subsequently, additional criteria have been added to the iron triangle, where appropriate, such as health, safety and environment, often as a result of experience and a reaction to incidents and risk mitigation (Golini et al., 2015; Jørgensen et al., 2021). Johnson and Badu (2018) ranked project scope changes, poor planning, inaccurate cost or schedule estimation, and poor selection of contractors as leading causes of cost



overruns (Johnson & Badu, 2018). These factors further highlight the iron triangle's importance for defining and measuring project success and for setting tangible project performance metrics (Johnson & Babu, 2018). This aligns with current knowledge across different industries that the variation between actual costs against the estimation costs is often used as a measure for project cost performance and VFM (Zidane & Olsson, 2017).

In the international development evaluations reviewed, only five of the ten international development organisations provide detailed cost variation data in the published secondary data reviewed in this research. To evaluate the project cost variation, and thereby explore VFM for the five organisations, the frequency of project costs falling into five thresholds above and below the Cost Estimate (CE) are collated and tabulated in Table 47 and shown in Figure 20. There are many reasons why actual costs may not align with the cost estimate in development projects. These include inaccurate cost estimation and budgeting methodology, unexpected site challenges, changes to scope, logistics issues and political constraints amongst others (Ahsan & Gunawan, 2010).

The results in Table 47 indicate that approximately 40% of projects fall within a 10% range above and below the CE. While this may appear as a relatively small percentage, it is not unusual in project management metrics, particularly in the example of construction industry (Flyvbjerg et al., 2018; Odeck, 2004). Secondly, over 50% of project costs, for international development projects have underruns, which are explained as actual costs that are below the CE. Potential explanations for frequent underruns include inaccurate cost estimation, changes to design or scope and fluctuations of the currency, all of which are common characteristics of development projects (Ahsan & Gunawan, 2010).

By comparison and contrast, Flyvbjerg et al. (2018) evaluated project cost performance in private sector infrastructure projects and found that nine out of ten projects reviewed were found to overrun, or exceed the CE (Flyvbjerg et al., 2018).

The findings build on current knowledge and examines cost performance based on methods used in the private sector and applied to international development projects in this research (Belay & Torp, 2017; Love et al., 2018; Park & Papadopoulou, 2012).

The results for cost performance represent new findings for project evaluation results in development and suggests that there is a disconnect between the current measurement methods for the overall project evaluation results, cost performance and VFM metrics.

As discussed in Section 3.10, the project evaluation criteria have broad definitions, do not distinguish between levels of the measurement of project performance, and the rating scales and processes are inconsistent (Clements et al., 2008; Gulrajani, 2014). The OECD project evaluation methodology and evaluation criteria, as described in Section 5.2.9, do not effectively capture VFM at the project activity level. Denizer (2013) argued that the evaluation ratings are vulnerable to bias as the initial evaluators are often the project managers who manage the projects. The project evaluation reports therefore do not provide conclusive answers to the question of “where does the money go?” and presents an opportunity for further research (Easterly & Pfitze, 2008).

As an important finding in this research, the cost variation results shown in Table 47, introduce an alternative perspective in evaluating project cost performance and a better understanding of VFM for development projects (Ahiaga-Dagbui & Smith, 2014).

In addition, the quality, consistency and relevance of the project evaluations need to be consistent for an accurate representation of project performance and the continual improvement DE and VFM at all levels of the hierarchy of performance measurement. Professional project managers and authors advocate that diligent documentation of scope changes and effective monitoring of actual costs against cost estimates leads to improved project performance (Baccarini, 2004; Kerzner, 2017; Klakegg & Lichtenberg, 2016).

To understand project success, Frefer et al. (2018) reviewed definitions and Critical Success Factors (CSFs) for project performance and drew attention to the importance of the iron triangle CSFs of cost, time and quality in addition to others.

The results in Section 5.2 address the second research question and provides an analysis of the project evaluation methodology and the measurement of project performance, DE and VFM. The research goes further to provide analysis of normalised project evaluation results and examine measurement of cost and other CSFs in the evaluation methodology. The results recommend improvement, augmentation and standardisation of the project evaluation and reporting process to provide more consistent, relevant and comparable criteria including meaningful CSFs. Consequentially, such adjustment through inclusion of CSFs to the project evaluation methodology, should help improve the measurement, monitoring and visibility of DE and VFM in international development projects.

The next section addresses the third research question dedicated to determining if the results of the Integrated Procurement Maturity Model (IPMM) correlates with the overall project evaluation results from the ten selected international development organisations.

#### **6.4 RELATIONSHIP BETWEEN PROCUREMENT MATURITY AND PROJECT PERFORMANCE**

The third research question examines whether the Integrated Procurement Maturity Model (IPMM) results, researched in Section 5.1, are related to project evaluation results, researched in Section 5.2, in international development organisations. This is an important next step in the inquiry as it gives insight into the relationship between procurement and project management and examines the theoretical framework shown in Figure 6. Cooke-Davies and Arzymanow (2003) suggest that prolonged pressure to reduce costs and improve efficiencies drive excellence in project management (Rendon, 2015). Cost reduction is often associated as a fundamental and intended objective of the

procurement process (Van Poucke et al., 2016). The theoretical framework proposes that the combination of the procurement and project management processes act together as key contributors, or mechanisms, for efficient and effective project performance (Habibi et al., 2019; Kwak & Ibbs, 2002).

As mentioned in Section 4.3, the IPMM model was developed to measure the procurement capacity and maturity of international development organisations due to the limitations with the procurement maturity data available from project evaluation reports. The IPMM results, therefore, do not measure procurement specific to individual projects, but rather measure the capacity and maturity of the procurement processes and procedures of the organisations assessed through the IPMM questionnaire results.

As discussed in Section 6.3, the research observed that OECD project evaluation criteria, and adaptations used by international development organisations, do not include detail of the procurement process performance in the project evaluation reports (Raymond, 2008).

Project performance measurement is taken from a sample of ( $n = 1,920$ ) project evaluation results from the ten international development organisations and the project evaluation results and procurement maturity results are summarised in Table 50.

The challenge in addressing the third research question, on the relationship between procurement and project performance, arises due to the limited reporting of detailed procurement data in the project evaluation results (de Araújo et al., 2017; OECD 2019).

As discussed in Section 5.2.9, the OECD project evaluation methodology and criteria do not include details or description on procurement maturity and have limited measurement of cost variation and financial performance (Clements et al., 2008; Picciotto, 2020).

Manyathi (2021) argued that effective, efficient and compliant procurement practices contribute substantially towards project performance and success. de Araújo et al. (2017) pointed out that suppliers should be selected by procurement for their capacity and

expertise against common CSFs for effective and efficient implementation of projects. The procurement process incorporates several Critical Success Factors (CSFs) including the measurement of cost, time and scope. In addition, there are several important CSFs and cost indicators common to both the procurement and project management processes. To explore how the individual evaluation criteria contribute to the overall project evaluation results, the results of Pearson's correlation test for the individual evaluation criteria and overall project evaluation ratings are tabulated and shown in Table 51. The results for the breakdown of the evaluation criteria show that the effectiveness and efficiency evaluation criteria have positive significant Pearson's correlation with the overall project evaluation results. By comparison, the relevance and sustainability evaluation criteria have a moderate positive Pearson's correlation with the overall project evaluation results. Therefore, the correlation results for evaluation criteria in this research, indicate that the effectiveness and efficiency criteria make a larger contribution towards the overall project evaluation results than the other two evaluation criteria. The finding is important as it indicates that effectiveness and efficiency criteria contain the majority of the critical indicators that determine the overall project evaluation results and measurement of project success. The results also suggest there are opportunities for additional and more detailed analysis and research of individual development projects. The research findings align with current knowledge suggest that there is a common thread of cost indicators and VFM factors incorporated predominantly within the effectiveness and efficiency evaluation criteria (Chianca, 2008; Picciotto, 2020). The theoretical implication and assumptions of the findings also confirms current knowledge, as noted by authors such as Frefer et al. (2018) and Ahsan (2012), that there is indeed a relationship between the procurement process, project management and VFM. The concept is represented in the theoretical framework in Figure 6.

The implications of findings are that there is a strong case for improved measurement of the iron triangle CSFs, such as cost, schedule and quality (common to procurement and project management) to be incorporated into the project evaluation process and criteria. The findings suggest that more consistent capture and reporting of project evaluations and cost-related data can provide enhanced visibility and identification of success factors that potentially lead towards the achievement of DE and VFM (Doucouliagos & Paldam, 2011; Metzger & Guenther, 2015). This is further discussed in the following sections.

#### **6.5 MEASUREMENT OF DEVELOPMENT EFFECTIVENESS AND VALUE FOR MONEY**

The fourth research question further examines whether procurement maturity and project performance contribute to DE and VFM for project evaluations in international development organisations. This is a major part of the research as it fundamentally determines if the procurement process contributes towards DE and VFM. Findings also reveal areas of opportunity, best practices and specific procurement criteria that can be applied to improve procurement maturity and performance to support DE and VFM. The effectiveness and outcomes levels of the hierarchy of performance measurement shown in Figure 2, include thematic, country and sector evaluation reports for DE conducted by each of the independent evaluation groups of the international development organisations. As discussed in Section 3.10, the thematic, country and sector evaluation reports are essential tools for the overall measurement of the development goals, objectives of DE as measured at the outcome level of development performance. This research, however, is primarily focussed on the output and spend activity levels of the hierarchy of performance measurement and examines development projects and the cumulative contribution of individual project performance towards DE and VFM. From initial analysis, the results do not appear to support current knowledge, discussed in the literature review of this research, that suggests procurement and project

management are closely related and that procurement performance contributes to project success (de Araújo et al., 2017; Schiele, 2007; Walker & Rowlinson, 2007). Authors such as Sundqvist et al. (2014) supports this perspective and argued that effective and efficient project management is essential for delivering successful projects, outputs and outcomes. Chianca (2008) reviewed the OECD-DAC project evaluation criteria and discussed the relative importance and the differences between the evaluation criteria. This research provides additional analysis of the contribution of individual evaluation criteria towards overall project evaluation results in Section 5.3 and shown in Table 51. The research confirms that the project evaluation methodology is the principal method to provide an assessment of development project performance. However, the project evaluation criteria do not include important elements, such as cost, schedule, quality and other factors and they are not applied consistently across the international development organisations (Chianca, 2008; Gulrajani, 2014). Therefore, the implications of the research findings suggest that existing project evaluation criteria do not provide reliable and meaningful measurement of project performance, DE and VFM (Clegg, 2015).

The research reveals a disconnect between the procurement process and the project evaluation results. Nevertheless, authors such as de Araújo et al. (2017) argued that the procurement and project management processes contribute in combination towards project performance as key mechanisms of delivery and implementation of projects. The concept is reflected and shown in the theoretical framework of the research in Figure 6.

Aulia and Isvara (2022) argued that procurement performance is important for effective and efficient supply chains and successful business outcomes. Other authors such as Zidane and Olsson (2017) claims that cost, schedule and quality are important procurement criteria that contribute towards to effective project performance and Value for Money (VFM). However, the findings of this research indicate that procurement, cost

and value details are not specifically measured in the OECD project evaluation process and criteria. In addition, the OECD project evaluations do not include commonly used Critical Success Factors (CSFs) noted by different authors for the evaluation of projects (Atkinson, 1999; Ika et al., 2012; Mir & Pinnington, 2014; Zidane & Olsson, 2017).

Research question four seeks to understand if procurement and project management performance contribute towards DE and VFM. The IPMM questionnaire has been developed to measure procurement performance and VFM in development organisations.

The IPMM has been distributed for the first time in this research to ten international development and three private sector organisations. The IPMM provides a preliminary independent set of results for the comparison of procurement maturity of international development organisations. The findings support current knowledge suggesting that procurement contributes towards achieving effective project outcomes and Value for Money (VFM) in development projects (Ahsan & Paul, 2018; Kakwezi & Nyeko, 2019).

The OECD project evaluation criteria are considered as the established measure of DE for development projects. However, several authors argue that there is inconsistency and bias in the project evaluation criteria and suggest opportunities for improvement of the measurement of DE (Chianca, 2008; Clements et al., 2008; Kilby & Michaelowa, 2019).

This research takes a new approach to understanding the evaluation of DE and provides analysis of 1,920 project evaluations by organisation, country and sector. In addition, the research examines how individual project evaluation criteria measure both DE and VFM.

The research confirms current knowledge on the measurement of DE measured at the project level, and the findings also suggest that there are several opportunities for improvement of the definitions of the project evaluation criteria. For example, the research recommends that the definitions of the relevance and sustainability criteria



should be revised as they add limited objective data and insight to the overall evaluation results. The effectiveness and efficiency criteria make a larger contribution to the overall project evaluation ratings than the other two criteria but contain limited information and detail on cost, schedule and quality as indicators of project performance. The secondary data reveals that only five of the ten organisations provide published secondary cost data for analysis in this research. Secondly, the limited quality and consistency of cost data from all five organisations implies that cost performance is not a priority in development project evaluations. From the 1,084 projects evaluated for cost performance, there were 657 cost overruns, 324 underruns and 103 projects that exactly met the CE. Possible causes of overruns and underruns are shown in Table 49 and include procurement and schedule delays, government approval issues, inaccurate CEs, and scope changes.

As discussed in Section 6.1, there is only limited detail of procurement performance in the OECD project evaluation reporting process and definitions and evaluation criteria do not measure cost performance and VFM (Clements et al., 2008; OECD-DAC, 2019).

This research introduces a new approach to understanding cost performance, as defined by Love et al. (2018), calculated from the variation of actual project costs from the Cost Estimates (CE) and the results are described in Section 5.2.10. The evaluation of cost performance provides a valuable perspective for understanding project performance and it is clear that this will be greatly improved with standardised project cost data collection.

The research proposes opportunities to improve the project evaluation methodology towards a more standardised and consistent approach to measuring procurement and project management performance. The new findings in this research suggest that enhanced project evaluations will strengthen the measurement and performance of these two critical project delivery processes and improve project outcomes, DE and VFM.

## 6.6 SUMMARY

Chapter 6 provides a discussion of the results and the findings in answer to the research questions with respect to the status of current knowledge. The IPMM questionnaire, developed and applied in this research, shows the comparative results in procurement maturity of international development organisations with private sector benchmarks. The research also provides analysis of the sample of ( $n = 1,920$ ) project evaluation results and describes and discusses the comparative project performance by organisation, country and sector between the ten selected international development organisations. Current knowledge suggests that there is a relationship between the procurement and project management processes in the selected international development organisations. However, as procurement is not specifically measured within the project evaluation criteria, the research recommends that procurement data and Critical Success Factors (CSFs), such as cost, schedule and quality are included in the project evaluation criteria. As discussed, the definitions of the project evaluation criteria are broad and often interpreted inconsistently, and the research suggests that the correlation results will be more meaningful and reliable if the project evaluation criteria are standardised. Secondly, the international development organisations use different even and odd number r-type scales which introduces numerical challenges with the comparison of ratings and results. DFID is the only organisation using an odd number five-point scale, which introduces an ordinal mid-point. The mid-point also creates a dilemma for the both definitions and the numerical mean, as the interpretation of the definition successful tends to fall at the mid-point, whereas for even number rating scales it falls above the mid-point. In the case of DFID, as discussed in Section 4.5.1 and Section 5.2.4, the distortion has the effect of reducing the ranking of DFID from first to tenth rank after normalising the results. The results arguably may not represent its accurate comparative

performance and offers the opportunity for further research on DFID results. Cummins and Gullone (2000) make a case to support a ten-point scale, which is in line with the normalised scale used in this research (Cummins & Gullone, 2000). The normalised scale used in this research introduces a new approach to the comparison of project evaluation results. The research recommends that there are opportunities to improve the assessment of the procurement performance and develop the project evaluation methodology to enable more accurate measurement and analysis of DE and VFM in international development organisations.

The definition of effectiveness, as discussed in Section 3.10, is “the extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups”. The definition of efficiency is defined as “the extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way” (Chianca, 2008, p. 43; OECD, 2019 pp. 5-12). Gulrajani (2014) argued that the evaluation process has an emphasis on cost and time in the efficiency criterion.

This research confirms that the Task Team Leaders (TTLs) utilise the efficiency criterion to include selected and limited narrative on costs in project evaluation reports. However, the research does not find substantial evidence of the data required for the measurement of cost and schedule Critical Success Factors (CSFs) in the project evaluation reports.

The absence of cost indicators included within the project evaluation criteria also implies that measurement of VFM is similarly inconsistent and difficult to identify and interpret.

Therefore, an important and far-reaching implication of the new findings in this research is that the project evaluation methodology, currently used by international development organisations, does not provide reliable and accurate measurement of either DE or VFM.

Further to these implications, given the scale of development aid funds lost to corruption, waste and inefficiency, is the important question of whether development organisations are able to accurately measure funds lost to corruption and waste (Kolstad et al., 2008; Ravelo, 2012). Following on from this, if the organisations cannot identify how and where funds are lost during project implementation, it is highly unlikely they are able to prevent such losses in future projects (Heggstad & Frøystad, 2011; Matthew et al., 2013).

Now that the four research questions have been addressed in terms of a reflection on how the findings of this research relate to current knowledge, Chapter 7 will offer concluding thoughts about the research. Chapter 7 summarises key findings, discusses the limitations of the research, elaborates upon the practical and theoretical implications of the research and explores ideas and suggestions for further research.

## **7 CONCLUSION**

Chapter 6 was dedicated to a discussion of results and demonstrated several new findings from the research. Chapter 7 provides detailed conclusion to the research. It summarises the research conducted for the thesis, discusses the limitations, outlines the implications from both practical and theoretical perspectives and makes proposals for future research.

### **7.1 INTRODUCTION**

Key findings of the research are derived from the inconsistent definition and measurement of project success across the international development organisations. The implications of these findings extend to questions on the approach to measurement of Development Effectiveness (DE) and Value for Money (VFM) in projects. Decisions to increase or extend development funding through projects should be based on successful performance and hence lack of confidence concerning project performance can have consequences. Some of these consequences include the risk that “good money will be spent after bad” and reinforce weak project management practices, poor ethical behaviour and potentially foster corruption (Kono & Montinola, 2013; Maren, 2009).

This research examines the measurement of procurement and project management performance and their relative contribution towards DE and VFM in international development. To understand whether development projects are effective, and if they provide value, it is necessary to carefully review how procurement and project performance is currently evaluated and measured. The first two research questions examine the procurement and project evaluation performance of ten selected international development organisations. The third research question provides analysis of the relationship between the procurement process and project management process. The fourth research question examines the effectiveness of development processes and explores the contribution of procurement towards project performance, DE and VFM.

The OECD based project evaluation methodology and criteria, used by the ten selected international development organisations, are intended to measure the performance of development activity at the project level. The analysis in the research, as shown in Figure 14, indicates that there is a substantial difference in the interpretation of project evaluation results between the different international development organisations. The research provides analysis of the aggregate project evaluation results and includes a break down and analysis of the four component evaluation criteria. The new research findings show that, while the overall evaluation results present ratings for the performance of development projects, it is more difficult to conclude whether these results measure or assess Development Effectiveness (DE) and Value for Money (VFM). This is a critical finding in the research, applicable to all the research questions, and implies that current project evaluation criteria do not comprehensively measure DE and VFM at the project level. Possible explanations include inconsistency of the project evaluation methodology, different interpretation of the evaluation criteria and absence of Critical Success Factors (CSFs), cost, schedule and quality within the evaluation criteria. As discussed in Section 3.12, Kelvin (1883, as cited in Kaplan, 2008) stated the belief that, “if you cannot measure it, you cannot improve it”, and remarked how this statement equally applied to management processes (Kaplan, 2009, p. 3; Kelvin, 1883). This research provides analysis of the measurement and evaluation of the procurement and project management processes in international development projects. The findings in this research reveal inconsistencies in the evaluation criteria definitions and implies that they should be addressed to better measure and improve project performance, DE and VFM. The research results indicate a very weak statistical correlation and relationship between the procurement and project management processes of international development organisations. As discussed in Section 6.4, this is an important and surprising finding as

these two processes are considered to act as the key mechanisms of project delivery, as reflected in the theoretical framework shown in Figure 6. However, the weak correlation result does not necessarily mean that there is a weak relationship between the procurement and project management processes themselves. To explore this further, detailed analysis is performed on the overall project evaluation results, the evaluation criteria, procurement maturity and their contribution to DE and VFM. The findings suggest that common links can be identified between the processes and that improvement in the measurement of procurement and project management performance can provide greater transparency and potentially lead to more effective DE and VFM.

The case was made in Section 3.3 that half of ODA is estimated to be delivered in the form of goods, services, and equipment (Ellmers, 2011). This provides further credence to the author's argument that the procurement process is important for implementing and delivering international development projects. Suppliers and contractors selected through the procurement process, design, implement and manage a large proportion of private sector and development project activity. Watt et al. (2010) argued that the effective selection and evaluation of suppliers and contractors in the procurement process is critical for successful project performance and outcomes (Watt, Kayis, & Willey, 2010). Furthermore, the effective monitoring and evaluation of procurement and spend activity through suppliers and contractors is important to ensure prudent disbursement of funds and to promote both DE and VFM on behalf of the contributing donors and taxpayers. The research discusses Critical Success Factors (CSFs) in the private sector for the achievement of project success, such as the "iron triangle" and suggests that there are best practices for improvement of the measurement of project performance, DE and VFM (Ika & Donnelly, 2017; Iriarte & Bayona, 2020; Jha & Iyer, 2007). The following section provides concluding remarks for the research questions and findings.

## 7.2 CONCLUDING REMARKS

The research took a mixed-method and abductive approach to analysis of procurement maturity and project performance in international development organisations.

For procurement it was concluded, after analysis of the literature and a review of existing PMA models, that there are several disparate procurement assessment models and very few suitable PMA model for assessing international development organisations. The different PMA models vary in structure and approach, and some are tailored towards the private sector and specific industries, while others are applied across a wide range of public and private sector organisations. The definition and classification of maturity in the different PMA models varies according to the target organisation, purpose and focus of the PMA questionnaire. For example, PMAs used in private sector organisations, classify criteria such as category management, technology and eProcurement systems at the top level of the procurement maturity model, shown in Figure 4. PMAs used in the public sector organisations have greater emphasis on procurement regulation and control with greater focus on governance and compliance criteria (Manyathi et al., 2021).

Phase one of this research reviews current knowledge on existing PMA models to develop a new IPMM relevant and applicable for measuring and comparing procurement capacity and maturity in international development organisations, as shown in Figure 4.

The first research question, outlined in Section 3.14, seeks to measure procurement maturity and VFM for ten international development organisations and three private sector benchmarks. For this research, a new IPMM was designed, standardised to assess the specific relevant phases and criteria across the procurement processes of both international development and private sector organisations (Brandmeier & Rupp, 2010). The inclusion of factors and weights for the IPMM is adapted from weighting techniques used for procurement in the private sector, such as the resources industry (Banda, 2019).



The results are important as they bring a new comparative assessment of international development organisations with the new Integrated Procurement Maturity Model (IPMM). The IPMM results indicate differences in procurement maturity between individual development organisations. Aulia and Isvara, (2021) argued that procurement maturity is correlated with project performance and hence the IPMM results can be assumed to measure the procurement capacity and performance of the organisations (Batenburg & Versendal, 2008). de Araújo et al. (2017) observed that excellence in the procurement process is required in order to achieve successful outcomes in any project. The connection between procurement maturity, procurement performance and project management implies that procurement performance contributes to project performance and subsequently towards DE and VFM (Frefer et al., 2018; Lindstrom, 2014).

Project performance and success is evaluated under the OECD-DAC project evaluation criteria explicitly to measure Development Effectiveness (DE) in development projects. However, the research findings for procurement performance and IPMM questionnaire results do not show a significant relationship with project performance and therefore, do not support the expectations and current knowledge derived from authors such as Aulia and Isvara, (2021) and de Araújo et al. (2017). The review of the OECD-DAC project evaluation process and criteria reveals that the criteria do not include evaluation of procurement performance in projects (OECD 2019). The findings of the research suggest that the disconnect between procurement maturity and project performance may be due to inconsistent measurement of Critical Success Factors (CSFs) such as cost, schedule and quality. In support of this finding, Clements (2020) and Clements et al. (2008) posited that project performance should be improved if project managers adopted a more reliable and consistent approach to measuring cost performance and efficiency.

Comparative results from the private sector, reveal best practices and opportunities for improvement of the procurement process for international development organisations.

The research also provides an analysis of the relative maturity of the specific procurement phase and criteria, as shown in Table 33 and shown in Table 34, for each organisation. The results indicate an important finding that the regulatory framework and procurement organisation phases contain the strongest performing procurement criteria for the procurement processes of the international development organisations.

The weighted IPMM results indicate that the procurement strategy and regulatory framework phases have a comparatively greater influence on VFM within the procurement process for the international development and benchmark organisations.

The identification of specific phases and criteria that perform well in both the development industry and private sector and reveals best practices and opportunities to improve procurement maturity and potentially to promote and support DE and VFM.

The IPMM results and findings from the literature review, suggest that the combination of strong maturity in the procurement strategy and regulatory framework phases and criteria, improves the overall procurement process, increases cost performance and contributes to VFM (Cooke-Davies, 2002; Manyathi et al., 2021; Tadelis, 2012).

Frefer et al. (2018) argued that the procurement process involves project-critical factors such as budget development, cost estimation, project design, specifications and contract management. These factors are key determinants for successful project outcomes and are related to the CSFs of the iron triangle including cost, schedule and quality (de Araújo et al., 2017; Jha & Iyer, 2007).

The research findings in Phase 1 supports current knowledge on the importance of the procurement process for project performance, DE and VFM. Major conclusions are that international development organisations should seek to enhance procurement maturity

and performance and adopt best practices from both public sector and private sector organisations. Secondly, they should continually improve the procurement process to maximise the phases and criteria that best achieve procurement effectiveness, efficiency and VFM. Thirdly, organisations should measure Critical Success Factors (CSFs), such as cost, schedule and quality in project evaluations and ensure that such factors are transparently and consistently measured and recorded.

In Phase 2 of the research, project management evaluation results were taken from the sample secondary data for ( $n = 1,920$ ) project evaluations of the ten international selected development organisations. The comparative analysis was conducted between the international development organisations for the project evaluation results overall, results by region and results by sector. As outlined in Section 5.2, the analysis seeks to address the second research question, which is to understand how the prevailing and established project evaluation assessment methodology measures DE and VFM.

The project evaluation results are further broken down to explore the component evaluation criteria and the relative contribution overall project evaluation results and DE.

In addition, an analysis of the project evaluation criteria was conducted to drill down further to examine cost variation and to further explore VFM and how money is spent.

The review and analysis of project evaluation results reveal several differences and inconsistencies in the application of the OECD-DAC project evaluation methodology between the international development organisations' results. The literature review discusses different measures of project success in industry and the private sector and seeks to identify best practices and benchmarks applicable to international development. The normalised project evaluation results, described in Section 4.5.1, provide a summary of project evaluation results for the selected international development organisations.

The best performing regions for all the international development organisations are EAP, ECA and SAR and the lowest ranked regional category is the Global category. The best performing sector is the multi-sector category followed by the Environment and Agriculture sectors and the lowest ranked sector is the Governance sector and these results provide a new and important perspective of project performance by region and sector. There are several possible explanations for the results and these include regional and sectoral differences in government policy, infrastructure and organisation structure, and these imply that there may be important variation in project management practices. There are differences between the organisations themselves and Biscaye et al. (2017) observed that there are variations in bilateral and multilateral development effectiveness across countries, regions and sectors. Serrat (2017) argued that there are numerous possible reasons for differences in development effectiveness by country and region, including the negative effects of corruption and inefficiency on development projects. The research found that there was stronger correlation between the project evaluation results in the Middle East and African countries, which have high corruption indicators. Similarly, there is variation in development effectiveness between the different sectors and the research findings suggested that the variation was frequently associated with the size and complexity of projects regardless of the sector. The research also found that increased focus and attention on the larger projects from senior management and sector experts improved project performance. Further research is recommended to better understand the different project evaluation results between countries, regions and sectors. Additional analysis was performed on the evaluation criteria and breakdown of actual costs from Cost Estimate (CE). The objective was to review how the individual project evaluation criteria contribute to the overall project evaluation results and to examine the measurement of cost and VFM for the ten international development organisations.

The research reveals an important new finding and a potential disconnect between the cost performance indicators and the project evaluations criteria. Secondly, the research shows that actual cost data and cost estimation data for projects are inconsistent or not recorded in evaluation reports for many international development organisations. Effective monitoring of costs are fundamental to understanding project management performance and how the money is spent (Kerzner, 2002; Pohl & Mihaljek, 1992). Project management best practices advocate that the diligent documentation of scope changes, and effective monitoring of actual project costs against CE leads to improvements in project performance and more effective project outcomes (Baccarini, 2004; Kerzner, 2017; Klakegg & Lichtenberg, 2016). The research reviews examples and best practices from the private sector, as described in the literature research in Chapter 3, and examines the variation of actual cost from the Cost Estimate (CE), as applied to five of the international development organisations.

Schiele (2007) makes the case that successful project performance requires the combined effective implementation of both the procurement and project management processes, as illustrated in the theoretical framework, in Section 3.13 and Figure 6.

The research findings support current knowledge that the OECD project evaluation definitions do not contain detailed descriptions of procurement activity or include consistent measurement CSFs, such as cost, schedule and quality (Frefer et al., 2018; Khagram et al., 2009). Secondly, as discussed in Section 3.10, the OECD-DAC (2019) definitions of project evaluation criteria are subjective and lack adequate measurable description and detail within the individual evaluation criteria (Chianca, 2008). As a result, the definition of project performance is inconsistent between individual projects and the interpretation of overall development outcomes and success can be unclear and inconclusive (Ika & Donnelly, 2017; Khang & Moe, 2008; Picciotto, 2019).

According to the findings, this research concludes that the private sector offers examples of best practices in procurement process that can improve the measurement of project performance and potentially lead to more effective project outcomes (Ika et al., 2012).

The implication of the findings suggests that Critical Success Factors (CSFs), such as cost, schedule and quality should be incorporated into the evaluation criteria and used to improve the measurement of project performance, DE and VFM (Frefer et al., 2018).

This research addresses the first two research questions on how procurement and project management performance is measured in international development organisations regarding DE and VFM. The findings conclude that the measurement of procurement maturity provides visibility on procurement performance and reveals best practices.

Secondly, procurement maturity assessment provides the comparison of the individual procurement criteria and their contribution towards project performance and VFM.

Thirdly, the findings suggest that cost performance serves as an important connector between the procurement and project management processes and therefore cost performance should be carefully measured within the project evaluation criteria.

The research concludes that international development organisations should adopt best practices and CSFs that include cost, schedule and quality in the project evaluation process and evaluation criteria. As mentioned, organisations should improve the measurement and monitoring of cost performance indicators from the early phases of the procurement process through to the project management and project evaluation process.

The research further addresses the third research question and explores the relationship between the procurement and the project management processes. The research concludes that the two mechanisms of project delivery, procurement and project management, shown in Figure 6, contribute to the successful implementation of development projects.

The research suggests that the successful implementation and transparent evaluation of development projects leads to the achievement of DE and VFM in development projects.

The fourth research question is much more difficult to examine. While the current project evaluation methodology used by international development organisations includes the assessment of development project performance, the findings do not show evidence that they comprehensively measure either DE or VFM. Nevertheless, as discussed earlier, the implications of the research are that improvement of project evaluation methodology and criteria should lead to achievement of both DE and VFM.

The fourth research question is therefore partly answered by the first three questions and the finding that the measurement of DE and VFM requires effective procurement and project management processes combined with reliable evaluation of performance.

### **7.3 LIMITATIONS OF THE RESEARCH**

The research questions and the comparison of qualitative and quantitative data in this research includes the use of the IPMM questionnaire and collection of secondary data.

There are several limitations identified in this research which emerge due to availability, consistency and comparison of the primary and secondary data selected for analysis.

The first limitation of the research concerns the first research question and the method of collection of primary data by electronic distribution of the IPMM survey instrument to the international development organisations. The head of procurement for each organisation was requested to forward the IPMM survey instrument to respondents or provide electronic contact details for direct distribution of the questionnaire. All the organisations responded, however, considerable effort was required to provide assurance of the research's confidentiality. The peculiar nature of the privileges, immunity and confidentiality of data, held by international development organisations, introduces challenges with obtaining such primary data. This may explain some variation in the

number of responses from different organisations. Secondly, even though the questions were randomised, there remains the possibility of bias in common with this type questionnaire and surveys of this kind used in similar research (Denizer et al., 2013).

The second limitation related to the second research question population and sample size due to the requirement that the organisations use the OECD-DAC project evaluation methodology in this research. This requirement for OECD-DAC project evaluation methodology is required so that the data and results can be compared with each other. However, the requirement reduces the sample size to those organisations that follow this methodology from the total population of international development organisations.

As discussed in Section 3.2, it is estimated that just over half of ODA is disbursed and distributed in the form of goods and services through the procurement processes toward the development projects (Ellmers, 2011). The scope of ODA in this research is therefore limited to development data taken from evaluated projects. Consequently, the research does not cover some other forms of aid disbursement, such as loans and grants, and concentrates on data analysis at the project level of international development.

Project evaluation results provide the principal measurement of procurement, project performance and expenditure and consequently goes some way to addressing Easterly and Pfutze's (2008) question of "where does the money go?" (Easterly & Pfutze, 2008).

The third limitation is the variable quality and inconsistency of the secondary project evaluation data and evaluation methodology used in this research, which introduces challenges for comparability and accuracy (Clements, 2020; Gulrajani, 2014). The OECD evaluation methodology does not mandate a standardised evaluation process or recommend common rating scales to assess project performance, DE and VFM in international development. The use of three different project evaluation rating scales used by the international development organisations make it necessary to normalise the



ratings to ensure that the evaluation results are comparable (Bayiley & Teklu, 2016). Finally, the broad definitions of the evaluation criteria evaluated in this research, introduces inconsistent interpretation including the potential for bias of the evaluation results for different evaluators (Denizer et al., 2013). The broad definitions often cause inconsistent results between different evaluators both within and between organisations. These inconsistencies make it more difficult to compare and benchmark project performance across the international development organisations (Clements et al., 2008). Despite the limitations of the research, it is possible to address the first and second research questions with the available primary and secondary data from the IPMM and the project evaluation data. The correlation and regression analysis between procurement and project evaluation results provide contributions towards the answer to the third research question. The analysis examines the relationship between the two critical process of procurement and project management of the theoretical framework. The fourth research question is more difficult to answer, given some of the research limitations mentioned. The research suggests that the procurement and project evaluation criteria contribute towards DE and VFM. However, the inconsistent definitions and assessment of the project evaluation criteria, make it difficult to conclude that there is currently a reliable measure of DE and VFM. The research suggests that improvement of these definitions and methods should lead to answers to the fourth research question on how the procurement and project management processes contribute to DE and VFM. In summary, the limitations of this research, include limitations of primary data collection shared by similar questionnaires and surveys and these can be largely mitigated through the management and distribution of the IPMM. Secondly, limitations of secondary data include variable interpretation and definition of the evaluation criteria, inconsistent application of evaluation methods and the use of different rating scales.

#### 7.4 CONTRIBUTIONS TO THEORY

The literature review describes the definitions and measurement of DE and VFM in international development and discusses the different levels of complexity of the measurement of both concepts. Whether development is effective, or Value for Money (VFM) has been achieved, is an epistemological challenge and often depends on who asks the question. Research findings confirm that the existing project evaluation methodology assesses the performance of development projects and evaluation reports incorporate some indicators of cost. However, the new findings from the research suggest that the project evaluation methodology does not fully extend to the measurement of the effectiveness or value achieved for funds disbursed by development projects. Accordingly, it cannot be concluded by this research and current knowledge that the project evaluation methodology and reports reliably measure either DE or VFM.

To address the the concept of Development Effectiveness (DE), the research builds on current knowledge that there is a distinction of the four levels of the hierarchy of performance measurement. The four levels include development effectiveness, outputs and spend activity as shown in Figure 2 (Biscaye et al., 2015). The first group is comprised of the top two levels of the hierarchy and includes the effectiveness and outcomes levels that measure the overall development goals, such as the Sustainable Development Goals (SDGs), thematic and country outcomes. This first group is often considered from the macro perspective of the donor governments, development agencies and the recipient government departments. The second group is comprised of the bottom two levels of the hierarchy, including the outputs and spend activity levels, which are more aligned to project outputs and how project funds are disbursed in practice. The second group is considered more from the micro perspective of donor development finances, participating stakeholders and recipient local government entities and

communities (Easterly & Pfutze, 2008). Ellmers (2011) noted that approximately 50% of development funds are spent through the procurement and project management process (Ellmers, 2011; OECD, 2017). Procurement and project performance is measured largely at the bottom two levels of the hierarchy of performance measurement. This research makes the case that effective measurement of both procurement and project performance is important to monitor and manage the disbursement of development project funds.

The research builds on current knowledge and concludes that improvement of the micro-level procurement and project management processes has a positive impact on project performance and potentially towards DE and VFM (Lockamy III & McCormack, 2004).

The concepts of DE emerged at the Millennium summit in Monterrey in 2000, and further developed at the March 2005 Paris declaration on aid effectiveness (Bourguignon & Sundberg, 2007; Eyben, 2013). Despite the best efforts of the participants at the Paris Declaration on Aid Effectiveness 2005, and the High-Level Forum for Aid Effectiveness in Busan 2011, no consensus has emerged on the meaning and definition of DE (Eyben, 2013; Gulrajani, 2014). The OECD project evaluation methodology is reviewed and discussed in Section 3.10, and the analysis and results in Section 5.2, confirm current knowledge that the definitions of the evaluation criteria are inconsistent and lack adequate description and detail (Clements, 2020; OECD-DAC, 2019). In addition to the issues of defining DE, there are several different definitions amongst development organisations for VFM, which include fitness for purpose, quality, total lifetime costs, risk and sustainability factors (Asian Development Bank, 2019; CIPS Australia, 2020).

The contributions to theory from this research are substantial. The new IPMM questionnaire developed in this research shows that the procurement process as a whole and the individual evaluation criteria contribute towards VFM in different ways.

For example, the findings suggest that the combination of procurement strategy, planning and regulatory framework criteria promote and increase procurement maturity and VFM.

The implications of these findings are that the criteria associated with early strategic planning of projects, and the criteria associated with the regulation and governance of active projects, are simultaneously required for procurement performance and VFM.

The research confirms current knowledge regarding inconsistency of the project evaluation criteria after the three different rating scales were normalised to enable more effective comparison of project evaluation results between the different organisations.

The research provides comparative analysis of the project evaluation results from different international development organisations by region and sector. For example, the new findings reveal relationships, albeit weak ones, between regional project evaluation results and corruption and recommends that further studies are needed to understand the country, regional and sectoral differences that affect project performance and success.

In addressing the third research question, the research confirms current knowledge that the relationships between procurement and project management are not well established. A key contribution of this research is to identify that potential links between these two processes are contained in Critical Success Factors (CSFs) such as cost, schedule and quality within the efficiency criterion. Additional conclusions of the research are that the connection between procurement and project management can be strengthened by improvements in the procurement, project management and project evaluation processes.

The new research findings confirm that, while the project evaluation methodology and criteria are intended to assess the performance of development projects, the ratings and criteria do not provide a reliable or meaningful measure of DE and VFM. Meaningful measurement and understanding of DE and VFM in development projects remains

elusive and continues to be a key topic for fora such as the 2005 Paris Declaration of Aid Effectiveness, the Accra Agenda for Action 2008, and High-Level Forum for Aid Effectiveness in Busan, 2011, and future events and studies. This research provides an original perspective and contribution to theory on the measurement of project performance, DE and VFM in the evaluation of development projects.

The research provides analysis at the project level of the hierarchy of performance measurement and concludes that the project evaluation process is critical for understanding both micro and macro levels of DE and VFM for development projects.

### **7.5 CONTRIBUTIONS TO PRACTICE**

The research includes development of a new IPMM for the assessment of procurement capacity and maturity that is relevant for the measurement of VFM and the comparison of international development organisations and comparable organisations. The literature review in Section 3.12 discusses 12 PMAs used by international development organisations to assess developing countries' procurement capacity. The literature review referenced 27 articles on the development of PMA, and examined nine common existing PMA models. Analysis of the PMA models outlines several differences between the various models, particularly between private and public sector models.

The IPMM was developed, in this research, to introduce a new PMA model, applicable to the assessment of procurement maturity for the specific procurement processes and criteria used in international development organisations as shown in Figure 5.

The IPMM questionnaire shown in Appendix 1 is closely aligned to the policy, process and procedures of procurement in international development organisations and includes flexible sensitivity analysis for VFM, which is adaptable for analysis of other factors.

The research contributes to practice with the development of a new PMA model and introduction of a flexible sensitivity, or weighting concept, which assigns different

values, or weights, to the individual procurement criteria that make up the model. The IPMM is a useful and adaptable tool for the assessment of procurement in development and for further research and benchmarking of international development organisations.

The project evaluation criteria have been in use since 1991. The OECD network on development evaluation reviewed the strengths and weaknesses of the evaluation criteria in 2019 and proposed several changes and updates (Chianca, 2008; OECD-DAC, 2019).

The research substantially contributes to practice and suggests that improved implementation and the standardisation of the evaluation criteria will provide greater accuracy and consistency of evaluation reports to measure DE and VFM. Furthermore, the research makes the case that best practices of project management and evaluation from industry and the private sector can potentially improve the measurement and performance and delivery of international development projects (Shrnur et al., 1997).

The focus on DE and VFM, in this research, provides a contribution to current knowledge regarding the role of project evaluation in measuring project performance in terms of effective implementation, delivery and the optimal use of development funds.

The macro-level perspectives of development results, such as effectiveness and outcomes, shown in Figure 2, are important for the assessment of thematic and country perspectives of project performance and DE. The project level micro-level perspectives of development results, such as outputs and spend activity, are equally important perspectives and provide detail on how development funds are disbursed and utilised.

The World Bank Group estimates that US\$20 billion to \$40 billion of ODA is lost from public budgets in developing countries to high level corruption each year (Baker, 2005; OECD, 2014). Ban-Ki-moon, the former Secretary General of the United Nations himself estimated that some 30% of developed aid was lost to corruption (OECD, 2014; Ravelo, 2012). Despite the difficulty in measuring the exact percentage lost to corruption

and inefficiency, this represents approximately \$48.36 billion of the total \$161.2 billion of ODA disbursed in 2020 (OECD, 2021). The research findings indicated that the correlation between the project evaluation results and the Corruption Index (CI) was stronger for countries with high CI indices than for countries with low CI indices.

Important contributions to practice of the research include the implication that effective measurement and implementation of the procurement and project processes improve project performance and potentially reduce the impact of funds lost to corruption and waste. Secondly, procurement measurement and project performance evaluation can provide reliable and meaningful indicators for DE and VFM at the project output level. The measurement of DE and VFM provides an essential source of truth, which equips donors and development organisations with tangible evidence on procurement performance and on how development funds are disbursed. Evidence of DE and VFM can be used in practice to evaluate the performance of current development projects and for the strategic planning and design of future development goals and objectives.

## **7.6 RECOMMENDATION FOR FURTHER RESEARCH**

Further research related to all of the research questions may benefit from access to additional data. For example, non-public data, normally considered confidential by the international development organisation, can be obtained by requests for undisclosed data. The challenge in obtaining confidential data that is not normally disclosed, is due to the internal security classification of data and the international development organisations' privileges and immunities (Miller, 2009). Direct communication in this way potentially increases the time for research, and there is no guarantee that requests for disclosure of confidential data would be successful. However, the approach will potentially obtain additional relevant data and documentation to enable further research analysis and to expand the research findings. Future research should concentrate more time to review the

details of specific documentation that would enable additional project-specific review and the comparison of data between the international development organisations.

The development of the IPMM is a particularly innovative contribution to new knowledge from this research in answer to the first research question. The IPMM questionnaire has been constructed to provide a standardised model for comparison and benchmarking of international development organisations. For future research, and to address the first research limitation, the IPMM questionnaire should be distributed to additional development organisations to potentially increase data and documentation. Further research may consider modification of the collection of feedback from a panel of procurement experts for the assessment and development of VFM factors in this research with the use of a Delphi survey. A Delphi survey can be used to increase the data collection from a larger number of subject matter experts and expand the feedback process with more review rounds. In addition, weighting factors can be adapted to prioritise other aspects of interest, such as Environmental, Social, Governance (ESG) or corruption as required by the development organisations (Cheema & Langa, 2022).

As discussed to answer the second research question, project evaluation data are comprised of data from a sample of individual projects ( $n = 1,920$ ), between 2015 and 2017, for ten international development organisations. The secondary data for the project evaluation reports includes overall project evaluation results, evaluation criteria results, cost estimates and actual costs for the development projects. The research findings conclude that there are inconsistencies with the project evaluation process and criteria and they do not effectively measure DE and VFM. As with the IPMM primary data collection, and to address the second and third research questions, further detailed analysis of specific projects should allow additional review of the secondary project data. In addition, it is recommended that project evaluation reports should include detailed



procurement performance metrics specifically related to the project under evaluation.

Development organisations dedicate a substantial effort to mitigate and address corruption in the disbursement and delivery of development funds and projects and often with mixed results (Harnois & Gagnon, 2021; Kuipers, 2021). Further research is recommended on the relationship between project evaluation results and corruption, which may provide additional insight into initiatives to mitigate corruption and waste.

The research acknowledges the continuous efforts to enhance the project evaluation methodology made by the international development organisations for procurement and project management process improvement (Clements et al., 2008; Picciotto, 2020).

To contribute to these efforts, the following recommendations are made for review and adjustment of the four current OECD-DAC (2019) project evaluation criteria:

- Review and adjust the OECD-DAC (2019) definitions and align the rating scales of the project evaluation criteria for improved objectivity and consistency for more tangible measurement and interpretation of the project evaluation results.
- The definition of the relevance criterion should be revised and adjusted to reflect the quality of project inputs, design, scope, specifications, budget preparation, planning and strategy of the development goals and objectives of the projects.
- The definition of the sustainability criterion should be adjusted to reflect the long-term impact and outputs of the project and connect the micro level project implementation and delivery results with overall macro DE goals and objectives.
- Review and adjust the effectiveness criterion to reflect the implementation and adherence to the design, quality and scope of the project objectives including the achievement of additional project goals, such as sustainability or ESG targets.

- The definition of the efficiency criterion should be adjusted to reflect time and cost as primary Critical Success Factors (CSRs) with improvement in recording and managing of Cost Estimates (CE) and monitoring of actual project costs.

Future research and analysis should then be able to address specific attributes of the measurement of procurement performance and project evaluation as identified in this research leading to improved ability to measure and report on DE and VFM in projects.

Clements (2020) supported the view that the measurement of cost-related data should provide further insight into DE and VFM (Ahsan & Gunawan, 2010).

Further research should expand the approach to cost estimation and provide analysis on funds not disbursed due to underruns and additional funds disbursed due to overruns.

Finally, future research questions should examine how the measurement of procurement performance and project evaluation contribute towards critical development and disbursement decisions to ensure that best practices are recognised and incentivised.

## **7.7 SUMMARY OF THE CHAPTER**

Chapter 7 draws a conclusion to the thesis by summarising the literature review, methodology, results and discussion of the thesis and discusses limitations and recommendations for further research that emerges from the research findings.

The chapter outlines contributions to theory and practice from the research that has been conducted to answer the four research questions described in Chapter 3. The research examines the relationship between procurement, project management and evaluation in the theoretical framework shown in Figure 6, and examines how the delivery mechanisms contribute to DE and VFM for projects in international development.

The procurement and project management processes are integral to the implementation, project and cost performance of projects in both the private sector and international

development projects (Schiele, 2007). There are four different levels identified for measurement of DE in development, as illustrated in the hierarchy of performance measurement in Figure 2 (Biscaye et al., 2015; Clements et al., 2008).

The introduction in this research provides an oversight and discussion on the estimated US\$161.2 billion dollars of Official Development Assistance (ODA), collectively disbursed towards global development in 2020 (OECD, 2017; 2021). Ackman and Motty (2020) argued that demand is dramatically increasing with the reported US\$5.7 trillion anticipated future investment in green infrastructure (Ackom & Motty, 2020). This includes US\$100 billion of annual financing for multilateral climate finance mechanisms such as the Global Climate Fund (GCF) and Global Environmental Facility (GEF) up to 2020 (Cui & Huang, 2018). In addition to this to these substantial numbers, GEF has provided US\$17 billion in grants, and US\$88 billion in joint financing since 1991, for over 4,000 projects in 160 developing countries (Cui, Sun, Song, & Zhu, 2020).

International development agencies can come under great pressure to meet disbursement targets, respond to demands for ODA and to promote national interests (Gulrajani, 2017). The pressure is doubtlessly driven by good intentions and to respond and resolve immediate and perceived development requirements in multiple regions and sectors, such as poverty and inequality, and mitigation and alleviation of disasters and emergencies. Easterly (2002) referred to this as a “cartel of good intentions” and also suggested that the numerous simultaneous demand from multiple organisations makes DE on aggregate, very difficult to evaluate. Development organisations such as MDBs and UN Agencies often suffer from disbursement pressure where there is motivation and behavioural incentives by donors to disburse funds, often rather quickly, to meet annual targets (De Haan, 2009). Clements et al. (2008) argued that rapid disbursement may achieve the quantity of development aid. However, without strong monitoring and evaluation, it is

unlikely to ensure delivery of either quality of goods, services and works or achieve effective development goals (Clements et al., 2008; Roodman, 2006).

This research supports current knowledge that both bilateral and multilateral agencies have not been effective in meeting their own development effectiveness commitments and that DE and VFM are still not effectively measured overall (Palagashvili, 2019).

The research findings indicate that the measurement of development projects at the bottom two levels of the hierarchy of performance measurement in Figure 2, contributes towards to the two top level macro-type indicators, such as thematic and country development goals. However, while both the macro and micro-levels of performance measurement are connected, macro-level metrics have different perspectives for the measurement of development performance. The research, therefore, makes the case that improvement of measurement of DE and VFM is necessary at both the macro-level and at the micro-level which includes the procurement and project management processes.

The research reveals an important finding that the development project delivery mechanisms of procurement and project management have very weak correlation with each other in the context of international development projects. However, the research results do not necessarily imply that there is no relationship between the two processes, and several explanations are made to interpret the correlation results, as discussed in Section 6.4. The research also identifies links, or common thread, between cost data and VFM that run predominantly through the efficiency project evaluation criterion.

A key implication of the research findings is the need for the improvement of the project evaluation methodology and the definitions of the evaluation criteria to provide more effective measurement of DE and VFM in development programmes and projects.

Another important implication and recommendation of the research is for development organisations to prioritise, plan and design project requirements and evaluate past project performance prior to the initiation of new development projects and the disbursement of funds (Ferry et al., 2020; Harnois & Gagnon, 2021; Yonehara et al., 2017).

In closing, the aspirations of this research were to understand how procurement contributes towards project performance, DE and VFM in development projects. The research also examined how project performance is evaluated and reported. While the research questions have been addressed, some limitations have been observed.

Overall, the research has filled a substantial gap in current knowledge with the ultimate message being that, despite the enormous effort and resources dedicated to assessment of DE and VFM, meaningful measurement remains inconsistent. The important implication for stakeholders and donors is that reliable measurement of project performance, DE and VFM should be established prior to setting objectives and disbursing development funds.

## REFERENCES

- Aberdeen Group. (2006). *Global Supply Chain Benchmark Report. industry Priorities for Visibility, B2B Collaboration, Trade Compliance and Risk Management November 2006*. Aberdeen Group, Boston, MA.
- Ackom, E., & Motty, M. (2019). Climate finance: Unlocking funds toward achievement of climate targets under the paris agreement. In *Leal Filho W., Azul A., Brandli L., Özuyar P., Wall T. (eds) Climate Action. Encyclopedia of the UN sustainable development goals*. Cham. Springer.
- Adams, A., & Cox, A. L. (2008). Questionnaires, in-depth interviews and focus groups. In *Research Methods for Human Computer Interaction*, (pp. 17-34). Cambridge, UK: Cambridge University Press.
- Adler, M., & Ziglio, E. (1996). *Gazing into the oracle: The Delphi method and its application to social policy and public health*. Cambridge, UK. Jessica Kingsley Publishers.
- Africa Development Bank. (2018). *AFDB Listing of Awarded Contracts 2011 to 2016*. Retrieved from <https://www.AFDB.org/en/documents/document/listing-of-awarded-contracts-from-2011-to-2016-january-2017-94157/>
- African Development Bank. (2015). *Methodology for Implementation of the Procurement Policy of the African Development Bank*. Retrieved from <https://www.AFDB.org/en/projects-and-operations/procurement/new-procurement-policy/>
- African Development Bank. (2018). *African Development Bank Summary of Procurement Awards (AFDB)*. Retrieved from <https://www.AFDB.org/en/documents/document/summary-of-procurement-statistics-by-year-september-2018-103664/>

- Ahiaga-Dagbui, D. D., & Smith, S. D. (2014). Rethinking construction cost overruns: cognition, learning and estimation. *Journal of Financial Management of Property and Construction*, 19(1), 38-54.
- Ahsan, K. (2012). Determinants of the performance of public sector development projects. *International Journal of Management*, 29(1), 77.
- Ahsan, K., & Gunawan, I. (2010). Analysis of cost and schedule performance of international development projects. *International Journal of Project Management*, 28(1), 68-78.
- Ahsan, K., & Paul, S. K. (2018). Procurement issues in donor-funded international development projects. *Journal of Management in Engineering*, 34(6), 04018041.
- AIAG. (2010). Automotive Industry Action Group. Retrieved from <https://www.aiag.org/>
- Al-Jundi, S. A., Shuhaiber, A., & Al-Emara, S. S. (2019). The effect of culture and organisational culture on administrative corruption. *International Journal of Economics and Business Research*, 18(4), 436-451.
- Allen, J. P. (2017). Information technology and wealth concentration, In *Technology and Inequality* (pp. 25-41). New York. Palgrave Macmillan.
- American Productivity and Quality Center APQC. (2018). *APQC Benchmarking*. Retrieved from <https://www.apqc.org/benchmarking>
- Ampratwum, E. F. (2008). The fight against corruption and its implications for development in developing and transition economies. *Journal of Money Laundering Control* 11(1), 76-87.
- Anantatmula, V., & Thomas, M. (2010). Managing global projects: A structured approach for better performance. *Project Management Journal*, 41(2), 60-72.

- Anderson, M., & Katz, P. (1998). Strategic sourcing. *The International Journal of Logistics Management*, 9(1), 1-13.
- Andrić, J. M., Mahamadu, A. M., Wang, J., Zou, P. X., & Zhong, R. (2019). The cost performance and causes of overruns in infrastructure development projects in Asia. *Journal of Civil Engineering and Management*, 25(3), 203-214.
- APICS. (2010). *APICS Supply Chain Operations Reference (SCOR) Model*. Retrieved from <http://www.apics.org/apics-for-business/frameworks/scor>
- Arlbjørn, J.S., & Freytag, P.V. (2012). Public procurement vs private purchasing: is there any foundation for comparing and learning across the sectors? *International Journal of Public Sector Management*, 25(3), 203-220.
- Arrowsmith, S. (2010). Horizontal policies in public procurement: a taxonomy. *Journal of Public Procurement*, 10(2), 149-186
- Ashforth, B. E., & Anand, V. (2003). The normalization of corruption in organizations. *Research in Organizational Behavior*, 25, 1-52.
- Asian Development Bank. (2015). *Asian Development Bank, Procurement Documents*. Retrieved from <https://www.adb.org/site/business-opportunities/operational-procurement/goods-services/documents>
- Asian Development Bank. (2017). *Asian Development Bank Major Contract Awards (ADB)*. Retrieved from <https://www.adb.org/projects/tenders/group/consulting/type/contracts-awarded-1541>
- Asian Development Bank. (2018). *Procurement Risk Framework; Guidance Note on Procurement*. Retrieved from <https://www.adb.org/business/how-to/does-my-executingimplementing-agency-have-staff-and-capability-work-adb>;



- Asian Development Bank. (2019). *Guidelines for Evaluation of Public Sector Operations*. Retrieved from <https://www.adb.org/site/evaluation/evaluations>
- Asian Development Bank. (2019). *Annual Evaluation Review 2019; Performance and Scorecards*. Retrieved from <https://www.adb.org/sites/default/files/evaluation-document/467896/files/aer-2019.pdf>
- Asian Development Bank. (2019). *Value for Money; Guidance Note on Procurement*. Retrieved from <https://www.adb.org/sites/default/files/procurement-value-money.pdf>
- Asian Development Bank. (2020). *Fact Sheets*. Retrieved from <https://www.adb.org/publications/series/fact-sheets#accordion-0-0>
- Asian Development Bank. (2022). *ADB Procurement Guidelines*. Retrieved from <https://www.adb.org/documents/procurement-guidelines>
- Asiedu, R. O., Frempong, N. K., & Alfen, H. W. (2017). Predicting likelihood of cost overrun in educational projects. *Engineering, Construction and Architectural Management* 24(1), 21-39.
- Asiedu, R. O., Manu, P., Mahamadu, A. M., Booth, C. A., Olomolaiye, P., Agyekum, K., & Abadi, M. (2021). Critical skills for infrastructure procurement: insights from developing country contexts. *Journal of Engineering, Design and Technology*.
- Assaf, S. A., & Al-Hejji, S. (2006). Causes of delay in large construction projects. *International Journal of Project Management*, 24(4), 349-357.
- Astley, W. G., & Van de Ven, A. H. (1983). Central perspectives and debates in organization theory. *Administrative Science Quarterly*, 28(2), 245-273.
- 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(6), 337-342.

- Atkisson, I. (2016). *The SDG Indicators: What are we measuring?* Retrieved from <http://atkisson.com/analysis-of-the-sdg-indicators/>
- Aulia, D., & Isvara, D. (2021). Strategies to Increase Procurement Maturity Level using Procurement Maturity Model to Improve Procurement Performance. *International Journal of Scientific and Research Publications*, 11(6), 489-501.
- Aulia, D., & Isvara, W. (2022, January). The Importance of Procurement Organization, Policy, and Value on Supply Chain Performance. In *2022 5th International Conference on Computers in Management and Business (ICCMB)* (pp. 190-194).
- Axelsson, B., Rozemeijer, F. A., & Wynstra, J. Y. F. (Eds.) (2005). *Developing sourcing capabilities : creating strategic change in purchasing and supply management*. Chichester, UK. John Wiley & Sons.
- Baccarini, D. (1999). The logical framework method for defining project success. *Project Management Journal*, 30(4), 25-32.
- Baccarini, D. (2004). Accuracy in estimating project cost construction contingency-a statistical analysis. *Paper presented at the Cobra 2004: RICS International Construction Conference, Responding to Change, September 2004*. Headingley Stadium, U.K.
- Baccarini, D., & Love, P. E. (2014). Statistical characteristics of cost contingency in water infrastructure projects. *Journal of Construction Engineering and Management*, 140(3), 04013063.
- Baker, R. W. (2005). *Capitalism's Achilles heel: Dirty money and how to renew the free-market system*. New Jersey, USA. John Wiley & Sons.
- Banda, W. (2019). An integrated framework comprising of AHP, expert questionnaire survey and sensitivity analysis for risk assessment in mining projects.

*International Journal of Management Science and Engineering Management*,  
14(3), 180-192.

Bardhan, P. (1997). Corruption and development: a review of issues. *Journal of Economic Literature*, 35(3), 1320-1346.

Barr, J., & Christie, A. (2015). *Improving the Practice of Value for Money Assessment*. Institute of Development Studies. Retrieved from  
[https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/5977/CDI%20PracticePaper\\_12.pdf](https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/5977/CDI%20PracticePaper_12.pdf)

Barry, J., Cavinato, J. L., Green, A., & Young, R. R. (1996). A development model for effective MRO procurement. *International Journal of Purchasing and Materials Management*, 32(2), 35-44.

Batenburg, R., & Versendaal, J. (2008). Maturity Matters: Performance Determinants of the Procurement Business Function. *Proceedings of the 16th European Conference on Information Systems, June 2008*. Galway, Ireland.

Bayiley, Y. T., & Teklu, G. K. (2016). Success factors and criteria in the management of international development projects. *International Journal of Managing Projects in Business*. 9(3), 562-582

Bayram, S., & Al-Jibouri, S. (2018). Cost forecasting using RCF: a case study for planning public building projects costs in Turkey. *International Journal of Construction Management*, 18(5), 405-417.

Beecham, S., Hall, T., Britton, C., Cottee, M., & Rainer, A. (2005). Using an expert panel to validate a requirements process improvement model. *Journal of Systems and Software*, 76(3), 251-275.

Belay, A. M., & Torp, O. (2017). Do Longer Projects Have Larger Cost Deviation Than Shorter Construction Projects? *Procedia Engineering*, 196, 262-269.

- Bell, E., Bryman, A., & Harley, B. (2018). *Business research methods*. Oxford, UK. Oxford University Press.
- Bergmann, T., & Karwowski, W. (2018, July). Agile project management and project success: A literature review. In *International Conference on Applied Human Factors and Ergonomics* (pp. 405-414). Springer, Cham.
- Bertisen, J., & Davis, G. A. (2008). Bias and error in mine project capital cost estimation. *The Engineering Economist*, *53*(2), 118-139.
- Bezzina, F., & Saunders, M. (2014). The pervasiveness of statistical misconceptions: A study of academics with a special interest in business research methods. *Leading Issues in Leading Issues in Business and Management Research*, *2*(2), 127.
- Bhargava, A., Anastasopoulos, P. C., Labi, S., Sinha, K. C., & Mannering, F. L. (2010). Three-stage least-squares analysis of time and cost overruns in construction contracts. *Journal of Construction Engineering and Management*, *136*(11), 1207-1218.
- Birdsall, N., Kharas, H. J., Mahgoub, A., & Perakis, R. (2010). *Quality of official development assistance assessment*. Center for Global Development, Washington, DC.
- Biscaye, LaFayette, & Martin. (2015). *Evaluating Donor-Level Results Measurement Systems*. University of Washington. Evans School Policy Analysis and Research. EPAR Technical Report. Retrieved from Microsoft Word - EPAR\_UW\_294\_Multilateral vs Bilateral Aid\_3.30.15.docx
- Biscaye, LaFayette, Martin, Richardson, True, & Anderson. (2017). Relative Effectiveness of Bilateral and Multilateral Aid on Development Outcomes. *Review of Development Economics*, *21*(4), 1425-1447.

- Blair, M. M. (2010). *Financial innovation and the distribution of wealth and income*.  
 Vanderbilt Law and Economics Research Paper No. 10-22, Vanderbilt Public  
 Law Research Paper No. 10-32. Retrieved from  
<https://ssrn.com/abstract=1656451>
- Bloch, A. (2011). *Procurement Maturity: A tool for Supply Chain improvement*. KTH  
 Royal Institute of Technology, Sweden.
- Boakye, L. G. (2015). *The Underlying Reasons Why International Development Projects  
 (IDPs) Fail: The Case of African Development Bank (AfDB)-Funded  
 Projects* [Master's thesis, University of Sydney]  
<https://ses.library.usyd.edu.au/handle/2123/14138>
- Bolstorff, P. (2001). *How do I use SCOR*. Retrieved from  
[http://www.any2any.net/RFIDAware/scmGroup/document/How%20to%20use%  
 20SCOR.pdf](http://www.any2any.net/RFIDAware/scmGroup/document/How%20to%20use%20SCOR.pdf)
- Boone, P. (1996). Politics and the effectiveness of foreign aid. *European Economic  
 Review*, 40(2), 289-329.
- Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of extension*, 50(2),  
 1-5.
- Borson, F. (2017). Reforms under the World Bank Procurement and the Policy  
 Implications for Developing Countries. *Eur. Procurement & Pub. Private  
 Partnership L. Rev.*, 12, 146.
- Bossu, V., Greset, D., Micaud, F., Pavie, A., Favre Bertin, M., & Estampe, D. (2004). Le  
 métier de supply chain manager. *Logistique & Management*, 12(1), 83-91.
- Bourguignon, F., & Sundberg, M. (2007). Aid effectiveness: opening the black box. *The  
 American Economic Review*, 97(2), 316-321.

- Brandmeier, R. A., & Rupp, F. (2010). Benchmarking procurement functions: causes for superior performance. *Benchmarking: An International Journal*, 17(1), 5-26.
- Brech, V., & Potrafke, N. (2014). Donor ideology and types of foreign aid. *Journal of Comparative Economics*, 42(1), 61-75.
- Brookes, N., Butler, M., Dey, P., & Clark, R. (2014). The use of maturity models in improving project management performance: An empirical investigation. *International Journal of Managing Projects in Business*, 7(2), 231-246.
- Brown, S. (2012). National development agencies and bilateral aid. In *Introduction to International Development: Approaches, Actors, and Issues, 2<sup>nd</sup> Edition* (pp. 143-158). Oxford. Oxford University Press.
- Brown, S. (2020). The rise and fall of the aid effectiveness norm. *The European Journal of Development Research*, 32(4), 1230-1248.
- Brugha, R., Donoghue, M., Starling, M., Ndubani, P., Ssenooba, F., Fernandes, B., & Walt, G. (2004). The Global Fund: managing great expectations. *The Lancet*, 364(9428), 95-100.
- Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research*, 6(1), 97-113.
- Bulman, D., Kolkma, W., & Kraay, A. (2017). Good countries or good projects? Comparing macro and micro correlates of World Bank and Asian Development Bank project performance. *The Review of International Organizations*, 12(3), 335-363.
- Burke, S., & Macdonald, S. (2014). Creativity and conservation: managing significance at the Sydney Opera House. *APT Bulletin*, 45(2), 3.

- Burnside, C., & Dollar, D. (2000). Aid, policies, and growth. *American Economic Review*, 90(4), 847-868.
- Cameron, R., & Molina-Azorin, J. F. (2011). The acceptance of mixed methods in business and management research. *International Journal of Organizational Analysis*, 19(3), 256-271.
- Cammish, R., & Keough, M. (1991). A strategic role for purchasing. *The McKinsey Quarterly* (3), 22-40.
- Campbell, S., & Cantrill, J. (2001). Consensus methods in prescribing research. *Journal of Clinical Pharmacy and Therapeutics*, 26(1), 5-14.
- CAPS Research. (2010). *CAPS Research Cross Industry Report of Standard Benchmarks*. Retrieved from <https://knowledge.capsresearch.org>
- CAPS Research. (2012). *CAPS Research Cross Industry Report of Standard Benchmarks*. Retrieved from <https://knowledge.capsresearch.org>
- CAPS Research. (2014). *CAPS Research Cross Industry Report of Standard Benchmarks*. Retrieved from <https://knowledge.capsresearch.org>
- Cavinato, J. L. (1999). Fitting purchasing to the five stages of strategic management. *European Journal of Purchasing & Supply Management*, 5(2), 75-83.
- Chan, A., & Chan, A. (2004). Key performance indicators for measuring construction success. *Benchmarking: An International Journal*, 11(2), 203-221.
- Chandes, J., Estampe, D., Berthomier, R., Courrié, L.-A., Han, L., & Marquevielle, S. (2003). *Logistics performance of actors in the wine supply chain*. In *Supply Chain Forum: An International Journal*, 4(1), 12-27.
- Charlesworth, H., & Chinkin, C. (2013). *The creation of UN Women*. RegNet research paper series, 2013/7. Regulatory Institutions Network (RegNet), Canberra, Australia.

- Chartered Institute of Procurement and Supply CIPS. (2020). Procurement Glossary. Retrieved from <https://www.cips.org/knowledge/procurement-glossary/v/>
- Chauvet, L., Collier, P., & Duponchel, M. (2010). *What explains aid project success in post-conflict situations?* Working Paper 5418, World Bank Development Research Group.
- Cheema, S., & Langa, M. (2022). Environment, Social, and Governance (ESG) and Sustainability. In *A Director's Guide to Governance in the Boardroom* (pp. 135-171). Routledge.
- Cherono, V., & Chepwony, J. (2021). *Journal of Procurement*, 8(1), 55-77.
- Cherono, K. J., & Patrick, M. Wangani. (2017). Influence of Procurement Management Process on Performance of Donor Funded Projects in Kenya (A case of African Development Bank in Kenya). *International Journal of Logistics and Procurement Management*, 1(1), 134-145.
- Chianca, T. (2008). The OECD/DAC criteria for international development evaluations: An assessment and ideas for improvement. *Journal of Multidisciplinary Evaluation*, 5(9), 41-51.
- Chrissis, M. B., Konrad, M., & Shrum, S. (2003). *CMMI guidelines for process integration and product improvement*. Boston, USA. Addison-Wesley Longman Publishing Co., Inc.
- CIPS Australia. (2020). Chartered Institute of Procurement and Supply Australia (CIPSA). Retrieved from <https://www.cips.org/en-au/knowledge/welcome-to-cips-knowledge/>
- Clegg, L. (2015). Benchmarking and blame games: Exploring the contestation of the Millennium Development Goals. *Review of International Studies*, 41(5), 947-967.



- Clements, P. (2020). Improving learning and accountability in foreign aid. *World Development, 125*, 104670.
- Clements, P., Chianca, T., & Sasaki, R. (2008). Reducing world poverty by improving evaluation of development aid. *American Journal of Evaluation, 29*(2), 195-214.
- Collier, P., & Dollar, D. (2004). Development Effectiveness: What have we learnt? *The Economic Journal, 114*(496), 244-271.
- Collier, D. A., & Evans, J. R. (2020). *Operations and supply chain management*. Boston. London. Cengage Learning.
- Comiskey, C., & Dempsey, O. (2013). 18 Analysing data from small and large samples and non-normal and normal distributions. In *Quantitative Health Research: Issues and Methods* (pp. 348). London. McGraw-Hill.
- Cooke-Davies. (2002). The “real” success factors on projects. *International Journal of Project Management, 20*(3), 185-190.
- Cooke-Davies, & Arzymanow, A. (2003). The maturity of project management in different industries: An investigation into variations between project management models. *International Journal of Project Management, 21*(6), 471-478.
- Cousins, P. D., Lawson, B., & Squire, B. (2006). An empirical taxonomy of purchasing functions. *International Journal of Operations & Production Management, 26*(7), 775-794.
- Crawford, P., & Bryce, P. (2003). Project monitoring and evaluation: a method for enhancing the efficiency and effectiveness of aid project implementation. *International Journal of Project Management, 21*(5), 363-373.
- Croxton, K. L., Garcia-Dastugue, S. J., Lambert, D. M., & Rogers, D. S. (2001). The supply chain management processes. *The International Journal of Logistics Management, 12*(2), 13-36.

- Cui, L., & Huang, Y. (2018). Exploring the schemes for green climate fund financing: international lessons. *World Development*, *101*, 173-187.
- Cui, L., Sun, Y., Song, M., & Zhu, L. (2020). Co-financing in the green climate fund: Lessons from the global environment facility. *Climate Policy*, *20*(1), 95-108.
- Cummins, R. A., & Gullone, E. (2000). Why we should not use 5-point Likert scales: The case for subjective quality of life measurement. *Paper presented at the Proceedings, Second International Conference on Quality of Life in Cities, National University of Singapore, March 2000* (pp. 74-93). Kent Ridge, Singapore
- Curtis, E. A., Comiskey, C., & Dempsey, O. (2016). Importance and use of correlational research. *Nurse Researcher*, *23*(6).
- Dancey, C. P., & Reidy, J. (2007). *Statistics without Maths for Psychology*. Harlow, UK Prentice Hall/Pearson Education.
- Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research*, *50*(1), 61-104.
- de Araújo, M. C. B., Alencar, L. H., & de Miranda Mota, C. M. (2017). Project procurement management: A structured literature review. *International Journal of Project Management*, *35*(3), 353-377.
- De Graaf, G. (2007). Causes of corruption: Towards a contextual theory of corruption. *Public Administration Quarterly*, *31*(1/2), 39-86.  
<http://www.jstor.org/stable/41288280>
- De Haan, A. (2009). *How the aid industry works: An introduction to international development*. Sterling, USA. Kumarian Press.
- de Montclos, M. A. P. (2012). Humanitarian action in developing countries: Who evaluates who? *Evaluation and Program Planning*, *35*(1), 154-160.

- de Souza, T. F., & Gomes, C. F. S. (2015). Assessment of maturity in project management: A bibliometric study of main models. *Procedia Computer Science*, 55, 92-101.
- De Vaus, D. (2013). *Surveys in Social Research*. London, UK. Routledge.
- De Wit, A. (1988). Measurement of project success. *International Journal of Project Management*, 6(3), 164-170.
- Denizer, C., Kaufmann, D., & Kraay, A. (2013). Good countries or good projects? Macro and micro correlates of World Bank project performance. *Journal of Development Economics*, 105, 288-302.
- Denscombe, M. (2008). Communities of practice: A research paradigm for the mixed methods approach. *Journal of Mixed Methods Research*, 2(3), 270-283.
- Department for International Development (DFID). (2007). *Procurement Capability Review Model and Standards Framework v1.0*. Retrieved from <https://www.gov.uk/government/organisations/department-for-international-development/about/procurement>
- Department for International Development (DFID). (2011). *DFID's Approach to Value for Money (VfM)*. Retrieved from [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/49551/DFID-approach-value-money.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/49551/DFID-approach-value-money.pdf);
- Department for International Development (DFID). (2015). *Why corruption matters: understanding causes, effects and how to address them*. Retrieved from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/406346/corruption-evidence-paper-why-corruption-matters.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/406346/corruption-evidence-paper-why-corruption-matters.pdf)
- Department for International Development (DFID). (2016). *Raising the standard: The Multilateral Development Review 2016*. Retrieved from

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/573884/Multilateral-Development-Review-Dec2016.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/573884/Multilateral-Development-Review-Dec2016.pdf)

Department for International Development (DFID). (2018). *Procurement at DFID*.

Retrieved from <https://www.gov.uk/government/organisations/department-for-international-development/about/procurement>

Department for International Development (DFID). (2019). *Evaluation Annual Report*.

Retrieved from <https://www.gov.uk/government/publications/dfid-evaluation-annual-report-2017-to-2018>

Department for International Development (DFID). (2020). *Evaluation Development*

*Tracker*. Retrieved from <https://devtracker.fcdo.gov.uk/department/FCDO>

Desouza, K. C., & Evaristo, J. R. (2006). Project management offices: A case of

knowledge-based archetypes. *International Journal of Information Management*, 26(5), 414-423.

Diallo, A., & Thuillier, D. (2005). The success of international development projects,

trust and communication: an African perspective. *International Journal of Project Management*, 23(3), 237-252.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and*

*mixed-mode surveys: The tailored design method*. New Jersey, USA. John Wiley & Sons.

Dixon, J. A., Carpenter, R. A., Fallon, L. A., Sherman, P. B., & Manipomoke, S. (2013).

*Economic analysis of the environmental impacts of development projects*.

London. Routledge.

Doloi, H. (2013). Cost overruns and failure in project management: Understanding the

roles of key stakeholders in construction projects. *Journal of Construction Engineering and Management*, 139(3), 267-279.

- Domberger, S., Hall, C., & Li, E. A. L. (1995). The determinants of price and quality in competitively tendered contracts. *The Economic Journal*, 1454-1470.
- Doucouliagos, H., & Paldam, M. (2008). Aid effectiveness on growth: A meta study. *European Journal of Political Economy*, 24(1), 1-24.
- Doucouliagos, H., & Paldam, M. (2009). The aid effectiveness literature: The sad results of 40 years of research. *Journal of Economic Surveys*, 23(3), 433-461.
- Doucouliagos, H., & Paldam, M. (2011). The ineffectiveness of development aid on growth: An update. *European Journal of Political Economy*, 27(2), 399-404.
- Driscoll, D. L., Appiah-Yeboah, A., Salib, P., & Rupert, D. J. (2007). Merging qualitative and quantitative data in mixed methods research: How to and why not. *Ecological and Environmental Anthropology*, 3(1), 19-28.
- Du Plessis, P., & Majam, T. (2010). Mixed method research-A new paradigm. *Journal of Public Administration*, 45(3), 456-472.
- Dubois, A., & Gadde, L.-E. (2002). Systematic combining: An abductive approach to case research. *Journal of Business Research*, 55(7), 553-560.
- Dyba, T. (2000). An instrument for measuring the key factors of success in software process improvement. *Empirical Software Engineering*, 5(4), 357-390.
- Easterly, W. (2002). The cartel of good intentions: The problem of bureaucracy in foreign aid. *The Journal of Policy Reform*, 5(4), 223-250.
- Easterly, W. (2003). Can foreign aid buy growth? *Journal of Economic Perspectives*, 17(3), 23-48.
- Easterly, W., & Pfutze, T. (2008). Where does the money go? Best and worst practices in foreign aid. *The Journal of Economic Perspectives*, 22(2), 29-52.

- Edwards, P., Roberts, I., Clarke, M., DiGuseppi, C., Pratap, S., Wentz, R., & Kwan, I. (2002). Increasing response rates to postal questionnaires: systematic review. *BMJ*, 324(7347), 1183.
- EFQM. (2018). EFQM Excellence Model. Retrieved from <http://www.efqm.org/>
- Elalfy, A., Weber, O. & Geobey, S. (2021), The Sustainable Development Goals (SDGs): a rising tide lifts all boats? Global reporting implications in a post SDGs world, *Journal of Applied Accounting Research*, Vol. 22 No. 3, pp. 557-575
- El-Reedy, M. A. (2016). *Project management in the oil and gas industry*, New Jersey, USA. John Wiley & Sons.
- El Emam, K., & Madhavji, N. H. (1996). An instrument for measuring the success of the requirements engineering process in information systems development. *Empirical Software Engineering*, 1(3), 201-240.
- Elhaniash, F. E. A., & Stevovic, S. (2016). Measurement the efficiency of building project management. *Ekonomika*, 62(4), 129.
- Elkjaer, B. and Simpson, B. (2011). Pragmatism: A lived and living philosophy. What can it offer to contemporary organization theory? *Philosophy and Organization Theory. Research in the Sociology of Organizations* 32, 55-84.
- Ellmers, B. (2011). *How to spend it. Smart procurement for more effective aid*. European Network on Debt and Development. Eurodad. Retrieved from [DEEEP-REPORT-2018-017.pdf \(concordeurope.org\)](https://concordeurope.org/DEEEP-REPORT-2018-017.pdf)
- Emmi, A., Ozlem, E., Maja, K., Ilan, R., & Florian, S. (2011). *Value for money: Current approaches and evolving debates*. London, UK, London School of Economics. [vfm-current-approaches-and-evolving-debates.pdf \(bigpushforward.com\)](https://bigpushforward.com/vfm-current-approaches-and-evolving-debates.pdf)
- Engen, L., & Prizzon, A. (2018). A guide to multilateral development banks. *Overseas Development Institute, Ed.*

- Ernst & Young. (2014). *Oil and Gas Megaprojects*. Retrieved from [http://www.ey.com/Publication/vwLUAssets/EY-spotlight-on-oil-and-gas-megaprojects/\\$FILE/EY-spotlight-on-oil-and-gas-megaprojects.pdf](http://www.ey.com/Publication/vwLUAssets/EY-spotlight-on-oil-and-gas-megaprojects/$FILE/EY-spotlight-on-oil-and-gas-megaprojects.pdf)
- Escadafal, R. (2014). Is the EIRR, the Economic Internal Rate of Return of Projects to Combat Desertification and Land Degradation Evaluation a Relevant Tool for Evaluation? *Planet@ Risk*, 2(1).
- Essmann, H., & Du Preez, N. (2009). An innovation capability maturity model—development and initial application. *International Journal of Industrial and manufacturing Engineering*, 3(5), 382-393.
- Estampe, D., Lamouri, S., Paris, J.-L., & Brahim-Djelloul, S. (2013). A framework for analysing supply chain performance evaluation models. *International Journal of Production Economics*, 142(2), 247-258.
- Esteves, P., & Soares, F. L. D. T. (2020). Regulated Improvisations: bringing the private sector back into the International Development Co-operation field. *Revista Brasileira de Política Internacional*, 63.
- European Bank for Reconstruction and Development. (2011). *Capacity Assessment Toolkit; Manual and Guide Note*. Retrieved from [http://www.ebrd.com/downloads/procurement/project/EBRD\\_Capacity\\_Building\\_Report.pdf](http://www.ebrd.com/downloads/procurement/project/EBRD_Capacity_Building_Report.pdf)
- European Bank for Reconstruction and Development. (2012). *Capacity building in procurement; A sample assessment in the EBRD region*. Retrieved from [http://www.ebrd.com/downloads/procurement/project/EBRD\\_Capacity\\_Building\\_Report.pdf](http://www.ebrd.com/downloads/procurement/project/EBRD_Capacity_Building_Report.pdf)
- European Bank for Reconstruction and Development. (2017). *EBRD Annual Review 2017*. Retrieved from <http://017.ar-ey.com/at-a-glance/>

- European Bank for Reconstruction and Development. (2020). *EBRD Annual Evaluation Review*. Retrieved from <https://www.ebrd.com/what-we-do/annual-evaluation-review.html>
- Eyben, R. (2013). Struggles in Paris: The DAC and the purposes of development aid. *The European Journal of Development Research*, 25(1), 78-91.
- Fabian, F. H. (2000). Keeping the tension: Pressures to keep the controversy in the management discipline. *Academy of Management Review*, 25(2), 350-371.
- Fafchamps, M., & Woodruff, C. (2016). Identifying gazelles: Expert panels vs. surveys as a means to identify firms with rapid growth potential. *The World Bank Economic Review*, 31(3), 670-686.
- Famiyeh, S., Amoatey, C.T., Adaku, E. and Agbenohevi, C.S. (2017), Major causes of construction time and cost overruns: A case of selected educational sector projects in Ghana, *Journal of Engineering, Design and Technology*, Vol. 15 No. 2, pp. 181-198.
- Feilzer, M. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, 4(1), 6-16.
- Ferry, L. L., Hafner-Burton, E. M., & Schneider, C. J. (2020). Catch me if you care: International development organizations and national corruption. *The Review of International Organizations*, 15(4), 767-792.
- Field, A. (2017). *Discovering statistics using IBM SPSS statistics: North American edition*. London, UK. Sage Publications.
- Findley, M. G., Milner, H. V., & Nielson, D. L. (2017). The choice among aid donors: The effects of multilateral vs. bilateral aid on recipient behavioral support. *The Review of International Organizations*, 12(2), 307-334.



- Fjeldstad, O.-H., & Isaksen, J. (2008). *Anti-Corruption Reforms: Challenges, Effects and Limits of World Bank Support. Public Sector Reform: What Works and Why? An IEG Evaluation of World Bank Support Working Paper 2008/7*, USA. World Bank Group.
- Fleming, F. (2013). *Evaluation methods for assessing value for money*. BetterEvaluation, published online. Retrieved from [https://beamexchange.org/uploads/filer\\_public/c7/4e/c74e0337-e2be-4583-af2c-7acd9c7de97a/evaluating\\_methods\\_for\\_assessing\\_vfm\\_compressed.pdf](https://beamexchange.org/uploads/filer_public/c7/4e/c74e0337-e2be-4583-af2c-7acd9c7de97a/evaluating_methods_for_assessing_vfm_compressed.pdf)
- Flyvbjerg, B., Skamris Holm, M. K., & Buhl, S. L. (2003). How common and how large are cost overruns in transport infrastructure projects? *Transport Reviews*, 23(1), 71-88.
- Flyvbjerg, B., Ansar, A., Budzier, A., Buhl, S., Cantarelli, C., Garbuio, M., ... & van Wee, B. (2018). Five things you should know about cost overrun. *Transportation Research Part A: Policy and Practice*, 118, 174-190.
- Fowler Jr, F. J. (2013). *Survey research methods*. University of Massachusetts, Boston, USA. Sage Publications.
- Francis, M. (2020). Global Structures of Inequality and Unequal Distribution of Wealth. In *Decent Work and Economic Growth* (pp. 490-498). Cham. Springer International Publishing.
- Freeman, V. T., & Cavinato, J. L. (1990). Fitting purchasing to the strategic firm: frameworks, processes, and values. *Journal of Purchasing & Materials Management*, 26(1), 6-11.
- Frefer, A., Mahmoud, M., Haleema, H., & Almamlook, R. (2018). Overview success criteria and critical success factors in project management. *Industrial Engineering & Management*, 7(1), 1-6.

- Fukuda-Parr, S., Greenstein, J., & Stewart, D. (2013). How should MDG success and failure be judged: Faster progress or achieving the targets? *World Development*, 41, 19-30.
- Fukuda-Parr, S., & McNeill, D. (2015). Post 2015: A new era of accountability? *Journal of Global Ethics*, 11(1), 10-17.
- Gartner. (2018). *Gartner Benchmark Analytics*. Retrieved from <https://www.gartner.com/en/consulting/gartner-benchmark-analytics>
- Gilmour, P. (1999). A strategic audit framework to improve supply chain performance. *Journal of Business & Industrial Marketing*, 14(5/6), 355-366.
- Giunipero, L. C., Hooker, R. E., Joseph-Mathews, S., Yoon, T. E., & Brudvig, S. (2008). A decade of SCM literature: past, present and future implications. *Journal of Supply Chain Management*, 44(4), 66-86.
- GIZ. (2016). *Annual Statement of Accounts 2016, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH*. Retrieved from <https://www.giz.de/en/downloads/giz2017-en-jahresabschluss-2016.pdf>
- GIZ. (2017). *Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH Knowing what works*. Retrieved from <https://www.giz.de/en/aboutgiz/516.html>
- Gladilina, I. P. (2017). Theory and practice of managing public procurement based on key professional competencies. *European Research Studies Journal*, XX(4B), 439-452.
- Glassman, A., Fan, V., Over, M., Silverman, R., McQueston, K., & Duran, D. (2013). *More health for the money: A practical agenda for the Global Fund and its partners* (pp 1-96). Center For Global Development. Retrieved from ([cgdev.org](http://cgdev.org)) Washington DC.

- Global Policy Forum. (2013). *Total UN System Contributions*. Retrieved from <https://www.globalpolicy.org/un-finance/tables-and-charts-on-un-finance/un-system-budget/27505-total-un-system-contributions.html>
- Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*, 21(2), 135-146.
- Golini, R., Kalchschmidt, M., & Landoni, P. (2015). Adoption of project management practices: The impact on international development projects of non-governmental organizations. *International Journal of Project Management*, 33(3), 650-663.
- Grant, K. P., & Pennypacker, J. S. (2006). Project management maturity: an assessment of project management capabilities among and between selected industries. *IEEE Transactions on engineering management*, 53(1), 59-68.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3), 255-274.
- Greener, S. (2008). *Business research methods*. London, UK. Ventus Publishing.
- Gulrajani, N. (2014). Organising for donor effectiveness: An analytical framework for improving aid effectiveness. *Development Policy Review*, 32(1), 89-112.
- Gulrajani, N. (2016). *Bilateral versus multilateral aid channels*. Overseas Development Institute. Retrieved from <https://www.odi.org/sites/odi.org.uk/files/resource-documents/10393.pdf>
- Gulrajani, N. (2017). Bilateral donors and the age of the national interest: What prospects for challenge by development agencies? *World Development*, 96, 375-389.

- Gupta, R., Karayil, A., & Rajendran, R. (2008). Contract Lifecycle Management. *The DNA of Procurement. Building Tomorrow's Enterprise*. Infosys. Retrieved from [Infosys - Contract Lifecycle Management | SCM \(schradersworld.com\)](https://www.schradersworld.com/Infosys-Contract-Lifecycle-Management-SCM)
- Habibi, M., Kermanshachi, S., & Rouhanizadeh, B. (2019). Identifying and measuring engineering, procurement, and construction (EPC) key performance indicators and management strategies. *Infrastructures*, 4(2), 14.
- Hair Jr, J., Wolfinbarger, M., Money, A., Samouel, P., & Page, M. (2015). *Essentials of business research methods*. New York, USA. Routledge.
- Hák, T., Janoušková, S., & Moldan, B. (2016). Sustainable Development Goals: A need for relevant indicators. *Ecological Indicators*, 60, 565-573.
- Hakim, C. (1987). *Research design: Strategies and choices in the design of social research*. London. Allen and Unwin.
- Halachmi, A., & Greiling, D. (2013). Transparency, e-government, and accountability: Some issues and considerations. *Public Performance & Management Review*, 36(4), 562-584.
- Han, X., Khan, H. A., & Zhuang, J. (2014). *Do governance indicators explain development performance? A Cross-Country Analysis, November 2014*. Working Paper Series 417. Manila, Philippines: Asian Development Bank Economics.
- Hansen, H., & Tarp, F. (2000). Aid effectiveness disputed: Foreign Aid and Development: Lessons Learnt and Directions for the Future. *Journal of International Development*, 12, 375-398.
- Harland, C., Telgen, J., Knight, L., Callender, G., & Thai, K. V. (2007). *24 Challenges facing public procurement. Public procurement: International cases and commentary*. London, UK. Routledge.

- Harnois, Y., & Gagnon, S. (2021). Corruption and international development: a review of project management challenges. *Journal of Financial Crime*, ahead-of-print.
- Hartmann, E., Kerkfeld, D., & Henke, M. (2012). Top and bottom line relevance of purchasing and supply management. *Journal of Purchasing and Supply Management*, 18(1), 22-34.
- Hassan. (2013). Review of the global oil and gas industry: a concise journey from ancient time to modern world. *Petroleum Technology Development Journal*, 3(2), 123-141.
- Hassan, Meyer, D., & Kot, S. (2019). Effect of institutional quality and wealth from oil revenue on economic growth in oil-exporting developing countries. *Sustainability*, 11(13), 3635.
- Hasson, F., & Keeney, S. (2011). Enhancing rigour in the Delphi technique research. *Technological Forecasting and Social Change*, 78(9), 1695-1704.
- Heggstad, K. K., & Frøystad, M. (2011). *The basics of integrity in procurement*. U4 Issue 10. Anti-Corruption Resource Centre. Retrieved from [www.U4.no](http://www.U4.no).
- Henchie, P. L. a. N. (2007). Worlds Apart: EPC and EPCM Contracts. Risk Issues and Allocation. *International Construction Law Review*, 24(1/4), 252.
- Hewson, C. (2003). Conducting research on the internet. *The Psychologist*, 16(6), 290-293.
- Hewson, C. (2015). 14 Research methods on the Internet. *Communication and Technology*, 5, 277.
- Hewson, C., Vogel, C., & Laurent, D. (2015). *Internet research methods*. London. Sage Publications.
- Hodgson, P. (2010). Usability Test Data. Retrieved from <https://www.userfocus.co.uk/articles/datathink.html>

- Hoekman, B. (1998). Using international institutions to improve public procurement. *The World Bank Research Observer*, 13(2), 249-269.
- Holsapple, C. W., & Joshi, K. D. (2000). An investigation of factors that influence the management of knowledge in organizations. *The Journal of Strategic Information Systems*, 9(2-3), 235-261.
- Holzappel, S. (2016). Boosting or hindering aid effectiveness? An assessment of systems for measuring donor agency results. *Public Administration and Development*, 36(1), 3-19.
- Horvarth & Partners. (2018). 360' Supply Management Efficiency Review. Retrieved from <https://www.horvath-partners.com/de/kompetenz/funktionale-kompetenz/operations-einkauf/einkauf/>
- Howes, S., Otor, S., & Rogers, C. (2011). *Is there a micro-macro paradox in international aid, or do the data deceive?* Australian National University, Canberra. Center for Development Policy, Crawford School.
- Huang, S. H., Sheoran, S. K., & Keskar, H. (2005). Computer-assisted supply chain configuration based on supply chain operations reference (SCOR) model. *Computers & Industrial Engineering*, 48(2), 377-394.
- Ibbs, C. W., & Kwak, Y. H. (2000). Assessing project management maturity. *Project Management Journal*, 31(1), 32-43.
- IBM. (2020). *Transforming different Likert scales to a common scale*. Retrieved from <https://www.ibm.com/support/pages/node/422073>
- ICAI. (2011). *ICAI's Approach to Effectiveness and Value for Money, Report 1*, London. Independent Commission on Aid Impact.
- ICAI. (2018). *Achieving value for money through procurement Part 2: DFID's approach to value for money through tendering and contract management*. Retrieved from

<http://devinit.org/post/aid-spending-by-development-assistance-committee-dac-donors-in-2016/#>

IEG. (2013). *IEG Results and Performance of the World Bank Group 2013 Background*.

Retrieved from

[https://ieg.worldbankgroup.org/Data/reports/RAP2013\\_Concept\\_Note\\_for\\_disclosure.pdf](https://ieg.worldbankgroup.org/Data/reports/RAP2013_Concept_Note_for_disclosure.pdf)

IEG. (2014). *The World Bank and Public Procurement—An Independent Evaluation*.

*Volume II: Achieving Development Effectiveness through Procurement in Bank Financial Assistance*. Washington, DC. World Bank. License. Creative

Commons Attribution CC BY 3.0. Retrieved from

<http://documents.worldbank.org/curated/en/702461468331906897/pdf/NonAsciiFileName0.pdf>

IEG. (2021). *World Bank Group: Impact Evaluations: Relevance and Effectiveness*.

Retrieved from <https://www.worldbank.org/en/programs/building-the-evidence-on-forced-displacement/impact-evaluations>

Ika, L. A. (2012). Project management for development in Africa: Why projects are failing and what can be done about it. *Project Management Journal*, 43(4), 27-41.

Ika, L. A., Diallo, A., & Thuillier, D. (2010). Project management in the international development industry: the project coordinator's perspective. *International Journal of Managing Projects in Business* 43(4), 27-41

Ika, L. A., Diallo, A., & Thuillier, D. (2012). Critical success factors for World Bank projects: An empirical investigation. *International Journal of Project Management*, 30(1), 105-116.

- Ika, L. A., & Donnelly, J. (2017). Success conditions for international development capacity building projects. *International Journal of Project Management*, 35(1), 44-63.
- Implement Consulting Group. (2018). *Implement Procurement Maturity Model*. Retrieved from <https://implementconsultinggroup.com/implement-procurement-maturity-model/>
- Integrity Vice Presidency (INT). (2013). *Fraud and Corruption Awareness Handbook*. World Bank Group. Retrieved from [http://siteresources.worldbank.org/INTDOII/Resources/INT\\_inside\\_fraud\\_text\\_090909.pdf](http://siteresources.worldbank.org/INTDOII/Resources/INT_inside_fraud_text_090909.pdf)
- Inter-American Development Bank. (2018). *Guide for the Acceptance of the Use Country Procurement Systems*. Retrieved from <https://www.iadb.org/en/procurement>
- Inter-American Development Bank. (2019). *Inter-American Development Bank Procurement Contract Awards (IADB)*. Retrieved from <https://www.iadb.org/en/procurement>
- Inter-American Development Bank. (2020). *Inter-American Development Bank Evaluation Reports*. Retrieved from <https://www.iadb.org/en/ove/home>
- International Labor Organisation (ILO). (2018). *Labor Standards*. Retrieved from <http://ilo.org/global/standards/lang--en/index.htm>
- International Monetary Fund. (2021). *About the IMF*. Retrieved from <http://www.imf.org/en/About>
- Iriarte, C., & Bayona, S. (2020). IT projects success factors: a literature review. *International Journal of Information Systems and Project Management*, 8(2), 49-78.



- Islamic Development Bank (ISDB). (2016). *Annual Evaluation Report 2015*. Retrieved from <https://www.ecgnet.org/document/annual-evaluation-report-year-1436-h-2015>
- Islamic Development Bank (ISDB). (2017). *Annual Evaluation Report 2016*. Retrieved [https://www.ecgnet.org/sites/default/files/2016\\_IDB-OED\\_AnnualReport.pdf](https://www.ecgnet.org/sites/default/files/2016_IDB-OED_AnnualReport.pdf)
- ISO 14044. Environmental management e life cycle assessment e life cycle impact assessment. Geneva: International Organization for Standardization; 2022.
- Israel, P. (2006). The flash of genius: Defining invention in the era of corporate research. *Paper presented at the Con/texts of Invention: Working Conference of the Society for Critical Exchange, Western Reserve University, April 2006* (pp. 20-23). Cleveland, Ohio.
- Itsubo, N., Murakami, K., Kuriyama, K., Yoshida, K., Tokimatsu, K., & Inaba, A. (2018). Development of weighting factors for G20 countries—explore the difference in environmental awareness between developed and emerging countries. *The International Journal of Life Cycle Assessment*, 23(12), 2311-2326.
- Iyer, K. C., & Banerjee, P. S. (2016). Measuring and benchmarking managerial efficiency of project execution schedule performance. *International Journal of Project Management*, 34(2), 219-236.
- Jackson, P. (2012). *Value for money and international development: Deconstructing myths to promote a more constructive discussion*. Organisation for Economic Co-operation and Development. Retrieved from <http://search.oecd.org/dac/effectiveness/49652541.pdf>
- Jain, P. (2016). Japan's foreign aid: old and new contests. *The Pacific Review*, 29(1), 93-113.

- Jajuga, K., & Walesiak, M. (2000). *Standardisation of data set under different measurement scales* (pp. 105-112). Classification and information processing at the turn of the millennium. Berlin. Springer.
- Janoušková, S., Hák, T., & Moldan, B. (2018). Global SDGs assessments: Helping or confusing indicators? *Sustainability, 10*(5), 1540.
- Jenkins, L., & Anderson, M. (2003). A multivariate statistical approach to reducing the number of variables in data envelopment analysis. *European Journal of Operational Research, 147*(1), 51-61.
- Jha, K., & Iyer, K. (2007). Commitment, coordination, competence and the iron triangle. *International Journal of Project Management, 25*(5), 527-540.
- JICA. (2013). *Study on Public Procurement Systems and Capacity in South Asian Countries Final Report*. Retrieved from [http://open\\_jicareport.jica.go.jp/pdf/12087599.pdf](http://open_jicareport.jica.go.jp/pdf/12087599.pdf)
- JICA. (2017). *JICA Annual Report 2017*. Retrieved from <https://www.jica.go.jp/english/publications/reports/annual/2017/index.html>
- JICA. (2020). *JICA Annual Evaluation Reports*. Retrieved from [https://www.jica.go.jp/english/our\\_work/evaluation/reports/index.html](https://www.jica.go.jp/english/our_work/evaluation/reports/index.html)
- Johnson, & Babu, R. (2018). Time and cost overruns in the UAE construction industry: A critical analysis. *International Journal of Construction Management, 20*(5), 402-411.
- Johnson, Leenders, M., & McCue, C. (2017). A comparison of purchasing's organizational roles and responsibilities in the public and private sector. *Journal of Public Procurement, 3*(1), 57-74.

- Joint Inspections Unit of the United Nations System. (2015). *Contract Management - Procurement Assessment Model (CM-PAM)*. Retrieved from <https://www.unjiu.org/>
- Jørgensen, M. Welde, M. and Halkjelsvik, T. Evaluation of Probabilistic Project Cost Estimates, in *IEEE Transactions on Engineering Management*.1-16
- Jugdev, K., & Müller, R. (2005). A retrospective look at our evolving understanding of project success. *Project Management Journal*, 36(4), 19-31.
- Kaiser, M. J. (2022). Offshore oil and gas records circa 2020. *Ships and Offshore Structures*, 17(1), 205-241.
- Kakwezi, P., & Nyeko, S. (2019). Procurement processes and performance: Efficiency and effectiveness of the procurement function. *International Journal of Social Sciences Management and Entrepreneurship (IJSSME)*, 3(1).
- Kanji, G. K. (2006). *100 statistical tests*. London. Sage Publications.
- Kaplan, R. S. (2009). Conceptual foundations of the balanced scorecard. *Handbooks of management accounting research*, 3, 1253-1269.
- Kaplan, R. S., & Norton, D. P. (2001). Transforming the balanced scorecard from performance measurement to strategic management: Part II. *Accounting Horizons*, 15(2), 147-160.
- Karaca, I., Gransberg, D. D., & Jeong, H. D. (2020). Improving the accuracy of early cost estimates on transportation infrastructure projects. *Journal of Management in Engineering*, 36(5).
- Kaufman, D. (2005). Myths about Governance and Corruption. *Finance and Development*, 42(3), 2005.

- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The worldwide governance indicators: methodology and analytical issues. *Hague Journal on the Rule of Law*, 3(2), 220-246.
- Kearney, A. T. (2016). *What Good Looks Like (ROSMA)*. Retrieved from <https://www.atkearney.com/procurement/rosma/full-report>
- Kellerman, M. (2019). The proliferation of multilateral development banks. *The Review of International Organizations*, 14(1), 107-145.
- Kelvin, W. T. (1883). *Electrical Units of Measurement: Being One of the Series of Lectures Delivered at The Institution of Civil Engineers, Session 1882-83*: Institution of Civil Engineers.
- Kent State University. (2013). *SPSS Tutorials: Perason Correlation*. Retrieved from <https://libguides.library.kent.edu/SPSS/PearsonCorr>
- Keough, M. (1993). *Buying your way to the top*. New Jersey. The McKinsey Quarterly(3), 41-63.
- Kerzner, H. (2002). *Strategic planning for project management using a project management maturity model*. New York. John Wiley & Sons.
- Kerzner, H. (2017). *Project management: Case studies*. New York. John Wiley & Sons.
- Khagram, S., Thomas, C., Lucero, C., & Mathes, S. (2009). Evidence for development effectiveness. *Journal of Development Effectiveness*, 1(3), 247-270.
- Khang, D. B., & Moe, T. L. (2008). Success criteria and factors for international development projects: A life-cycle-based framework. *Project Management Journal*, 39(1), 72-84.
- Kilby, C. (2009). Donor influence in international financial institutions: Deciphering what alignment measures measure. *Paper presented at the Political Economy of International Organizations Meeting, January 2009 (1-29)*. Geneva, Switzerland.

- Kilby, C., & Michaelowa, K. (2019). What influences World Bank project evaluations? In *Lessons on Foreign Aid and Economic Development* (pp. 109-150). Palgrave Macmillan, Cham.
- Kim, M., Suresh, N., & Kocabasoglu-Hillmer, C. (2015). A contextual analysis of the impact of strategic sourcing and E-procurement on performance. *Journal of Business & Industrial Marketing* 30(1), 1-18.
- Kim, S. Y., Tuan, K. N., Lee, J. D., Pham, H., & Luu, V. T. (2018). Cost overrun factor analysis for hospital projects in Vietnam. *KSCE Journal of Civil Engineering*, 22(1), 1-11.
- Klakegg, O. J., & Lichtenberg, S. (2016). Successive cost estimation—successful budgeting of major projects. *Procedia-Social and Behavioral Sciences*, 226, 176-183.
- Knack, S., & Rahman, A. (2007). Donor fragmentation and bureaucratic quality in aid recipients. *Journal of Development Economics*, 83(1), 176-197.
- Knack, S., Rogers, F. H., & Eubank, N. (2011). Aid quality and donor rankings. *World Development*, 39(11), 1907-1917.
- Knight, L., Harland, C., Telgen, J., Thai, K. V., Callender, G., & McKen, K. (2012). *Public procurement: International cases and commentary*. London. Routledge.
- Knudsen, D. (1999). *Procurement performance measurement system*. Sweden. Lund University Publications.
- Kolstad, I., Fritz, V., & O'Neil, T. (2008). *Corruption, anti-corruption efforts and aid: Do donors have the right approach*. Research project (RP-05-GG) of the Advisory Board for Irish Aid: Overseas Development Institute. Retrieved from <https://cdn.odi.org/media/documents/2267.pdf>

- Kono, D. Y., & Montinola, G. R. (2013). The uses and abuses of foreign aid: development aid and military spending. *Political Research Quarterly*, 66(3), 615-629.
- Kovács, G., & Spens, K. M. (2005). Abductive reasoning in logistics research. *International Journal of Physical Distribution & Logistics Management*, 35(2), 132-144.
- KPMG. (2012). *The Power of Procurement. A Global Survey of Procurement Functions*. Retrieved from <https://assets.kpmg/content/dam/kpmg/pdf/2012/07/the-power-of-procurement-a-global-survey-of-procurement-functions.pdf>
- Kraljic, P. (1983). Purchasing must become supply management. *Harvard business review* 61(5), 109-117.
- Kuipers, S. (2021). Rethinking anti-corruption efforts in international development. *Journal of Financial Crime*, ahead-of-print.
- Kumar, A., Ozdamar, L., & Peng Ng, C. (2005). Procurement performance measurement system in the health care industry. *International Journal of Health Care Quality Assurance*, 18(2), 152-166.
- Kwak, Y. H., & Ibbs, C. W. (2002). Project management process maturity (PM) 2 model. *Journal of Management in Engineering*, 18(3), 150-155.
- La Chimia, A., & Trepte, P. (Eds.). (2019). *Public procurement and aid effectiveness: A roadmap under construction*. Bloomsbury Publishing.
- Lahti, M., Shamsuzzoha, A. H. M., & Helo, P. (2009). Developing a maturity model for Supply Chain Management. *International Journal of Logistics Systems and Management*, 5(6), 654-678.

- Lamhauge, N., Lanzi, E., & Agrawala, S. (2012). *Monitoring and evaluation for adaptation: lessons from development co-operation agencies*. OECD Environment Working Papers, No. 38. OECD Publishing.
- Lapide, L. (2000). *What about measuring supply chain performance*. Achieving Supply Chain Excellence Through Technology, 2(2), 287-297. AMR Research Retrieved from <http://lapide.ASCET.com>
- Lardenoije, E. J., van Raaij, E. M., & van Weele, A. J. (2005). Performance management models and purchasing: relevance still lost. *Paper presented at the Researches in Purchasing and Supply Management, the 14th IPSERA Conference, March 2005* (pp. 687-97). Archamps, France.
- LeBaron, G., & Rühmkorf, A. (2017). Steering CSR through home state regulation: A comparison of the impact of the UK bribery act and modern slavery act on global supply chain governance. *Global Policy*, 8, 15-28.
- Lederman, D., Loayza, N. V., & Soares, R. R. (2005). Accountability and corruption: Political institutions matter. *Economics & Politics*, 17(1), 1-35.
- Leung, S.-O. (2011). A comparison of psychometric properties and normality in 4, 5, 6, and 11-point Likert scales. *Journal of Social Service Research*, 37(4), 412-421.
- Liebetrueth, T., Melneck, W., & Pilsl, A. (2016). *Process maturity-assessments in strategic procurement—guideline for developing an advanced methodology* (pp. 273-291) Supply Management Research. Gaber, Wiesbaden. Springer.
- Lim, C., & Mohamed, M. Z. (1999). Criteria of project success: an exploratory re-examination. *International Journal of Project Management*, 17(4), 243-248.
- Lindstrom, D. (2014). *Procurement project management success: achieving a higher level of effectiveness*. Florida, USA. J. Ross Publishing.

- Linstone, H. A., & Turoff, M. (1975). *The delphi method: Techniques and Applications*. Reading, MA, USA. Addison-Wesley.
- Lloyd, R., Poate, D., & Villanger, E. (2014). Results measurement and evaluability: a comparative analysis. *Journal of Development Effectiveness*, 6(4), 378-391.
- Lockamy III, & McCormack. (2004). The development of a supply chain management process maturity model using the concepts of business process orientation. *Supply Chain Management: An International Journal*, 9(4), 272-278.
- Long, N. D., Ogunlana, S., Quang, T., & Lam, K. C. (2004). Large construction projects in developing countries: a case study from Vietnam. *International Journal of Project Management*, 22(7), 553-561.
- Lonsdale, Cordelia. 2016. *Aligning blended finance with the Busan principles of development effectiveness*. Bristol, UK. Development Initiatives.  
<http://devinit.org/post/aligning-blended-finance-to-busan-principles>
- Love, P. E. D., Ahiaga-Dagbui, D. D., Smith, S. D., Sing, M. C. P., & Tokede, O. (2018). Cost Profiling of Water Infrastructure Projects. *Journal of infrastructure systems*, 24(4).
- Lucas, R. E. (2002). The industrial revolution: Past and future. Lectures on economic growth, *American Institute for Economic Research* 44(8), 109-188.
- Luzon, B., & El-Sayegh, S. M. (2016). Evaluating supplier selection criteria for oil and gas projects in the UAE using AHP and Delphi. *International Journal of Construction Management*, 16(2), 175-183.
- Manyathi, S., Burger, A. P., & Moritmer, N. L. (2021). Public sector procurement: A private sector procurement perspective for improved service delivery. *Africa's Public Service Delivery & Performance Review*, 9(1), 11.
- Maren, M. (2009). *The road to hell*. New York. Simon and Schuster.



- Matthew, K., Patrick, K., & Denise, K. (2013). The effects of fraudulent procurement practices on public procurement performance. *International Journal of Business and Behavioral Sciences*, 3(1), 17-27.
- Mauro, P. (1995). Corruption and growth. *The Quarterly Journal of Economics*, 110(3), 681-712.
- Mavroidis, P. C. (2005). *The general agreement on tariffs and trade: a commentary*. Oxford, UK. Oxford University Press.
- McCrudden, C. (2007). *Corporate social responsibility and public procurement*. Oxford Legal Studies Research Paper No. 9/2006: Cambridge University Press.
- McGillivray, M., & Morrissey, O. (2000). Aid fungibility in assessing aid: red herring or true concern? *Journal of International Development*, 12(3), 413-428.
- McKechnie, A. (2017). *Six recommendations for reforming multilateral development banks*. Overseas Development Institute. Retrieved from <https://cdn.odi.org/media/documents/11908.pdf#page=21>
- McKee, C., Blampied, C., Mitchell, I., & Rogerson, A. (2020). *Revisiting Aid Effectiveness: A New Framework and Set of Measures for Assessing Aid Quality*" (Vol. 524). Center for Global Development.
- McLiesh, C., & Arizti, P. (2006). *The doing business project*. MfDR Principles in Action: Sourcebook on Emerging Good Practices, 1, 107-111.
- McMillan, S. S., King, M., & Tully, M. P. (2016). How to use the nominal group and Delphi techniques. *International Journal of Clinical Pharmacy*, 38(3), 655-662.
- McMullan, M. (1961). A theory of corruption. *The Sociological Review*, 9(2), 181-201.
- Meier, W. (2003). *Results based management: towards a common understanding among development co-operation agencies*. Documento de discusión preparado para la Agencia Canadiense para el Desarrollo Internacional, División de Evaluación del

Desempeño, para ser considerado por el Grupo de Trabajo de CAD sobre la Eficacia de la Ayuda y la Armonización, Ottawa, Canadá.

- Mena, C., Christopher, M., & van Hoek, R. (2014). *Leading Procurement Strategy: Driving Value Through the Supply Chain*. London, UK. Kogan Page Publishers.
- Mensah, J. L. A. (2016). The Impact Of Procurement Professionals And Functions In The Strategic Level Of Corporate Organization. *International Journal of Innovative Research Studies (IJIRAS)*, 3(10), 85-89.
- Metzger, L., & Guenther, I. (2015). How to assess the effectiveness of development aid projects: evaluation ratings versus project indicators. *Journal of International Development*, 27(8), 1496-1520.
- Miller. (2009). The privileges and immunities of the United Nations. *International Organizations Law Review*, 6(1), 7-115.
- Miller, & Tsang, E. W. K. (2011). Testing Management Theories: Critical Realist Philosophy and Research Methods. *Strategic Management Journal*, 32(2), 139-158.
- Miliband, D., & Gurumurthy, R. (2015). Improving Humanitarian Aid: how to make relief more efficient and effective. *Foreign Affairs*, 94(4), 118-129.
- Min, H., Thakkar, J., Kanda, A., & Deshmukh, S. (2009). Supply chain performance measurement framework for small and medium scale enterprises. *Benchmarking: An International Journal*, 16(5), 702-723.
- Mir, F. A., & Pinnington, A. H. (2014). Exploring the value of project management: linking project management performance and project success. *International Journal of Project Management*, 32(2), 202-217.

- Misangyi, V. F., Weaver, G. R., & Elms, H. (2008). Ending corruption: The interplay among institutional logics, resources, and institutional entrepreneurs. *Academy of Management Review*, 33(3), 750-770.
- Mitchell, G. E. (2014). Why will we ever learn? Measurement and evaluation in international development NGOs. *Public Performance & Management Review*, 37(4), 605-631.
- Mitchell, V. (2019). A proposed framework and tool for non-economic research impact measurement. *Higher Education Research & Development*, 38(4), 819-832.
- Molino, J. M. (2019). The New Asian Development Bank Procurement policy and regulations: Promoting innovation in public procurement in Asia? In *Joint public procurement and innovation: Lessons across borders*, 27(20), 537-564
- Moradi, S., Kähkönen, K., & Aaltonen, K. (2020). From past to present-The development of project success research. *The Journal of Modern Project Management*, 8(1).
- Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1(1), 48-76.
- Morra-Imas, L. G., & Rist, R. C. (2009). *The road to results: Designing and conducting effective development evaluations*: World Bank Publications.
- Mosley, P. (1986). Aid-effectiveness: The Micro-Macro Paradox. *Institute of Development Studies Bulletin*, 17(2), 22-27.
- Mullen, P. M. (2003). Delphi: myths and reality. *Journal of Health Organization and Management*, 17(1), 37-52.
- Murray, J. (2013). Likert data: what to use, parametric or non-parametric? *International Journal of Business and Social Science*, 4(11).

- Murray, J. G. (2014). Procurement Fraud Vulnerability: A Case Study. *EDPACS*, 49(5), 7-17.
- Muzio, D., Faulconbridge, J., Gabbioneta, C., & Greenwood, R. (2016). Bad apples, bad barrels and bad cellars: a ‘boundaries’ perspective on professional misconduct. *Organizational wrongdoing* (p.141-175). UK. Cambridge University Press.
- Negra, C., Remans, R., Attwood, S., Jones, S., Werneck, F., & Smith, A. (2020). Sustainable agri-food investments require multi-sector co-development of decision tools. *Ecological Indicators*, 110, 105851.
- Netland, T. H., Alfnes, E., & Fauske, H. (2007). How mature is your supply chain? - A supply chain maturity assessment test. *Paper presented at the Proceedings of the 14th International EurOMA Conference Managing Operations in an Expanding Europe, June 2007* (pp. 7-20). Bilkent, Ankara, Turkey.
- Neumayer, E. (2003). The determinants of aid allocation by regional multilateral development banks and United Nations agencies. *International Studies Quarterly*, 47(1), 101-122.
- Nielsen, R. (2010). Does aid follow need? Humanitarian motives in aid allocation. *Paper presented at the AidData Conference, March 2010. Vol 57* (pp. 2). Oxford, UK .
- Noh, J., & Lee, K. M. (2003). Application of multi-attribute decision-making methods for the determination of relative significance factor of impact categories. *Environmental Management*, 31(5), 0633-0641.
- Odeck, J. (2004). Cost overruns in road construction—what are their sizes and determinants? *Transport Policy*, 11(1), 43-53.
- Odette. (2020). Odette International. Retrieved from <https://www.odette.org/>
- Odhiambo, W., & Kamau, P. (2003). *Public Procurement. Working Paper No. 208*, OECD Development Paper. OECD Publishing.

- OECD-DAC. (2012). *Assessing the Development Effectiveness of Multilateral Organizations: Guidance on Methodological Approach*. Retrieved from <http://www.oecd.org/dac/evaluation/dcdndep/50540172.pdf>
- OECD-DAC. (2019). *OECD DAC Evaluation Criteria: Summary of consultation responses. OECD Enabling Development Effectiveness*. Retrieved from <http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>
- OECD-DAC. (2020). *Global Consultation on Evaluation Criteria*. Retrieved from <http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>
- OECD. (2010). *Methodology for Assessing Procurement Systems (MAPS)*. Retrieved from <http://www.oecd.org/governance/ethics/Methodology-Assessment-Procurement-System-Revised-Draft-July-2016.pdf>
- OECD. (2014). *The Rationale for Fighting Corruption*. Retrieved from <https://www.oecd.org/cleangovbiz/49693613.pdf>
- OECD. (2015). *Mapping of Current Initiatives on Public Procurement Assessment*. Retrieved from <https://www.oecd.org/governance/procurement/toolbox/search/mapping-assessments-methods-formatted.pdf>
- OECD. (2017). *Development aid rises again in 2017*. Retrieved from <http://www.oecd.org/dac/financing-sustainable-development/development-finance-data/ODA-2016-detailed-summary.pdf>
- OECD. (2018). *Results in Development Co-operation*. Retrieved from [www.oecd.org › dac › docs › results-pse-results-workshop-apr-18](http://www.oecd.org/dac/docs/results-pse-results-workshop-apr-18)

- OECD. (2019). *DAC Criteria for Evaluating Development Assistance*. Retrieved from <http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>
- OECD. (2020). *Organisation for Economic Co-operation and Development (OECD) Development Data*. Retrieved from <https://stats.oecd.org/Index.aspx>
- OECD. (2021). *COVID 19 spending helped to lift foreign aid to an all-time high in 2020*. Retrieved from <http://www.oecd.org/dac/financing-sustainable-development/development-finance-data/ODA-2020-detailed-summary.pdf>
- OECD Inter Agency Task Force on Financing for Development. (2017). *Official Development Assistance (ODA)*. Retrieved from [http://www.un.org/esa/ffd/wp-content/uploads/2016/01/ODA\\_OECD-FfDO\\_IATF-Issue-Brief.pdf](http://www.un.org/esa/ffd/wp-content/uploads/2016/01/ODA_OECD-FfDO_IATF-Issue-Brief.pdf)
- Okafor, C., & Udibe, K. U. (2021). Foreign Aid Management Challenges in Nigeria (2010-2020). *Journal of the Management Sciences (Jomas)*, 57, 26.
- Olaniran, O. J., Love, P., Edwards, D., Olatunji, O., & Matthews, J. (2017). Chaos theory: implications for cost overrun research in hydrocarbon megaprojects. *Journal of Construction Engineering and Management*, 143(2), 05016020.
- Olaniran, O. J., Love, P. E., Edwards, D., Olatunji, O. A., & Matthews, J. (2015). Cost overruns in hydrocarbon megaprojects: A critical review and implications for research. *Project Management Journal*, 46(6), 126-138.
- Open Data Institute (2021). *Open Data Institute*. Retrieved from <https://theodi.org/>
- Open Knowledge Foundation (2021). *A fair, free and open future*. Retrieved from <https://okfn.org/>
- Pache, G., & Spalanzani, A. (2007). *" La gestion des chaines logistiques multi-acteurs: perspectives stratégiques*. Presses Universitaires de Grenoble, Grenoble. PUG.

- Palagashvili, L. (2019). *Evaluating Aid Agencies: Challenges, Comparisons, and Causes of Best Aid Practices* (pp. 85-107). Lessons on Foreign Aid and Economic Development. Switzerland AG. Springer.
- Pallas, C. L., & Wood, J. (2009). The World Bank's Use of Country Systems for Procurement: A Good Idea Gone Bad? *Development Policy Review*, 27(2), 215-230.
- Papadaki, M., Gale, A., Rimmer, J., Kirkham, R., Taylor, A., & Brown, M. (2014). Essential factors that increase the effectiveness of project/programme risk management. *Procedia-Social and Behavioral Sciences*, 119, 921-930.
- Park, Y. I., & Papadopoulou, T. C. (2012). Causes of cost overruns in transport infrastructure projects in Asia: their significance and relationship with project size. *Built Environment Project and Asset Management* 2(2), 195-216.
- Parker, P. L., McDaniel, H. S., & Crumpton-Young, L. L. (2002). Do research participants give interval or ordinal answers in response to Likert scales? In *IIE Annual Conference. Proceedings* (p. 1). Institute of Industrial and Systems Engineers (IISE).
- Paterson, A. S., Changwony, F., & Miller, P. B. (2019). Accountin Accounting control, governance and anti-corruption initiatives in public sector organisations. *The British Accounting Review*, 51(5), 100844
- Patrucco, A. S., Moretto, A., Ronchi, S., & Luzzini, D. (2019). Organisational choices in public procurement: what can public management learn from the private sector? *Local Government Studies*, 45(6), 977-1000.
- Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative Social Work*, 1(3), 261-283.

- Paulraj, A., Chen, I. J., & Flynn, J. (2006). Levels of strategic purchasing: impact on supply integration and performance. *Journal of Purchasing and Supply Management*, 12(3), 107-122.
- Pennypacker, J. S., & Grant, K. P. (2003). Project management maturity: An industry benchmark. *Project Management Journal*, 34(1), 4-11.
- Phellas, C. N., Bloch, A., & Seale, C. (2011). Structured methods: interviews, questionnaires and observation. *Researching Society and Culture*, 3, 181-205.
- Picciotto, R. (2002). *Development co-operation and performance evaluation: the Monterrey challenge. OED: The First Thirty Years*. WB OED. Washington DC, World Bank.
- Picciotto, R. (2018). *Making development work: Development learning in a world of poverty and wealth*. New York. Routledge.
- Picciotto, R. (2020). Towards a 'New Project Management' movement? An international development perspective. *International Journal of Project Management*, 38(8), 474-485.
- Piketty, T. (2018). *Capital in the 21st Century*. Cambridge, MA and London, England. Harvard University Press.
- Pinto, J., Leana, C. R., & Pil, F. K. (2008). Corrupt organizations or organizations of corrupt individuals? Two types of organization-level corruption. *Academy of Management Review*, 33(3), 685-709.
- Pizzi, S., Caputo, A., Corvino, A., & Venturelli, A. (2020). Management research and the UN sustainable development goals (SDGs): A bibliometric investigation and systematic review. *Journal of Cleaner Production*, 276, 124033.



- Pizzol, M., Laurent, A., Sala, S., Weidema, B., Verones, F., & Koffler, C. (2017). Normalisation and weighting in life cycle assessment: quo vadis? *The International Journal of Life Cycle Assessment*, 22(6), 853-866.
- Pohl, G., & Mihaljek, D. (1992). Project evaluation and uncertainty in practice: A statistical analysis of rate-of-return divergences of 1,015 World Bank projects. *The World Bank Economic Review*, 6(2), 255-277.
- Pollock, A. M., Price, D., & Player, S. (2007). An examination of the UK Treasury's evidence base for cost and time overrun data in UK value-for-money policy and appraisal. *Public Money and Management*, 27(2), 127-134.
- Poluha, R. G. (2007). *Application of the SCOR model in supply chain management*. Youngstown, New York, USA. Cambria Press.
- Pope, A. (1716). *An essay on criticism*: London. Printed for W. Lewis in Russel-Street, p. 1711.
- Potage, J. (2017). Managing procurement value creation with a maturity model. *Logistique & Management*, 25(4), 303-315.
- Powell, B. (2005). State development planning: did it create an east Asian miracle? *The Review of Austrian Economics*, 18(3), 305-323.
- Qualtrics. (2020). Qualtrics. Retrieved from <https://www.qualtrics.com/>
- Quinlan, C., Babin, B., Carr, J., & Griffin, M. (2019). *Business research methods*. Andover, UK. South Western Cengage.
- Quinot, G., & Arrowsmith, S. (2013). *Public procurement regulation in Africa: UK*: Cambridge University Press.
- Quyên, Đ. T. N. (2014). Developing university governance indicators and their weighting system using a modified Delphi method. *Procedia-Social and Behavioral Sciences*, 141, 828-833.

- Raschke, R. L., & Ingraham, L. R. (2010). *Business process maturity's effect on performance*. AMCIS 2010 Proceedings. 402. Lima, Peru. Retrieved from <http://aisel.aisnet.org/amcis2010/402>
- Ravelo, J. L. (2012). *30 percent of aid lost to corruption – Ban Ki-moon*. Retrieved from <https://www.devex.com/news/30-percent-of-aid-lost-to-corruption-ban-ki-moon-78643>
- Raymond, J. (2008). Benchmarking in public procurement. *Benchmarking: An International Journal*, 15(6), 782-793.
- Reck, R. F., & Long, B. G. (1988). Purchasing: a competitive weapon. *Journal of Purchasing and Materials Management*, 24(3), 2-8.
- Reinsberg, B., & Westerwinter, O. (2021). The global governance of international development: Documenting the rise of multi-stakeholder partnerships and identifying underlying theoretical explanations. *The Review of International Organizations*, 16(1), 59-94.
- Renard, R., & Lister, S. (2015). *Measuring and reporting on value for money: A conceptual framework for MDBs*. Oxford, UK. Mokoro Ltd.
- Rendon, R. G. (2008). Procurement process maturity: Key to performance measurement. *Journal of Public Procurement*, 8(2), 200-214.
- Rendon, R. G. (2010). Critical success factors in government contract management. In *Proceedings of the 4th International Public Procurement Conference (IPPC), August 2010* (pp. 1-32). Seoul, South Korea.
- Rendon, R. G. (2015). Benchmarking contract management process maturity: a case study of the US Navy. *Benchmarking: An International Journal* 22(7), 1481-1508.

- Riddell, R. C. (2009). 3 Does Foreign Aid Work? In *Scientific Council for Government Policy (WRR)* (pp. 47-70). Amsterdam. Amsterdam University Press.
- Rimkūnienė, D. (2013). Modern procurement: strategic role and competitive advantage. *Innovative Infotechnologies for Science, Business and Education*, 2(15), 14-18.
- Roberto Aiello. (2016). *IFI Business Opportunities. IDB Financing in Latin America and the Caribbean*. Retrieved from [http://www.cmm.ca/~media/Files/presentations/2016/Seminaire\\_banque/7%20Roberto%20Aiello%20%20Banque%20InterAmricaine%20de%20Dveloppement%20%20Opportunits%20daffaires%20en%20Amrique%20latine%20%20Carabe%20s.pdf](http://www.cmm.ca/~media/Files/presentations/2016/Seminaire_banque/7%20Roberto%20Aiello%20%20Banque%20InterAmricaine%20de%20Dveloppement%20%20Opportunits%20daffaires%20en%20Amrique%20latine%20%20Carabe%20s.pdf)
- Rohde, L., Steen Larsen, T., Jensen, R. L., Larsen, O. K., Jønsson, K. T., & Loukou, E. (2020). Determining indoor environmental criteria weights through expert panels and surveys. *Building Research & Information*, 48(4), 415-428.
- Rohwer, A. (2009). Measuring corruption: a comparison between the transparency international's corruption perceptions index and the World Bank's worldwide governance indicators. *CESifo DICE Report*, Institut für Wirtschaftsforschung an der Universität München, München 7(3), 42-52.
- Roland Berger. (2014). *The CPO Agenda*. Retrieved from [www.rolandberger.com/media/pdf/Roland\\_Berger\\_T\\_he\\_CPO\\_Agenda\\_20140130.pdf](http://www.rolandberger.com/media/pdf/Roland_Berger_T_he_CPO_Agenda_20140130.pdf).
- Rönnbäck, Å. (2012). Quality in the public procurement process. *The TQM Journal*, 24(5), 447-460.
- Roodhooft, F., & Van den Abbeele, A. (2006). Public procurement of consulting services: Evidence and comparison with private companies. *International Journal of Public Sector Management*, 19(5), 490-512.

- Roodman, D. (2006). *An index of donor performance*. Center for Global Development Working Paper 67, Center for Global Development.
- Roodman, D. (2007). *Macro aid effectiveness research: A guide for the perplexed*. Center for Global Development Working Paper 134, Center for Global Development.
- Rowley, J. (2014). Designing and using research questionnaires. *Management Research Review*, 37(3), 308-330.
- Sachs, J. D. (2012). From millennium development goals to sustainable development goals. *The Lancet*, 379(9832), 2206-2211.
- Safari, H., Etezadi, S., Moradi-Moghadam, M., & Fathi, M. R. (2021). Maturity evaluation of supply chain procedures by combining SCOR and PST models. *International Journal of Process Management and Benchmarking*, 11(5), 707-724.
- Samardžija, J., & Kralj, D. (2010). EFQM excellence model 2010 solid framework for introducing environmental innovation. In *Proceedings of the International Conference on Circuits, Systems, Signals, September 2010* (pp. 164-152), Malta.
- Santiso, C. (2001). Good governance and aid effectiveness: The World Bank and conditionality. *The Georgetown Public Policy Review*, 7(1), 1-22.
- Santos, C., Santos, V., Tavares, A., & Varajão, J. (2020). Project management in public health: A systematic literature review on success criteria and factors. *Portuguese Journal of Public Health*, 38(1), 37-48.
- Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research methods for business students* (Fifth Edition ed.). Harlow, England. Pearson.
- Saunders, M., Lewis, P., & Thornhill, A. (2012). *Research methods for business students* (6th Edition ed.). Harlow, England. Pearson.

- Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research methods for business students* (7th Edition ed.). Harlow, England. Pearson.
- Savedoff, W. D., Levine, R., & Birdsall, N. (2005). *When will we ever learn? Recommendations to improve social development through enhanced impact evaluation*. Washington, DC. Center for Global Development, 15.
- Schiele, H. (2007). Supply-management maturity, cost savings and purchasing absorptive capacity: Testing the procurement–performance link. *Journal of Purchasing and Supply Management*, 13(4), 274-293.
- Scholte, J. A., & Söderbaum, F. (2017). A changing global development agenda? In *Forum for Development Studies*, 44(1), 1-12. Retrieved from <http://www.adlittle.no/en/insights/viewpoints/seven-value-drivers-future-procurement>; <https://www.toolshero.com/strategy/adl-matrix/>
- Schuh, C., Kromoser, R., Strohmer, M. F., Pérez, R. R., & Triplat, A. (2009). Using the Purchasing Chessboard™. *The Purchasing Chessboard*, (pp. 35-53). New York, USA. Springer.
- Schuh, C., Raudabaugh, J. L., Kromoser, R., Strohmer, M. F., Triplat, A., & Pearce, J. (2017). The Way Forward *The Purchasing Chessboard* (pp. 207-226). New York, USA. Springer.
- Schweiger, J. (2014). A theory-based perspective on Maturity Models in Purchasing and Supply Management. In *Innovative Methods in Logistics and Supply Chain Management: Current Issues and Emerging Practices. Proceedings of the Hamburg International Conference of Logistics (HICL), 2014 Vol 18* (pp. 531-554). Berlin. epubli GmbH.

- Schweiger, J. (2015). Development of a Purchasing and Supply Management Maturity Framework. *Operations and Supply Chain Management: An International Journal*, 8(1), 11-21.
- Seiler, N., & Madir, J. (2012). Fight against corruption: Sanctions regimes of multilateral development banks. *Journal of International Economic Law*, 15(1), 5-28.
- Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill building approach*. New York, USA. John Wiley & Sons.
- Serrat, O. (2017). Political economy analysis for development effectiveness. In *Knowledge solutions* (pp. 207-222). Springer, Singapore.
- Sharp, S., Valters, C., & Whitty, B. (2019). *How DFID can better manage complexity in development programming*. Overseas Development Institute. Retrieved from <https://www.odi.org/publications/11315-how-dfid-canbetter-manage-complexity-development-programming>.
- Shokri-Ghasabeh, M., & Kavouusi-Chabok, K. (2009). Generic project success and project management success criteria and factors: Literature review and survey. *World Scientific and Engineering Academy and Society* 8(6), 456-468.
- Shrnhur, A. J., Levy, O., & Dvir, D. (1997). Mapping the dimensions of project success. *Project Management Journal*, 28(2), 5-13.
- Soares, S. R., Toffoletto, L., & Deschênes, L. (2006). Development of weighting factors in the context of LCIA. *Journal of Cleaner Production*, 14(6-7), 649-660.
- Stapleton, D., Hanna, J. B., Yagla, S., Johnson, J., & Markussen, D. (2002). Measuring logistics performance using the strategic profit model. *The International Journal of Logistics Management*, 13(1), 89-107.
- Stern, N. (2003). Scaling up: The Challenge of Monterrey. In *Towards Pro-Poor Policies, Aid, Institutions and Globalisation. 5<sup>th</sup> Annual World Bank Conference*

- on Development Economics, May 2003* (pp. 14-42) Paris, France: Co-publication of World Bank and Oxford University Press.
- Stijns, J.-P. (2012). Can we still Achieve the Millennium Development Goals? *From Costs to Policies*. Paris, France. OECD Publishing.
- Stilger, P. S., Siderius, J., & Van Raaij, E. M. (2017). A comparative study of formulas for choosing the economically most advantageous tender. *Journal of public procurement*, 17(1), 89-125
- Sumner. (2010). *Global Poverty and the New Bottom Billion: What if Three-Quarters of the World's Poor Live in Middle-income Countries?* Institute of Development Studies Working Paper 349, Brighton. IDS.
- Sumner, A., & Tribe, M. A. (2008). International development studies: Theories and methods in research and practice. *International Politik* 66(4),687-690.
- Sundqvist, E., Backlund, F., & Chronéer, D. (2014). What is project efficiency and effectiveness? *Procedia-Social and Behavioral Sciences*, 119, 278-287.
- Suraraksa, J., & Shin, K. S. (2019). Comparative analysis of factors for supplier selection and monitoring: The case of the automotive industry in Thailand. *Sustainability*, 11(4), 981.
- Survey Monkey. (2020). Survey Monkey. Retrieved from <https://www.surveymonkey.com/>
- Suryahadi, A., Hadiwidjaja, G., & Sumarto, S. (2012). Economic growth and poverty reduction in Indonesia before and after the Asian financial crisis. *Bulletin of Indonesian Economic Studies*, 48(2), 209-227.
- Tadelis, S. (2012). Public procurement design: Lessons from the private sector. *International Journal of Industrial Organization*, 30(3), 297-302.

- Tassabehji, R., & Moorhouse, A. (2008). The changing role of procurement: Developing professional effectiveness. *Journal of Purchasing and Supply Management*, 14(1), 55-68.
- Tayefeh Hashemi, S., Ebadati, O. M., & Kaur, H. (2020). Cost estimation and prediction in construction projects: a systematic review on machine learning techniques. *SN Applied Sciences*, 2(10), 1-27.
- Tierney, M. J., Nielson, D. L., Hawkins, D. G., Roberts, J. T., Findley, M. G., Powers, R. M., . . . Hicks, R. L. (2011). More dollars than sense: Refining our knowledge of development finance using AidData. *World Development*, 39(11), 1891-1906.
- Toutouchian, S., Abbaspour, M., Dana, T., & Abedi, Z. (2018). Design of a safety cost estimation parametric model in oil and gas engineering, procurement and construction contracts. *Safety science*, 106, 35-46.
- Transparency International. (2004). *Global Corruption Report 2004: Political Corruption*. Retrieved from <https://www.transparency.org/en/publications/global-corruption-report-2004-political-corruption>
- Transparency International. (2021). *Corruption perceptions index 2019*. Retrieved from [https://images.transparencycdn.org/images/CPI2021\\_Mapindex\\_EN.pdf](https://images.transparencycdn.org/images/CPI2021_Mapindex_EN.pdf)
- Trigo, A., & Varajão, J. (2020, July). IT project management critical success factors. In *International Conference on Computational Science and Its Applications* (pp. 714-724). Cham, Springer.
- Uitto, J. (2021). *Evaluating environment in international development*. London, UK. Taylor & Francis.
- UNCTAD. (2020). *Public procurement is a powerful tool for sustainable development – UN report*. Retrieved from <https://unctad.org/news/public-procurement-powerful-tool-sustainable-development-un-report>



- United Nations. (2002). *Monterrey Consensus on Financing for Development*.  
Retrieved from <http://www.un.org/esa/ffd/monterrey/MonterreyConsensus.pdf>
- United Nations. (2015). *Sustainable Development Goals*. Retrieved from  
<https://www.un.org/sustainabledevelopment/sustainable-development-goals/>
- United Nations. (2019). *About the United Nations*. Retrieved from  
<http://www.un.org/en/sections/about-un/funds-programmes-specialized-agencies-and-others/index.html>
- United Nations Development Program. (2010). *UNDP Procurement Capacity Assessment Guide*. Retrieved from  
<http://www.undp.org/content/undp/en/home/procurement.html>
- United Nations Development Program. (2019). *UNDP Evaluation Guidelines*.  
Retrieved from <http://web.undp.org/evaluation/guideline/index.shtml>
- United Nations Development Program. (2020). *UNDP Evaluation Resource Centre*.  
Retrieved from <https://erc.undp.org/>
- United Nations International Children's Emergency Fund. (2018). *UNICEF Procurement Services*. Retrieved from  
[https://www.unicef.org/supply/index\\_procurement\\_services.html](https://www.unicef.org/supply/index_procurement_services.html)
- United Nations Office for Project Services. (2018). *Procurement Efficiency Assessment Tool (PEAT)*. Retrieved from <https://www.unops.org/business-opportunities/how-we-procure>
- United Nations Office for Project Services. (2020). *United Nations Office for Project Services (UNOPS) Annual Statistical Report 2020*. Retrieved from  
<https://www.ungm.org/public/asr>

- United Nations Statistics Division (UNSD). (2022). *Standard Country or Area Codes for Statistical Use (M49)*. Retrieved from <https://unstats.un.org/unsd/methodology/m49/>
- United States Institute of Peace. (2018). *International Organisations*. Retrieved from <https://www.usip.org/publications/international-organizations>
- USAID. (2020). *Evaluations at the DEC*. Retrieved from <https://dec.usaid.gov/dec/content/evaluations.aspx>; <https://explorer.usaid.gov/>
- Van Looy, A., De Backer, M., Poels, G., & Snoeck, M. (2013). Choosing the right business process maturity model. *Information & Management*, 50(7), 466-488.
- Van Maanen, J., Sørensen, J. B., & Mitchell, T. R. (2007). The interplay between theory and method. *Academy of Management Review*, 32(4), 1145-1154.
- Van Poucke et al. (2016). Enhancing cost savings through early involvement of purchasing professionals in sourcing projects: Bayesian estimation of a structural equation model. *Journal of Purchasing and Supply Management*, 22(4), 299-310.
- Van Roy, E. (1970). On the theory of corruption. *Economic Development and Cultural Change*, 19(1), 86-110.
- Van Toan, N., Hà, T. T., & Chau, N. N. (2012). From theory to application: using the Economic Internal Rate of Return (EIRR) for assessing the profitability efficiency of the rural infrastructure investments in huong khe district, ha tinh province. *Journal of Science*, Hue University, 70(1), 255-263.
- Van Weele, A., Rozemeijer, F., & Rietveld, G. (1998). Professionalizing purchasing organization: towards a purchasing development model. *Paper presented at the Proceedings 7th IPSERA Conference, April 1998* (pp. 687-97). London, Great Britain.

- Van Weele, A. J. (2010). Value creation and purchasing strategy. In *International trade forum* (4), 34. International Trade Centre.
- Verbrugge, L., de Hoop, L., Aukema, R., Beringen, R., Creemers, R., van Duinen, G., . . . Spikmans, F. (2019). Lessons learned from rapid environmental risk assessments for prioritization of alien species using expert panels. *Journal of Environmental Management*, 249, 109405.
- Verhulst, S. G., & Young, A. (2017). *Open data in developing economies: Toward building an evidence base on what works and how*. Cape Town, SA. African Minds.
- Voegele, A. R., & Schwientek, R. (2002). Purchasing EmPowerment—Bestleistungen im Einkauf *Handbuch Industrielles Beschaffungsmanagement* (pp. 301-318). Wiesbaden. Springer.
- Vrchota, J., Řehoř, P., Maříková, M., & Pech, M. (2020). Critical success factors of the project management in relation to industry 4.0 for sustainability of projects. *Sustainability*, 13(1), 281.
- Vu, T. Q., Pham, C. P., Nguyen, T. A., Nguyen, P. T., Phan, P. T., & Nguyen, Q. L. H. T. T. (2020). Factors Influencing Cost Overruns in Construction Projects of International Contractors in Vietnam. *The Journal of Asian Finance, Economics and Business*, 7(9), 389–400.
- Walker, D., & Rowlinson, S. (2007). *Procurement systems: a cross-industry project management perspective*. Oxon, UK. Routledge.
- Walliman, N. (2017). *Research methods: The basics*. Abingdon, Oxon, UK. Routledge.
- Watt, D., Kayis, B., & Willey, K. (2010). The relative importance of tender evaluation and contractor selection criteria. *International Journal of Project Management*, 28(1), 51-60.

- Wei, S.-J. (1999). *Corruption in economic development: Beneficial grease, minor annoyance, or major obstacle?* World Bank Policy Research Working Paper 2048, USA. Harvard University and National Bureau of Economic Research.
- Welde, M., & Odeck, J. (2017). Cost escalations in the front-end of projects—empirical evidence from Norwegian road projects. *Transport Reviews*, 37(5), 612-630.
- Wendler, R. (2012). The maturity of maturity model research: A systematic mapping study. *Information and Software Technology*, 54(12), 1317-1339.
- Westerveld, E. (2003). The Project Excellence Model®: linking success criteria and critical success factors. *International Journal of Project Management*, 21(6), 411-418.
- White, H. (2005). *Challenges in evaluating development effectiveness*. Working paper series, 242. Brighton: IDS. Retrieved from <https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/4107>.
- Williams, D. A., & Isaksen, J. (2016). *Corruption and state-backed debts in Mozambique: What can external actors do?* U4 Issue 6. Anti-Corruption Resource Centre. Retrieved from [www.U4.no](http://www.U4.no).
- Wood, B., Betts, J., Etta, F., Gayfer, J., Kabell, D., Ngwira, N., . . . Samaranayake, M. (2011). *The Evaluation of the Paris Declaration, Phase 2*. Final Report. Copenhagen. Danish Institute for International Studies.
- Wood, B., Kabell, D., Muwanga, N. K., & Sagasti, F. R. (2008). *Paris Declaration on Aid Effectiveness. Evaluation of the Implementation of the Paris Declaration: phase one: Synthesis report: Joint Progress Towards Aid Effectiveness*. High Level Forum. Paris.

- Woods, N., & Narlikar, A. (2001). Governance and the Limits of Accountability: The WTO, the IMF, and the World Bank. *International Social Science Journal*, 53(170), 569-583.
- World Bank. (2009). *Multilateral Development Banks' Common Performance Assessment System 2009 COMPAS REPORT*. Retrieved from [http://siteresources.worldbank.org/INTDOII/Resources/INT\\_inside\\_fraud\\_text\\_090909.pdf](http://siteresources.worldbank.org/INTDOII/Resources/INT_inside_fraud_text_090909.pdf)
- World Bank Group. (2010). *Corruption Hunters Rally for Action Against Fraud*. Retrieved from <https://www.worldbank.org/en/news/feature/2010/12/06/corruption-hunters-rally-for-action-against-fraud>
- World Bank Group. (2012). *The World Bank and Public Procurement. An Independent Evaluation*. Approach Paper. Washington, DC: World Bank. Retrieved from <http://documents.worldbank.org/curated/en/702461468331906897/pdf/NonAsciiFileName0.pdf>
- World Bank Group. (2013). *The Bank's procurement policies and procedures: Initiating discussion and approach papers*. Washington, DC: World Bank Group. Retrieved from <http://documents.worldbank.org/curated/en/413861468182330960/The-World-Banks-procurement-policies-and-procedures-proposed-review-initiating-discussion-paper>
- World Bank Group. (2014). *Fraud and Corruption Handbook*. Retrieved from [http://siteresources.worldbank.org/EXTDOII/Resources/F&C\\_Awareness\\_Handbook\\_Civil\\_Servants\\_Procurement.pdf](http://siteresources.worldbank.org/EXTDOII/Resources/F&C_Awareness_Handbook_Civil_Servants_Procurement.pdf)

- World Bank Group. (2015). *Independent Evaluation Group World Bank Project Performance Ratings*. Retrieved from [https://ieg.worldbankgroup.org/sites/default/files/Data/reports/ieg-wb-project-performance-ratings-codebook\\_092015.pdf](https://ieg.worldbankgroup.org/sites/default/files/Data/reports/ieg-wb-project-performance-ratings-codebook_092015.pdf)
- World Bank Group. (2016). *Behind the Mirror: A report on the self evaluation systems of the World Bank Group*. Retrieved from <http://documents.worldbank.org/curated/en/902331469736885125/pdf/107274-WP-REVISED-PUBLIC.pdf>
- World Bank Group. (2016). *New Procurement Framework and Regulations for Projects After July 1, 2016*. Retrieved from <http://www.worldbank.org/en/projects-operations/products-and-services/brief/procurement-new-framework>
- World Bank Group. (2019). *World Bank Group Finances Major Contract Awards*. Retrieved from <https://finances.worldbank.org/Procurement/Major-Contract-Awards/kdui-wcs3/data>
- World Bank Group. (2020). *Annual Report*. Retrieved from <http://www.worldbank.org/en/about/annual-report/fiscal-year-data>
- World Bank Group. (2020). *Results and Performance of the World Bank Group 2020*. Retrieved from <https://ieg.worldbankgroup.org/sites/default/files/Data/Evaluation/files/RAP2020.pdf>
- World Bank Group. (2020). *Who we are*. Retrieved from <http://www.worldbank.org/en/where-we-work>
- World Bank Group. (2021). *Evaluation Database*. Retrieved from <https://ieg.worldbankgroup.org/ieg-search>

World Bank Group. (2022). *Fraud and Corruption Handbook*. Retrieved from [http://siteresources.worldbank.org/INTDOII/Resources/INT\\_inside\\_fraud\\_text\\_090909.pdf](http://siteresources.worldbank.org/INTDOII/Resources/INT_inside_fraud_text_090909.pdf)

Yonehara, A., Saito, O., Hayashi, K., Nagao, M., Yanagisawa, R., & Matsuyama, K. (2017). The role of evaluation in achieving the SDGs. *Sustainability Science*, *12*(6), 969-973.

Yuan, J., Zeng, A. Y., Skibniewski, M. J., & Li, Q. (2009). Selection of performance objectives and key performance indicators in public–private partnership projects to achieve value for money. *Construction Management and Economics*, *27*(3), 253-270.

Yves-Adou, J. (2016). *Applications and limitations of value for money in development activities* (pp. 16-20). Abidjan, Ivory Coast: African Development Bank.  
Retrieved from Evaluation Matters Q3\_EN.pdf (AFDB.org)

Zidane, Y.J.-T. and Olsson, N.O.E. (2017), Defining project efficiency, effectiveness and efficacy, *International Journal of Managing Projects in Business*, *10*(3), 621-641.

Zikmund, W. G., Babin, B. J., Carr, J. C., & Griffin, M. (2013). *Business research methods*. Southwestern, Cincinnati, OH. Cengage Learning.

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.

## APPENDIX 1. IPMM QUESTIONNAIRE

### **Integrated Procurement Maturity Assessment Model (IPMM)**

A1 Employer organization

A2 Country of organisation head quarters

A3 Country of primary/base work location (participant completing survey)

A4 Procurement group name (participant completing survey)

A5 Procurement group description (participant completing survey)

A6 Procurement group function (participant completing survey)

A7 Total number of employees in the organisation

A8 Number of employees in procurement group

A9 Organisation currency

A10 Total annual disbursement/expenditure of your organisation  
(Loans/Grants/Technical Assistance)

A11 Total annual procurement expenditure managed or reviewed by procurement group

A12 Total annual operating and administrative cost of procurement group

#### **Questions:**

#### **PHASE 1: Procurement Strategy**

Q1 Select the best description of the Procurement Planning and Forecasting Procedure of your organisation (and required/expected of Implementation Partners (IPs))

- No annual procurement planning and forecasting procedure (1)\*
- Limited/partial annual procurement planning and forecasting procedure (2)
- Annual procurement planning and forecasting procedure is fully established (3)
- Annual procurement planning and forecasting procedure is integrated with strategic business objectives (4)



Q2 Select the best description of the Procurement Planning and Forecasting Process of your organisation (and required/expected of IPs)

	None (1)	Reactive (2)	Annual/Quarterly (3)	Continually Updated (4)
Procurement planning process (schedule)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q3 Select the best description of the Specifications and Scope inputs to the procurement process of your organisation (and required/expected of IPs)

- No specifications and scope inputs to procurement process (1)
- Limited specifications and scope inputs to procurement process (2)
- Detailed specifications and scope inputs to procurement process (3)
- Specifications and scope inputs to procurement process are strategic and innovative (4)

Q4 Select the best description of the Financial Inputs to the procurement process of your organisation (and required/expected of IPs)

	Basic (1)	Limited (2)	Established (3)	Validated/Reliable (4)
Budget provision to procurement is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cost estimation for procurement is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Procurement savings capture process is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q5 Select the alignment of the Procurement Strategy to the business development objectives of your organisation (and required/expected of IPs) *Select all that apply*

- Procurement strategy is independent from business objectives (1)
- Procurement strategy is aligned to business objectives (1)
- Procurement strategy includes justification of selection of the procurement methods (1)
- Procurement strategy includes performance indicators and delivery metrics (1)

Q6 Select the approach to Market Intelligence (MI) in your organisation (and required/expected of IPs)

- MI is not included in the procurement procedures (1)
- MI is limited in the procurement procedures (2)
- MI is established in the procurement procedures (3)
- MI is integrated and aligned with strategic business objectives (4)

Q7 Select the best description of the Supplier Registration and Eligibility in your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Strategic (4)
Supplier registration and eligibility procedures are:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Supplier registration and eligibility system/database:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q8 Select the best description of the Supplier Qualification Procedures of your organisation (and required/expected of IPs)

- No supplier qualification procedures and guidelines (1)
- Limited/basic supplier qualification procedures and guidelines (2)
- Comprehensive/established supplier qualification procedures and guidelines (3)
- Supplier qualification procedures and guidelines include multiple risk criteria (e.g. Finance, Health, Safety and Environment (HSE), Corporate Social Responsibility (CSR) and Local Business Development (LBD)) (4)

Q9 Select the approach to Health, Safety and Environment (HSE) in the procurement process of your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Strategic/Integrated (4)
HSE in the procurement procedures is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
HSE criteria in supplier qualification assessment is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q10 Select the approach to Corporate Social Responsibility (CSR) in the procurement process of your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Strategic/Integrated (4)
CSR in the procurement procedures is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
CSR criteria in supplier qualification assessment is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q11 Select the approach to Local Business Development (LBD) in the procurement process of your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Strategic/Integrated (4)
LBD in the procurement procedures is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
LBD criteria in supplier qualification assessment is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**PHASE 2: Procurement Organisation**

Q12 What is the best description of the Procurement Organisational Structure of your organisation (and required/expected of IPs)

- Fully Centralised (1)
- Fully Decentralised (2)
- Hybrid of centralised and decentralised (3)
- Other (please describe) (4)

Q13 The Procurement Group reports to which level in your organisation (and required/expected of IPs)

- Top Leadership Level (e.g. CEO, President, Vice President, MD) (1)
- Senior Management Level (e.g. Director, Chief/Head of Department) (2)
- Middle Management Level (e.g. Manager, Team Leader) (3)
- Other (please describe) (4)

Q14 Select the best description of the Procurement Resources and Support in your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Strategic/Mandatory (4)
Senior staff qualifications required (e.g. CIPS, ISM, Masters, Bachelors)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Procurement support functions (e.g. administration, spend analytics, eProcurement)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q15 Select the level of understanding and awareness of the Procurement Roles and Responsibilities across your organisation (and required/expected of IPs)

0 (1); 1 (1); 2 (2); 3 (2); 4 (2); 5 (3); 6 (3); 7 (3); 8 (4); 9 (4); 10 (4)

Q16 Select the best description of Procurement and Customer Collaboration in your organization (and required/expected of IPs)

	None (1)	Limited (2)	Regular/Routine (3)	Strategic (4)
Collaboration (Pre-Contract Award)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Collaboration (Post-Contract Award)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q17 Select the description of Category Management Implementation in your organisation (and required/expected of IPs) *Select all that apply*

- Spend analysis reports are displayed by category (1)
- Category management is fully implemented (1)
- Procurement organisation structure is organised by category (1)
- Category management plans are regularly reported to senior management (1)

Q18 Select the best description of Procurement Spend Analysis in your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Strategic (4)
Procurement spend analysis reporting is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Procurement spend analysis technology tools are:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q19 Select the best description of the Procurement Systems and Technology in your organisation (and required/expected of IPs)

- No procurement systems and technology (1)
- Procurement systems have limited procurement functionality (2)
- Procurement process is managed with automated systems and technology (3)
- Procurement process utilizes innovative procurement systems and technology (e.g. cloud technology) (4)

Q20 Select the level of automation of the Procurement Management System across your organisation (and required/expected of IPs)

0 (1); 1 (1); 2 (2); 3 (2); 4 (2); 5 (3); 6 (3); 7 (3); 8 (4); 9 (4); 10 (4)

Q21 Select the implementation of eProcurement (e.g. eTendering, eCatalogs and eAuctions) in your organisation (and required/expected of IPs)

- No eProcurement systems and tools (1)
- Ad Hoc utilisation of eProcurement systems and tools (e.g. eAuctions) (2)
- Regular utilisation of multiple eProcurement systems and tools (e.g. eCatalog/eTendering) (3)
- Procurement process utilizes innovative eProcurement systems and tools (e.g. cloud technology) (4)

Q22 Select the best description of the implementation of Procure-to-Pay Systems in your organisation (and required/expected of IPs)

0 (1); 1 (1); 2 (2); 3 (2); 4 (2); 5 (3); 6 (3); 7 (3); 8 (4); 9 (4); 10 (4)

### **PHASE 3: Procurement Regulatory Framework**

Q23 Select the best description of the Procurement Regulatory Framework of your organisation (and required/expected of IPs)

- No Procurement policy and procedures (1)
- Limited/Partial procurement policy and procedures (2)
- Comprehensive/established procurement policy, procedures and guidelines (3)
- Strategic integration of procurement policy, procedures with the key principles (integrity, fairness, transparency and Value for Money (VFM)) (4)

Q24 Select the best description of Contract General Terms and Conditions Templates of your organisation (and required/expected of IPs)

- No contract general terms and conditions templates (1)
- Limited/partial contract general terms and conditions templates (2)
- Consistent/established contract general terms and conditions templates (3)
- General and specialised contract terms and conditions templates are optimised (fit-for-purpose) (4)

Q25 Select the best description of the Tender Evaluation Procedures (TEP) of your organisation (and required/expected of IPs)

- There is no procurement TEP (1)
- There is limited and informal procurement TEP (2)
- Procurement TEP is fully established (3)
- Procurement TEP supports quality and Value for Money (VFM) principles (4)

Q26 Select the best description of the Procurement Methods used in your organisation (and required/expected of IPs)

*Select all that apply*

- Procurement methods promote competitive practices (1)
- There are multiple procurement methods (e.g. Quality/cost based and competitive negotiation) (1)
- Procurement methods are optimised to support business strategy and Value for Money (VFM) (1)
- There are alternative procurement arrangements and methods (e.g. Framework contracts, Best and Final Offer - BAFO, Public Private Partnerships - PPPs) (1)

Q27 Select the best description of procurement Key Performance Indicators (KPIs) and Benchmarking of your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Strategic/Integrated (4)
Procurement compliance and controls process is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Procurement compliance and controls metrics are:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q28 Select the best description of the procurement Compliance and Controls of your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Strategic (4)
Procurement KPIs and metrics are: (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Procurement process benchmarking is: (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Q29 Select the best description of the procurement Delegation of Authority and Approvals of your organisation (and required/expected of IPs)

- No formal procurement delegation of authority and approvals procedure (1)
- Limited/partial procurement delegation of authority and approvals procedure (2)
- Established procurement delegation of authority and approvals procedure and 'manual' (3)
- Delegation of authority and approvals are automated and integrated across the organisation (4)

Q30 Select the best description of Procurement Committees in your organisation (and required/expected of IPs)

*Select all that apply*

- There are no Procurement Committees in my organisation (0)
- Procurement Committees review compliance to procurement procedures (1)
- Procurement Committees include cross functional representation (quorum) (1)
- Procurement Committees provide strategic procurement advice and oversight (1)

Q31 Select the best description of the procurement Data and Information Disclosure Procedure of your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Compliant to OPD, OKF, OD4D (4)
Data and information disclosure procedure is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Public disclosure of data and information is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q32 Select the best description of the procurement Integrity, Corruption, Fraud and Debarment Policy of your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Strategic/Mandatory (4)
Integrity, corruption and fraud policy is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Supplier debarment policy is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q33 Select the best description of procurement Complaints, Disputes and Arbitration Procedures of your organisation (and required/expected of IPs)

- No procurement complaints, disputes and arbitration procedures (1)
- Limited procurement complaints, disputes and arbitration procedures (2)
- Comprehensive/established procurement complaints, disputes and arbitration procedures (3)
- Procurement complaints, disputes and arbitration process is integrated with legal advice and support (4)

Q34 Select the best description of Contract Risk Management Procedures of your organisation (and required/expected of IPs)

- No contract risk management process and procedures (1)
- Limited contract risk management process and procedures (2)
- Comprehensive/established contract risk management process and procedures (3)
- Contract risk management process and procedures are integrated with strategic business objectives (4)

Q35 Select the best description of the Supplier Risk Management Procedures of your organisation (and required/expected of IPs)

- No supplier risk management process and procedures (1)
- Limited supplier risk management process and procedures (2)
- Comprehensive/established supplier risk management process and procedures (3)
- Supplier risk management process and procedures are integrated with strategic business objectives (4)

Q36 Select the best description of the Contract Amendments and Variations Management process of your organisation (and required/expected of IPs)

*Select all that apply*

- Management of contract amendments and variations are regularly monitored (1)
- Contract amendments and variations management procedures are fully established (1)
- Contract amendments and variations management includes review of contract fundamentals (e.g. contract value and term) (1)
- Contract amendments and variations management includes review of key procurement principles (e.g. contract performance, compliance and Value for Money (VFM)) (1)

Q37 Select the Contract Management Performance measurement process of your organisation (and required/expected of IPs)

0 (1); 1 (1); 2 (2); 3 (2); 4 (2); 5 (3); 6 (3); 7 (3); 8 (4); 9 (4); 10 (4)

Q38 Select the Supplier Management Performance measurement process of your organisation (and required/expected of IPs)

0 (1); 1 (1); 2 (2); 3 (2); 4 (2); 5 (3); 6 (3); 7 (3); 8 (4); 9 (4); 10 (4)

**PHASE 4: Contract Management**

Q39 Select the best description of the procurement Invoice Management Procedures of your organisation (and required/expected of IPs)

- No procurement invoice management procedures (1)
- Limited/partial procurement invoice management procedures (2)
- Consistent/established procurement invoice management procedures (3)
- Procurement invoice management is integrated with strategic key performance indicators (KPIs) (4)

Q40 Select the best description of the procurement Invoice Management System of your organisation (and required/expected of IPs)

- No procurement invoice management system technology (1)
- Procurement invoice management system is integrated with manual invoice management process (2)
- Fully automated electronic procurement invoice management system across the organisation (3)
- Procurement invoice management system includes innovative technology (e.g. Cloud Technology and Optical Scanning Technology (OCR)) (4)

Q41 Select the best description procurement Document Management Procedures of your organisation (and required/expected of IPs)

	None (1)	Limited (2)	Established (3)	Consistent/Compliant (4)
Document management procedure is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Document security/risk classification process is:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q42 Select the best description of the procurement Document Management System (DMS) of your organisation (and required/expected of IPs)

*Select all that apply*

- Procurement DMS is managed by dedicated procurement staff and resources (1)
- Procurement DMS is used as the official archive of confidential documents (1)
- Procurement DMS is integrated and automated across the organisation (1)
- Procurement DMS includes innovative technology (e.g. scanning technology, eSystems, cloud technology) (1)

Q43 Select the Procurement Training Programme format required for procurement staff in your organisation (and recommended of IPs)

*Select all that apply*

- Procurement training programme includes core procurement skills content (1)
- Specialist training programmes are in place for specific skills (e.g. Negotiation, Spend Analysis) (1)
- Regular scheduled procurement training programmes are in place for procurement staff (1)
- Comprehensive procurement training programme includes multiple levels, modules and competencies (1)

Q44 Select the best description of the Professional Procurement Accreditation and Certification required for procurement staff in your organisation (and recommended of IPs)

- No professional procurement accreditation and certification (1)
- Professional procurement accreditation and certification is provided but not mandatory for staff (2)
- Professional procurement accreditation and certification is a requirement for the recruitment of staff (3)
- Professional procurement accreditation and certification is a requirement for the recruitment and career progression of staff (4)

Q45 Select the best description of the Contract Management Training Programme required of your organisation (and recommended for IPs)

- No contract management training programme (1)
- Ad Hoc/limited contract management training programme (2)
- Regular/established contract management training programme (3)
- Contract management training is integrated to staff continuous improvement programmes (4)

Note: \* The numbers in brackets indicate the score allocated for each question

## **APPENDIX 2. QUESTIONS FOR PROCUREMENT EXPERT PANEL ON VFM**

This study aims to examine the effectiveness and efficiency of money spent on development aid by international development organisations through procurement considering that over half of ODA spent on goods and services through the procurement process. The procurement process represents an important channel through which a significant proportion of development aid is spent and hence procurement performance can be an indicator of Development Effectiveness (DE) and Value for Money (VFM).

The research and analysis is intended to demonstrate the impact of procurement performance on DE and VFM and lead to further research on procurement performance and the evaluation of development effectiveness.

Kindly complete the questionnaire independently and submit upon completion guided by the instructions and requested timelines indicated in the accompanying invitation message.

Taking part in a research project is voluntary. It is your choice to take part and you do not have to agree to respond if you do not feel comfortable.

Your responses will be treated in strict confidence and only the research team will have access to the completed questionnaires and individual responses.

Curtin University Human Research Ethics Committee (HREC) has approved this study (HREC number 06/2018). Should you wish to discuss the study with someone not directly involved, in particular, any matters concerning the conduct of the study or your rights as a participant, or you wish to make a confidential complaint, you may contact the Ethics Officer on (08) 9266 9223 or the Manager, Research Integrity on (08) 9266 7093 or email [hrec@curtin.edu.au](mailto:hrec@curtin.edu.au).

## **Requirements**

In order to complete the requirements do the following:

Step 1: Read the definition of Value for Money (VFM) in **Tab 1 - Definition of Value for Money (VFM) Asian Development Bank** (ADB).

Step 2: Enter the Percentage (%) Weight Criteria (in participant's opinion) in **Tab 2 Column J** - that best represents relative influence of VFM for each criteria and then fill the sub-criteria. The Percentage (%) Weight reflects the comparative influence the specific criteria has on total VFM following the ADB Definition.

**Total weight for all criteria should total**

**100%.**



POLICY	CONTRACT PHASE	KEY CRITERIA	SUB-CRITERIA (TASKS)	Combined Percentage % (VFM)
<p>INTEGRITY, ETHICS &amp; ACCOUNTABILITY</p> <p>PROCUREMENT POLICY</p> <p>BINDING AUTHORITY</p> <p>TRANSPARENCY, FAIRNESS, COMPETITION, BEST VALUE</p>	STRATEGY & PLANNING	PLANNING & FORECASTING	Planning and Forecasting	
			Definition, Standards, Specifications, Scope and Budget Estimation	
			Procurement Strategy and Business Alignment	
			<b>Percentage (%) Weight for Procurement Strategy Criteria</b>	17.79
		SUPPLIER QUALIFICATION & ELIGIBILITY	Market Intelligence	
			Supplier Registration and Database Management	
	Supplier Qualification, Capacity and Capability Assessment			
	<b>Percentage (%) Weight for Supplier Qualification &amp; Eligibility Criteria</b>	9.92		
	SUSTAINABLE PROCUREMENT	Health, Safety & Environment (HSE)		
		Corporate Social Responsibility (CSR)		
		Local Business Development (LBD)		
	<b>Percentage (%) Weight for Sustainable Procurement Criteria</b>	5.46		
	ORGANISATION & RESOURCES	ORGANISATION STRUCTURE & ALIGNMENT	Organisational Structure and Reporting Alignment	
			Resources and Staffing	
			Roles and Responsibilities	
		<b>Percentage (%) Weight for Procurement Organisation Criteria</b>	8.54	
		CUSTOMER MANAGEMENT	Cross Functional Integration/Collaboration	
			Procurement Category Segmentation & Prioritisation	
	Spend Analysis & Reporting Framework			
	<b>Percentage (%) Weight for Customer Management Criteria</b>	5.92		
	PROCUREMENT SYSTEMS	Procurement Management System (ERP)		
		eProcurement (eTendering, eAuctions & eCatalogues)		
		Procurement Systems Integration (Procure to Pay)		
	<b>Percentage (%) Weight for Procurement Systems Criteria</b>	6		
	PROCUREMENT REGULATORY FRAMEWORK	PROCUREMENT POLICY & PROCEDURES	Policy and Regulatory Framework	
			Procurement Procedures, Guidelines and Contract Templates	
			Procurement Process, Methods and Evaluation	
			Procurement Process Performance Metrics	
		<b>Percentage (%) Weight for Procurement Policy &amp; Procedures Criteria</b>	10.5	
		PROCUREMENT GOVERNANCE	Compliance and Controls (Audit and Legal)	
Approval Authorities and Delegations (Boards and Committees)				
Disclosure of Data and Information				
	Integrity, Ethics, Corruption & Fraud			
	Complaints Disputes and Arbitration			
<b>Percentage (%) Weight for Procurement Governance Criteria</b>	8.92			
PROCUREMENT RISK MANAGEMENT	Contract Risk Management Framework			
	Supplier Risk Management Process			
<b>Percentage (%) Weight for Procurement Risk Management Criteria</b>	5.33			
CONTRACT PERFORMANCE MANAGEMENT	Contracts Management Process (Mobilization, Variation and Close-out)			
	Contract Monitoring and Performance Measurement			
	Supplier Performance Management			
	<b>Percentage (%) Weight for Contract Performance Management Criteria</b>	10.25		
CONTRACTS ADMINISTRATION	Invoice Processing			
	Document Management and Archiving			
<b>Percentage (%) Weight for Contracts Administration Criteria</b>	4.83			
TRAINING & CERTIFICATION	Procurement Training			
	Professional Accreditation			
	Contract Management Training			
<b>Percentage (%) Weight for Training and Certification Criteria</b>	6.54			
<b>TOTAL MUST EQUAL 100%</b>				<b>100</b>

### **APPENDIX 3. COMPARISON OF TRANSFORMATION EQUATIONS FOR SCALE NORMALISATION**

The tables in Appendix 3 show the normalised six, five and four-point scales used in this research, as discussed in Section 4.5 and shown in Table 23, and secondly the alternative normalised ratings scale used to test the proportionality of the scale used for reliability and consistency. In addition, the frequency distribution and analysis of both normalised scales for the project evaluation results is shown below for comparison between the overall calculated outputs. (Dawes, 2008; IBM, 2020; Leung, 2011).

#### **PROJECT EVALUATION NORMALISED RATINGS**

6-point scale	Normalised 6-point scale	5-Point scale	Normalised 5-point scale	4-point scale	Normalised 4-point scale
HS	1	A++	1	HS	1
S	0.800	A+	0.750	S	0.666
PS	0.600	A	0.500	PS	0.333
PU	0.400	B	0.250	U	0
U	0.200	C	0		
HU	0				

*Note: Highly Satisfactory (HS), Satisfactory (S), Partially Satisfactory (PS), Partially Unsatisfactory (PU), Unsatisfactory (U), Highly Unsatisfactory (HU)*

The alternative normalised ratings scale, as discussed in Section 4.5, and identified in the literature research and calculated from the transformation equation, is shown in the table below and used to test the reliability of the transformation equation used in this research (IBM, 2020).

#### **ALTERNATIVE PROJECT EVALUATION NORMALISED RATINGS**

6-point scale	Normalised 6-point scale	5-point scale	Normalised 5-point scale	4-point scale	Normalised 4-point scale
HS	1	A++	1	HS	1
S	0.820	A+	0.775	S	0.700
PS	0.640	A	0.550	PS	0.400
PU	0.460	B	0.325	U	0.100
U	0.280	C	0.100		
HU	0.100				

*Note: Highly Satisfactory (HS), Satisfactory (S), Partially Satisfactory (PS), Partially Unsatisfactory (PU), Unsatisfactory (U), Highly Unsatisfactory (HU)*

The results for the alternative transformation equation are calculated for the project evaluation results and compared with the results of the transformation equation used in this research.

MEAN RESULTS FOR NORMALISED PROJECT EVALUATION RATINGS

<i>Statistic</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	Skewness	Kurtosis	<i>p</i> -value	Rank
WBG	.595	.600	.201	-0.697	0.300	.000	7
ADB	.635	.667	.168	-0.837	2.788	.000	4
AFDB	.740	.800	.173	-1.880	3.570	.000	2
EBRD	.617	.600	.272	0.519	-0.180	.000	6
IADB	.628	.600	.171	0.827	0.375	.002	5
ISDB	.569	.667	.203	-1.303	0.252	.000	9
UNDP	.777	.800	.155	-0.310	-0.246	.000	1
UNICEF	.576	.667	.208	-0.010	-0.082	.000	8
DFID	.527	.500	.142	1.000	3.250	.000	10
JICA	.725	.667	.278	-0.851	0.202	.000	3

The results for the frequency distribution of the project evaluation results as used in this research is shown in Table 37. The results in the table below show no difference in the rank order and, although there is a minor increase across the means, the difference is under 2% and consistent for all the results in the table.

MEAN RESULTS FOR ALTERNATIVE NORMALISED PROJECT EVALUATION RATINGS

<i>Statistic</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	Skewness	Kurtosis	<i>p</i> -value	Rank
WBG	.634	.640	.185	-0.856	0.949	.000	7
ADB	.670	.700	.157	-1.146	4.422	.000	4
AFDB	.766	.820	.156	-1.880	3.570	.000	2
EBRD	.649	.640	.258	-0.759	0.624	.000	6
IADB	.666	.640	.154	-0.827	0.375	.002	5
ISDB	.610	.700	.189	-0.536	1.073	.000	9
UNDP	.799	.820	.139	-0.310	-0.246	.000	1
UNICEF	.618	.700	.192	-0.277	0.476	.000	8
DFID	.574	.550	.126	1.000	3.400	.000	10
JICA	.752	.700	.251	-0.871	0.278	.000	3

The alternative transformation equation results for the statistical distribution and the tests of normality are the same for both transformation equations and therefore the results and comparison that original transformation equation used is appropriate for this research.