



**Effectiveness of Evaluation Practices in
Supporting Regional Development
Planning in Indonesia:
The Cases of DI Yogyakarta and West
Sumatra Provinces**

By

Dwi Ratih Suryantining Esti

Thesis

Submitted to Flinders University

for the degree of

Doctor of Philosophy

College of Business, Government and Law

October 2021

DECLARATION

I certify that this thesis:

1. does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and
2. to the best of my knowledge and belief, does not contain any material previously published or written by another person except where due reference is made in the text.

A handwritten signature in blue ink, appearing to read 'Dwi', with a stylized flourish extending from the end.

Dwi Ratih Suryantining Esti

24 May 2021

ACKNOWLEDGEMENTS

Praise be to Allah the Almighty for giving me the blessing, knowledge and strength to complete my PhD thesis. I am immensely indebted to many people who have been supporting me upon the completion of this thesis. First of all, I would like to express my gratitude to my supervisors: Associate Professor Noore Alam Siddiquee, Professor Gerry Redmond, and Jo Baulderstone who have assisted me in formulating ideas through constructive discussions. Thank you very much for your continuous support, I have learned a lot from you.

I would like to acknowledge the support from the Australia Award Scholarship. I also would like to acknowledge my institution, the Ministry of National Development Planning of Indonesia (BAPPENAS), that has allowed me to pursue my doctoral study in Australia. My gratitude also goes to the College of Business, Government and Law, Flinders University (particularly Caitlin Hughes), the Office of Graduate Research (OGR) Flinders University (particularly Kate Wilson), and International Student Service (ISS) Flinders University, particularly Cassie Ankers and Shervin Shokri, for their kind and helpful assistance.

My gratitude also goes to my friends and colleagues in Flinders University, including Tui (my dear friend from the beginning of our PhD journey in Adelaide), PPIA friends, FLIPSA friends, and also the Biggles (especially Reem, Felicia, and Nadeeka). I also want to thank David Langdon for helping me in editing my thesis. I am also thankful for all the support provided by the Indonesian communities in Adelaide for their warm welcome, knowledge and friendships. I would also like to acknowledge the support from many institutions and research participants during my fieldwork in Indonesia.

My heartiest gratitude goes to my family. My husband Yuda Kamsi Abadi, my children Atifa Machiko Putri and Adela Dwi Haniya (born during the writing of the thesis) who always generate a lovely ambiance by their presence during my PhD journey. The PhD journey was a unique experience for all of us. We went through it when the world was experiencing a great difficulty and uncertainty (the global pandemic), but it was also a time of great resilience and positive change for us. Without the love and support from my family, my thesis would have not been accomplished. Yes, we made it together!

My deep gratitude also goes to my parents, Zaenal Mardi Chamid and Hanipah; my sisters Iin Innayathi S. Murthi and Tri Tania Lastri; and my parents-in-law, Edi Yudomo and Fatimah Resmiati, for their sincere thoughts and endless prayers during my study.

Finally, to everybody else who I have not mentioned who supported me during my studies. Thank you very much!

TABLE OF CONTENTS

DECLARATION.....	2
ACKNOWLEDGEMENTS.....	3
Table of Contents	5
List of Tables	9
List of Figures.....	11
Acronyms / Abbreviations/Terms	12
ABSTRACT.....	13
CHAPTER 1 INTRODUCTION.....	14
1.1 Research Background.....	15
1.2 The Indonesian Context	17
1.3 Brief Results of the Systematic Scoping Review.....	20
1.4 Research Hypothesis.....	21
1.5 Research Questions and Conceptual Framework.....	23
1.6 Research Significance.....	25
1.7 Structure of Thesis	26
CHAPTER 2 LITERATURE REVIEW	28
2.1 The relationships between policies, plans, programs, and projects.....	29
2.2 Development Planning	31
2.2.1 Definition	32
2.2.2 Approaches.....	32
2.2.3 Types of Development Plans	34
2.3 The Systematic Scoping Review.....	35
2.3.1 Review of the Stages	35
2.3.2 Results.....	39
2.4 Evaluation Practices in Development Planning.....	43

2.4.1	The Concepts	44
2.4.2	Potential Roles.....	45
2.4.3	Major Approaches	47
2.4.4	Utilisation.....	48
2.4.5	Actors	49
2.4.6	The Influencing Factors.....	50
2.5	Effective Evaluation.....	56
2.5.1	The concept of effectiveness.....	56
2.5.2	Effectiveness Criteria	57
2.5.3	Evaluation standards.....	58
2.6	Realistic Evaluation.....	59
2.6.1	The Main Concepts.....	61
2.7	Conclusion.....	65
CHAPTER 3 RESEARCH METHODOLOGY		67
3.1	Realist-Informed Research.....	68
3.2	Intervention Theory or CMO Configuration	70
3.3	Effectiveness Framework Adaptation.....	71
3.4	Case Study Selection.....	75
3.5	Data Collection	79
3.6	Data Analysis	83
3.7	Conclusion.....	85
CHAPTER 4 DEVELOPMENT PLANNING AND EVALUATION PRACTICES IN INDONESIA		86
4.1	Development Planning in Indonesia	86
4.1.1	National Development Planning.....	88
4.1.2	Regional Development Planning	91
4.2	Evaluation Systems and Practices in Indonesia.....	93
4.2.1	Regulatory Framework	97

4.3	Conclusion.....	101
CHAPTER 5 OVERVIEW OF CASE STUDY LOCATIONS.....		103
5.1	West Sumatra Province	103
5.1.1	Geographical Context.....	103
5.1.2	Socioeconomic Conditions	104
5.1.3	Governance in West Sumatra Province	110
5.1.4	The Regional Development Planning Agency (RDPA).....	113
5.1.5	Types of Evaluation Practices	117
5.2	DI Yogyakarta Province	120
5.2.1	Geographical Context.....	120
5.2.2	Socioeconomic Condition.....	121
5.2.3	Special Status of DI Yogyakarta Province	125
5.2.4	Governance in DI Yogyakarta Province	127
5.2.5	The Regional Development Planning Agency (RDPA).....	129
5.2.6	Types of Evaluation Practices	133
5.3	Conclusion.....	134
CHAPTER 6 RESEARCH FINDINGS		135
6.1	Effective evaluation practice in regional development planning.....	136
6.1.1	Effectiveness Framework of Evaluation Practices in Case Study Locations	136
6.2	What generates evaluation practice effectiveness, under which circumstances and why	142
6.2.1	Effectiveness of Evaluation Practices in West Sumatra Province	142
6.2.2	Effectiveness of Evaluation Practices in DI Yogyakarta Province.....	160
6.2.3	Identified Mechanisms and Contexts	178
6.3	Conclusion.....	184
CHAPTER 7 DISCUSSION.....		186
7.1	Working Definition of an Effective Evaluation Practice.....	187
7.2	Role of Evaluation Practices in Supporting the Regional Development Planning.....	192

7.3	What Generates Effective Evaluation and Under Which Circumstances: Refined Explanation	199
7.3.1	The Refined Context-Mechanism-Outcome Configuration (CMOc)	200
7.3.2	Unpacking the Mechanisms	212
7.3.3	The Importance of Context.....	240
7.4	Conclusion.....	243
CHAPTER 8 CONCLUSION		245
8.1	Key Findings of the Research	245
8.2	Implications for Policy and Practice	251
8.3	Limitations of the Research	252
8.4	Directions for Future Research.....	254
References.....		256
APPENDICES.....		267

LIST OF TABLES

Table 1.1 Initial Intervention Theory (Hypothesis).....	22
Table 2.1 Database Search Results (June 3, 2017).....	37
Table 2.2 Constructs of Mechanism at Levels of Systems.....	64
Table 3.1 Operational Stages.....	69
Table 3.2 The Number and Representation of Interview Respondents.....	81
Table 4.1 Planning Stages in Indonesia.....	88
Table 5.1 Population in West Sumatra Province 2015 – 2018.....	105
Table 5.2 Human Development Index (HDI) in West Sumatra Province.....	107
Table 5.3 Human Development Index (HDI) in West Sumatra Province.....	108
Table 5.4. Provincial Government Accountability Performance Top Ten Ranks.....	111
Table 5.5. Indonesia Government Index for West Sumatra Province (Range: 0-10).....	112
Table 5.6. Human Resources Composition in the Bappeda of West Sumatra Province.....	116
Table 5.7. Population in DI Yogyakarta Province by Gender.....	121
Table 5.8. Number of Tourist in DI Yogyakarta Province.....	123
Table 5.9. Indonesia Government Index for DI Yogyakarta Province (Range: 0-10).....	128
Table 5.10. Human Resources Profiles in Bappeda DI Yogyakarta Province.....	131
Table 6.1 CMO Configuration in West Sumatra Province under the Procedural Effectiveness Category.....	144
Table 6.2 CMO Configuration in West Sumatra Province under the Transactive Effectiveness Category.....	153
Table 6.3 CMO Configuration in West Sumatra Province under the Substantive Effectiveness Category.....	155
Table 6.4 CMO Configuration in DI Yogyakarta Province under the Procedural Effectiveness Category.....	161
Table 6.5 CMO Configuration in DI Yogyakarta Province under the Transactive Effectiveness Category.....	168
Table 6.6 CMO Configuration in DI Yogyakarta Province under the Substantive Effectiveness Category.....	173
Table 6.7 Summary of Identified Mechanisms and Elements of Context.....	179
Table 6.8 The Identified Positive Mechanisms in Each Case Study Location.....	180
Table 6.9 Relationships between Elements of Context with the Identified Mechanisms.....	181

Table 7.1 Summary of Identified Positive Mechanisms	201
Table 7.2 Proposed CMOC for the Transactive Effectiveness Category.....	209
Table 7.3 Proposed CMOC for the Substantive Effectiveness Category	210

LIST OF FIGURES

Figure 1.1 Conceptual Framework	24
Figure 2.1 Illustration of Policies, Plans, Programs, and Projects	30
Figure 2.2 Illustration of Relationships between Policies, Plans, Programs, and Projects	31
Figure 2.3 The Systematic Scoping Review Process	38
Figure 2.4 Realist Family Tree.....	60
Figure 2.5 The Realist Causation Explanation of the CMO Configuration	62
Figure 3.1 Research Design	68
Figure 3.2 The Developed Effectiveness Framework for Evaluation Practice	74
Figure 3.3 Representation of Case Studies.....	78
Figure 3.4 Location of Case Studies.....	78
Figure 4.1 Overview of Indonesia’s National Development Planning System	90
Figure 4.2 Stages of Indonesian Policy-making	94
Figure 5.1 Administration Map of West Sumatra Province.....	104
Figure 5.2 Organisational Structure of West Sumatra Provincial Government.....	110
Figure 5.3 Organisational Structure of the RDPA in West Sumatra Province	115
Figure 5.4 Quarterly Evaluation Report of West Sumatra Province	118
Figure 5.5 Evaluation Reports in the RDPA of West Sumatra Province.....	119
Figure 5.6 Administration Map of DI Yogyakarta Province	121
Figure 5.7 Organisational Structure of DI Yogyakarta Provincial Government	128
Figure 5.8 Organisational Structure of Bappeda DI Yogyakarta Province	130
Figure 5.9 Human Resources Profiles in Bappeda DI Yogyakarta Province	132
Figure 6.1 Effectiveness of Evaluation Practices in Regional Development Planning.....	138
Figure 6.2 Effectiveness of Evaluation Practices in Regional Development Planning.....	140
Figure 7.1 Proposed Comprehensive CMOc	202
Figure 7.2 Proposed CMOc for the Procedural Effectiveness Category	204
Figure 7.3 Groups of Mechanisms	213
Figure 7.4 Commitment Relationships with the Majority of Other Identified Mechanisms	224

ACRONYMS / ABBREVIATIONS/TERMS

- APBD : *Anggaran Pendapatan Belanja Daerah* (Regional Revenue and Expenditure Budget)
- APBN : *Anggaran Pendapatan Belanja Negara* (State Revenue and Expenditure Budget)
- Bappenas : Ministry of National Development Planning/National Development Planning Agency
- Bappeda/RDPA : Regional Development Planning Agency
- CMOc : Context-Mechanism-Outcome configuration
- DI Yogyakarta/DIY : *Daerah Istimewa* Yogyakarta/ Special Region of Yogyakarta
- DKI Jakarta : *Daerah Khusus Ibukota*/ Special Capital Region
- DPR : *Dewan Perwakilan Rakyat*/National Parliament
- DPRD : *Dewan Perwakilan Rakyat Daerah* (Local Parliament)
- GBHN : *Garis-Garis Besar Haluan Negara* (Supreme State Policy Guidelines)
- HDI : Human Development Index
- LKPP : *Lembaga Kebijakan Pengadaan Barang Jasa Pemerintah* (National Public Procurement Agency)
- MoF : Ministry of Finance
- MoHA: Ministry of Home Affairs
- Musrenbang : *Musyawarah Perencanaan Pembangunan* (Development Planning Consultative Forum)
- NGO : Non-Governmental Organisation
- OECD : Organisation for Economic Co-operation and Development
- Permendagri : Regulation of the Minister Of Home Affairs
- RKP : *Rencana Kerja Pemerintah* (National Government Work Plan)
- RKPD : *Rencana Kerja Pemerintah Daerah* (Regional Government Work Plan)
- RPJMD : *Rencana Pembangunan Jangka Menengah Daerah* (Regional Medium-Term Development Plan)
- RPJMN : *Rencana Pembangunan Jangka Menengah Nasional* (National Medium-Term Development Plan)
- RPJP : *Rencana Pembangunan Jangka Panjang Nasional* (National Long-Term Development Plan)
- SKPD/OPD : *Satuan Kerja Perangkat Daerah /Organisasi Perangkat Daerah* (Regional Working Unit)
- SPPN Law : *UU Sistem Perencanaan Pembangunan Nasional* (National Development Planning System Law)

ABSTRACT

This research examines the effectiveness of evaluation practices in supporting regional development planning in Indonesia using two case studies at the provincial level. Many publications on the topic of evaluation practice already exist, indicating its importance in evaluation literature. However, there are no studies that have discussed the effectiveness of evaluation practices, with all the prerequisites, in planning processes and particularly in the context of developing countries. On the other hand, in many countries, including Indonesia, there has been a growing call for public policy in regional development planning to be more evidence-based. Consequently, there are major gaps in the literature about what occurs in government agencies with regard to evaluation practice and its role as part of research-based evidence in regional development planning. Therefore, this research will assist in filling a major gap in the literature by revealing the processes, elements of context, and mechanisms through which evaluation practice works to improve its effectiveness. Realist methodology was selected for this study to better understand causal mechanisms within a given context of the research. In-depth interviews and focus group discussions with major stakeholders in the provinces of West Sumatra and DI Yogyakarta were conducted to help identify the key mechanisms and elements of context which influence the effectiveness of evaluation practice in regional development planning. This research found that considerable inadequacies in practice undermined the effectiveness of evaluation in case study locations, leading to the utilisation of evaluation predominantly for reporting or accountability purposes rather than for learning and improvement. An important contribution of this study is the working definition of an effective evaluation practice in regional development planning context where effectiveness is categorised into procedural, transactive, substantive, and normative effectiveness based on the Effectiveness Framework developed. Even though no single, general strategy can be applied in ensuring effective evaluation practice across all contexts, the intervention theories developed through this research highlight important areas for attention when designing evaluation practice. Future efforts should prioritise contextually sensitive, evidence-informed strategies, in order to optimise the role of evaluation practice in supporting regional development planning. These findings are expected to make a valuable contribution to the literature, stimulate discussions, and encourage further research in this key area of governance. They also have significant implications for theory and practice in this field.

CHAPTER 1 INTRODUCTION

This research investigates evaluation practices in regional development planning at provincial level in Indonesia. Specifically, it identifies and explains the underlying mechanism and specific contextual factors that lead to effective evaluation practice in supporting regional development planning. This is based on the background that the existing literature does not discuss the effectiveness of evaluation practices comprehensively, as far as the planning processes are concerned, particularly in the context of developing countries. Therefore, this research will make a significant contribution to fill the gap in the literature, coupled with recommendations to improve the effectiveness of evaluation practices in supporting regional development planning. It is one of the first studies in a developing country context to examine and answer questions about the effectiveness of evaluation practices in informing regional development planning: how, why, and for whom?

Evaluation in regional development in this study is referred to as the systematic activities by government agencies to assess, make judgements, or determine the value of a program or policy, or plan, in order to improve future practice. Evaluation is a broad concept that can be examined in more detail in terms of evaluation practice, evaluation theory, and evaluation policy. In this research, the emphasis is on the evaluation practices where definitions vary between different contexts. However, in planning context, evaluation practice is usually referred to as the everyday task of doing evaluation, such as assessing the plans or the planning processes or the outcomes.

The aim of this chapter is to describe the background, hypothesis, research questions, conceptual framework, and significance of this research. It consists of seven major sections. Section 1.1 provides a brief background of this research, and is then followed by Section 1.2, which provides the background about the Indonesian context. Section 1.3 briefly summarises the results of the systematic scoping literature search that was conducted in the early stages of this project to identify gaps in the literature. This systematic search found that there are no studies that have examined and discussed the effectiveness of evaluation practice in relation to regional development planning, particularly in the context of Asian developing countries. Therefore, there is a clear gap

in the literature on what occurs in Asian regional or local government agencies with regard to evaluation practice and its effectiveness in supporting regional development planning process.

The results of the systematic scoping literature search outlined the hypotheses of this study which are described in Section 1.4. Section 1.5 presents research questions, and the conceptual framework, which are formulated based on the backgrounds, systematic search results, and the research hypothesis. Section 1.6 explains the significance of the research in this thesis. Lastly, Section 1.7 describes the structure of the whole thesis.

1.1 Research Background

A shift in public sector management has been occurring to further acknowledge evaluation as an essential part of the development management process, especially in evidence-based policymaking. Shaxson (2016: 43) has identified evaluations as one of the five main types of evidence that policymakers tend to use and access besides data from statistical offices, administrative data from service providers, evidence from citizens, and other research-based evidence. In this study, evaluation is referred to as the systematic activities to assess, make judgements, or determine the value of program or policy or plan, in order to improve future practice. Many studies highlight the important role of evaluation in public policymaking (Sanderson, 2002; Segone, 2008; Head, 2016; Hellmut, 2017), including in planning (Oliveira and Pinho, 2010a; Mueller and Hersperger, 2015; Guyadeen and Seasons, 2018). Evaluation is considered to make a significant contribution to good governance, especially by fostering transparency, although it is context-specific (Dahler-Larsen and Boodhoo, 2019). Furthermore, evaluation supports the policymaking process by providing information to policy makers, the public, and other stakeholders about the worth of interventions from public sector organisations (Chouinard, 2013; Cousins *et al.*, 2014). In the planning area, evaluation can be utilised for understanding the impacts of planning; to increase the legitimacy of plans (Oliveira and Pinho, 2010a, 2011); to determine lessons, which can be utilised to guide future planning; and to support capacity building to better design, implement and value planning interventions (Faludi, 2000; Roberts, 2006).

However, researchers worldwide are concerned about the limited utilisation of evaluation results, where one of the internal factors is the lack of a supportive institutional environment (Bamberger *et al.*, 2004; Olejniczak, 2013). Guyadeen and Seasons (2016) observed that evaluations in government planning practice are underused. Even when public officials have access to good evaluations, there is no guarantee that they will use them, since they may prefer to make adjustments based on political decisions that are to their advantage (Head, 2016). Hence, the recommendations of evaluations may be ignored, making its utilisation limited by other, underlying practices. This issue of limited utilisation indicates that the value of the role of evaluation practice in supporting policymaking, including the planning process, is doubtful.

Many publications on the topic of evaluation practice already exist, indicating its importance in evaluation literature (Guyadeen and Seasons, 2018). However, research on evaluation practices has primarily discussed their utilisation aspect in the context of OECD countries (Shulha and Cousins, 1997; Seasons, 2002b), while only limited studies discuss evaluation practices in Asian countries, especially in Indonesia. On the other hand, in the regional development planning of many countries, including Indonesia, there have been growing calls for public policy to be more evidence-based (Sutmuller and Setiono, 2011; Zhang, 2015). Additionally, Kaşkol *et al.* (2018) state that cultural similarities concerning complex relationships and social conditions increase the chances of the effective implementation of solutions between countries with culturally similar societies. Unfortunately, no studies could be identified that discuss the effectiveness of evaluation practices in Asian countries. Considering Indonesia is one of the Asian countries which tends to have stronger culture-determined value systems, it is worthwhile to investigate the evaluation practices in Indonesia to provide perspectives about Asian countries and make additional contribution to the literature.

To begin this investigation of evaluation practice, two initial questions arise: How effective is the current evaluation practice in Asian countries' public agencies, specifically in Indonesia's regional development planning agencies? What are the factors that influence the effectiveness of the evaluation practice?

This research aims to answer these important questions and will emphasise the implications of the context to the applicability of the answers.

1.2 The Indonesian Context

Indonesia has undergone a massive transformation from centralised to decentralised government and regional autonomy. Decentralisation may be defined as a transfer of power, authority, and responsibility from the upper to the lower level of government, and also the sharing of authority and responsibility between government and private sector and civil society in economic activities, for instance, promoting economic growth and social development (Cheema and Rondinelli, 2007). The decentralisation policy in Indonesia was developed following the successful adoption of decentralisation process in other parts of Asia and several African countries. Decentralisation policy has been fashionable in the administration of developing countries in Africa and Asia as a response to the failure of the centralised system to address local concerns over development issues, such as inequality and low quality of services. The decentralisation in Indonesia aimed to improve the quality of governance and economic conditions of the country by empowering provincial and local governments, local parliaments and local communities (Firman, 2009; Holzhaecker *et al.*, 2016). This transformation is part of wider public sector reforms which aim to promote greater transparency and accountability in government affairs. As with such goals, the Republic of Indonesia is committed to implementing evaluation in the planning process (Haryana, 2013).

The Indonesian government consists of three levels: the national or central government, the provincial government, and the city or regency government. Based on the Law No. 23 Year 2014 on Local Government, both provincial and the city/regency governments are considered as local governments in Indonesia. It is important to note that this study mainly focuses on the provincial government. Indonesia has 34 provinces and, therefore, 34 provincial governments. Each province consists of cities and regencies. Based on the Regulation of the Minister of Home Affairs of the Republic of Indonesia Number 72 of 2019 concerning Amendments to the Regulation of the Minister of Home Affairs of the Republic of Indonesia Number 137 of 2017 concerning Code and Data for Government Administration Areas, there are 416 regencies and 98 cities in Indonesia.

Under the decentralisation policy put in place in 1999, provincial and city/regency governments enjoy greater power to manage regional development including planning, implementation, budgeting, monitoring and evaluation (Landiyanto, 2015). Since the start of the decentralisation era, the roles of provincial governments have changed several times. Under the Law No. 22 Year

1999 on Local Governments, the national government delegated authority to local government. Nevertheless, five years later, under the new Law No. 32 Year 2004 and Law No. 23 Year 2014 that superseded it 10 years later, the provincial government regained its role as the representative of the national government at the regional level. In evaluation practice, the provincial governments also have a larger role where, in addition to evaluating their regional development plans, they also are required to evaluate the regional development plans of the local governments in their province. As one effect arising from these changes in responsibility, Regional Development Planning Agencies (RDPA) have become more important players than previously in regional development planning, especially in conducting evaluation practices to support the planning process. Due to their responsibility for the whole regional planning cycle, RDPAs conduct evaluations in order to prioritise and monitor the effectiveness of policies (Zhang, 2015).

Additionally, Indonesia has national development targets which are stated in the National Mid-Term Development Plan that can only be achieved with support at the regional level. Based on existing laws, provincial governments are the representatives of the national government. Therefore, evaluation in regional development planning, particularly at the provincial level, is required to ensure regional development plans and programs are in line with national priorities and contribute to national development targets.

Monitoring and evaluation are fairly new development functions in Indonesia, having been introduced to the government institutions since the enactment of Law Number 25 of 2004 regarding the National Development Planning System (NDPS) (Haryana, 2013). Previously, evaluations were often sporadic and tended not to be implemented. There was regular reporting and monitoring, but they were only conducted at the Planning Ministry/BAPPENAS and in some cases at the Ministry of Finance only (Landiyanto, 2015). There were no specific working units to conduct evaluations either, especially at the local government level.

Since 2004, monitoring and evaluation has become an integral part of the planning cycle, including in local governments. As one of the consequences, all RDPAs in Indonesia are obligated to conduct evaluations of their regional development plans. This is regulated in the NDPS Law, and then specifically stated in the Government Regulation Number 8 of 2008 on Stages, Procedures of Preparation, Control and Evaluation of Regional Development Plan, especially in Articles 48.

In applying this Government Regulation, the Ministry of Home Affairs issued the Regulation No. 54 of 2010 on Implementation of Government Regulation Number 8 of 2008, which was revised by the Regulation No. 86 of 2017. These regulations stipulate that each regional government will evaluate their own regional development plans which consist of Regional Long-Term Development Plan (RPJPD), Regional Medium-term Development Plan (RPJMD), and Regional Government Work Plan (RKPD) (Government of Indonesia, 2008) to be input for their future planning. More details regarding evaluation practices in Indonesia after the decentralisation era and the related regulations will be explained further in Chapter 4.

However, recent research has demonstrated that evaluation results in RDPAs in Indonesia are underutilised. As shown by Zhang (2015) in her study on several local governments in Indonesia, there was almost no indication that evaluation results were used in the policymaking process in regional development planning. Specifically, in RDPAs, most of the evaluation results are not entirely used as input for further planning. Policy makers are more concerned about how to fulfil the administrative and legislative requirements than whether the policies they make will fulfil the needs of their citizens. Consequently, most evaluations were made only to fulfil their administrative obligations (Zhang, 2015). As a result, significant resources, including money and time, are wasted for years. In addition, poorly informed decision making in planning might lead to misallocation of limited resources, missed opportunities, and damage to political and professional reputations (Seasons, 2003). These conditions indicate that the evaluation practices hardly support planning process in most RDPAs.

Based on the descriptions above, it can be concluded that there are significant issues which may be affect the effectiveness of evaluation practices in supporting regional development planning in Indonesia. Therefore, this research is needed to better understand this and to enhance evaluation practice in the Indonesian context.

1.3 Brief Results of the Systematic Scoping Review

A systematic scoping review (SSR) was conducted in the early stages of this research to gauge the scope of the existing literature and to identify the gaps in the literature by obtaining relevant literature related to the research topic. This review is important to formulate more specific questions and hypotheses in this research. More detailed information on each stage of the systematic search can be found in Chapter 2.

The scoping review was based on a systematic literature search that employed a rigorous approach in searching a specific topic where all the steps involved in the process are predefined (Kitchenham, 2004). It helped to guide further exploration of relevant literature through identification of keywords, main theories or theorists, and appropriate databases.

The scoping review produced eight final results that were categorised into three major themes. The first theme is evaluation in policymaking. There are four studies in this theme, which are Martin and Sanderson (1999), Saunders (2012), Worthen and Schmitz (1997), and Widmer and Neuenschwander (2004). The second theme is evaluation and evidence-based policy. There are two studies in this theme, which are Sanderson (2000) and Hayton (2015). The third theme is effective evaluation in policymaking. This is the most relevant theme to this research topic, but only two studies, Seasons (2003) and Olejniczak (2013), belong in this category.

Results from the scoping review show that there is no explicit definition of effectiveness of evaluation practice in supporting evidence-based policymaking in regional development planning. However, drawing on Seasons (2003), it is proposed that evaluation practice would be effective if the obstacles when developing the evaluation process are well addressed. The identified impeding factors in the Canadian municipal planning context are resource constraints; an organizational culture that avoids criticism, does not support risk taking, and relies on limited evaluation results; a heavy dependence on quantitative research methods; indicators with unclear meaning and application; inadequate justification for evaluation; plan contents that are difficult to evaluate (vaguely worded); and insufficient communication of the evaluation purpose and approach in publication (Seasons, 2003).

Olejniczak (2013) identified four groups of mechanisms and factors that have facilitated the development of the system and utilisation of evaluation as a learning tool in the Polish

policymaking process. The first mechanism is growth in EU funding. The second is stability of institutions. The third is incentives. The final mechanism that shapes the system of evaluation is its architecture which has two aspects: the network of evaluation units, and the loose connection and relatively isolated position of evaluation units in relation to strategic and operational functions.

The initial scoping review gave a comprehensive overview on what the existing literature discusses in relation to the effectiveness of evaluation practice in evidence-based policymaking in regional development planning. This overview confirms the initial hypotheses of this research (as explained in the research background section) that almost no study discusses and examines the effectiveness of evaluation practice in assisting regional development planning in the context of Asian developing countries. Therefore, there is a clear gap in the literature on what occurs in Asian regional or local government agencies with regard to evaluation practice and its effectiveness in supporting regional development planning processes.

1.4 Research Hypothesis

The systematic scoping review results outlined the program or intervention theory which will guide the hypotheses in this study. Considering the focus of this research is not a program, ‘intervention theory’ will be the preferable term. In realist approach (this will be discussed in more detail in the following chapters), an initial intervention theory should be developed, and then used as a guide to build hypotheses to focus the research questions, select relevant data collection methods, and then guide the data analysis.

Basically, an intervention theory is a set of tangible assumptions that explain how the researcher envisages the intervention to achieve its expected outcomes (White, 2009; Masterson-Algar *et al.*, 2014). The initial intervention theory in this research was developed in the form of a Context-Mechanism-Outcome (CMO) configuration. The CMO configuration outlines a link between the context, the mechanisms, and the expected outcome which is effective evaluation practice in supporting regional development planning. From the systematic scoping review, two studies were identified to be significantly relevant to this research: Seasons (2003) and Olejniczak (2013). However, the differences in contexts between Canada as a developed country in the North American region, Poland as a developed country in the European region, and Indonesia as a

developing country in the Asian region should be kept in mind. Both studies identified key factors which determine an effective evaluation. Olejniczak (2013) identified four groups of mechanisms and factors that have facilitated the development of the system and utilisation of evaluation as a learning tool in the policymaking process. From a realist perspective, these key factors and mechanisms can be categorised into key mechanisms and elements of context. Additionally, conclusions from the other six studies as part of the scoping review were also considered in developing this initial intervention theory. However, due to their lower relevancy with the main focus of this research, the input from those six studies is not as significant as the inputs from Seasons (2003) and Olejniczak (2013). Table 1.1 illustrates the initial intervention theory in this research which was constructed based on the eight studies from the SSR results.

Table 1.1 Initial Intervention Theory (Hypothesis)

Context	Mechanism	Outcome
<ul style="list-style-type: none"> • Sufficient resources (time, money, and expertise) • Good organisational culture • Good level of support from politicians • High stability of institutions • Adequate justification for evaluation • Good network of evaluation units • Sufficient capability of practitioners • Good evaluability of plans • Appropriate indicators are available • Highly articulated roles and responsibilities for evaluation 	<ul style="list-style-type: none"> • Understanding • Commitment • Inclusiveness 	Effective evaluation practices in supporting regional development planning process

Based on Pawson (2013) and (Pawson and Tilley, 1997b), a CMO configuration is a hypothesis that the outcome (O) emerges because of the action of some underlying mechanisms (M), which only come into operation in particular contexts (C). Therefore, the proposed main hypothesis in this research is that evaluation practice is effective in supporting regional development planning if the identified underlying mechanisms and elements of context, especially the most influencing ones, are sufficiently present.

1.5 Research Questions and Conceptual Framework

From the systematic scoping review, it can be noted that there is a clear gap regarding evaluation practice support in regional development planning. Based on the backgrounds, scoping review results, and hypothesis discussed in previous sections, this research aims to address policy questions regarding improving evaluation practice and its effectiveness in supporting regional development planning: “How and to what extent are evaluation practices effective in supporting regional development planning processes?”

To answer the research question, related sub-questions are necessary:

1. What is an effective evaluation practice in the context of regional development planning?
2. How do evaluation practices support the regional development planning processes?
3. What are the circumstances that aid evaluation practices to be effective especially in supporting regional development planning process? What are the factors that create an enabling environment for effective evaluation?

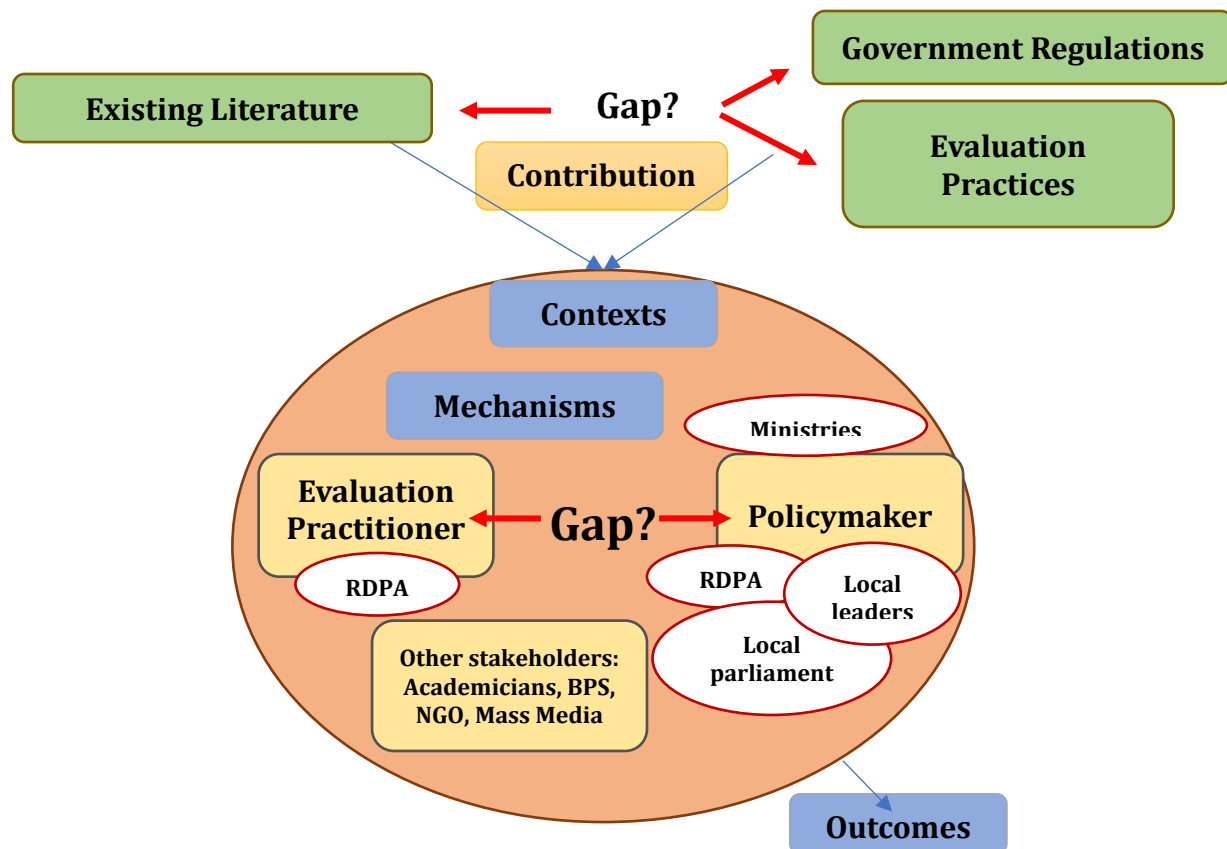


Figure 1.1 Conceptual Framework

As illustrated in Figure 1.1, the framework consists of interrelated blocks including:

1. The existing literature, government regulations, and current evaluation practices will be examined in this research. The literature is discussed in detail in Chapter 2 while the regulations and current evaluation practices in case study locations are discussed in detail in Chapters 4-6.
2. The realistic evaluation concepts: contexts, mechanisms, outcomes. These concepts will be the tool in analysis, especially to identify the factors of evaluation practice that influence its effectiveness in supporting regional development planning processes. The analysis and interpretation of these concepts will unpack the issues of whether evaluation practices are effective or not, why, for whom, and in what contexts. Chapters 2 and 3 provide thorough descriptions of these concepts, while the detailed discussions on these concepts in case study locations are presented in Chapter 6 and 7.
3. The national and regional context of development planning practice will be discussed in

this research. Inside these two broad contexts, there will be multiple units of analysis including the evaluation practitioners, policymakers, and other related stakeholders. The national context is discussed in detail in Chapter 4 while the regional contexts are discussed in detail in Chapters 5 and 6.

1.6 Research Significance

This research will contribute to the regional development planning and evaluation literature by developing intervention theories for evaluation practices in regional development planning. It will do so by identifying the necessary mechanisms and elements of context. Moreover, this research will develop a working definition of an effective evaluation practice in regional development planning based on the perspective of the main stakeholders, especially the evaluation practitioners and users. As concluded in the scoping review, hardly any study discusses and examines the effectiveness of evaluation practice in regional development planning in the context of Asian developing countries. Therefore, there is a clear gap in the literature on what occurs in Asian regional government agencies with regard to evaluation practice and its effectiveness in supporting regional development planning process. Therefore, this research will fill this gap. On a wider scale, this research will contribute to raising evaluation practitioners' understanding and capacity development around the factors related to evaluation effectiveness in regional development planning.

Evaluation practice is not new and Indonesia has realised how important it is, especially after the beginning of the decentralisation era, by formulating several policies related to evaluation in government institutions. This research will be useful for the Indonesian government, especially at the regional level, to recognise the concept and potential benefits and challenges of evaluation practice effectiveness in regional development planning. As noted by Landiyanto (2015), only a few studies on the Indonesian monitoring and evaluation system are available, such as Barbarie (1998), in the context of monitoring and evaluation before decentralisation, and Haryana (2013), in the context of monitoring and evaluation after decentralisation. Additionally, the available studies, including Barbarie (1998), Haryana (2013), and Landiyanto (2015), focus on the national level monitoring and evaluation system. Therefore, this study will make a significant contribution

to the literature, coupled with recommendations to improve the effectiveness of evaluation practices in supporting regional development planning.

In addition, as mentioned in Section 1.2, RDPAs in Indonesia are obliged to undertake evaluations of their regional development plans; however, the evaluation results remain under-utilised. Therefore, improving effectiveness of evaluation practice is expected to optimise the role of evaluation in supporting regional development planning. Ultimately, by having effective evaluation practices, regional development planning in Indonesia could support achievement of the ultimate goal, which is to improve the quality-of-life standards of its citizens.

1.7 Structure of Thesis

The remainder of this thesis consists of seven chapters and is structured as follows:

Chapter 2 reviews the theories and main concepts related to development planning, especially the evaluation practices. This chapter describes the relationships between policies, plans, programs, and projects, and reviews development planning which provide definitions, approaches, and types of plans. Chapter 2 further discusses the theoretical frameworks of evaluation practices in planning, includes the concepts, major approaches, potential roles, utilisation, actors, and influencing factors. This chapter also discusses the effective evaluation, divided into three subsections: the concept of effectiveness, effectiveness criteria, and evaluation standards. In the final section, this chapter discusses the realistic evaluation theories that are used as part of the methodology in this study.

Chapter 3 presents the research methodology. It specifically discusses the realist-informed research, intervention theory or CMO configuration that is developed and analysed further in Chapters 6 and 7, effectiveness framework adaptation which is used in Chapter 6, case study selection, and the research methods for data collection and the data analysis employed in this research.

Chapter 4 presents an overview of development planning and evaluation in Indonesia under decentralisation. The overview describes national and regional development planning and monitoring and evaluation systems and practices in Indonesia.

Chapter 5 introduces the case study locations (West Sumatra Province and DI Yogyakarta Province), including geographical context, socioeconomic conditions, the Regional Development Planning Agency, governance, types of evaluation practices, and the special status of DI Yogyakarta Province. The overview leads to the fieldwork result discussed in Chapter 6.

Chapter 6 presents the findings from the case study locations and provides a working definition of an effective evaluation practice in a regional development planning context. The chapter also examines what generates effectiveness in evaluation practices, for whom, under what circumstances and why, by focusing on the elements of context, the mechanisms, interrelations between mechanisms and elements of context, and the outcomes.

Chapter 7 focuses on analysing the major findings from this research. It presents the relationship between the literature and empirical findings, and the relationship between findings from case studies.

Chapter 8 highlights the main contributions of this research. It also summarises recommendations for policy and practice, especially for policymakers and practitioners in Indonesia. Finally, this chapter also identifies the limitations of this research and points to future research directions.

CHAPTER 2 LITERATURE REVIEW

The aim of this chapter is to review the related literature on the topic of this research. It consists of five major sections. Section 2.1 describes the relationships between policies, plans, programs, and projects. This section will try to illustrate the relationships between the overarching policies which direct the plans, specific policies inside the plans which direct the programs, programs as operationalisations of plans, and projects as implementations of programs. This is also expected to help improve understanding about the relationships between policymaking and planning process.

This study focuses on evaluations within the development planning process. Therefore, Section 2.2 reviews development planning which includes definitions, approaches, types of plans, and description of the planning process. Currently, there is only limited literature that discusses the relationship between national and regional planning. Therefore, it is important to describe these types of development planning processes together, where regional development planning is the main area of this research and the national development planning has a significant relationship to it.

Section 2.3 discusses the systematic scoping review. This review was conducted in the early stages of this research to gauge the scope of the existing relevant literature and to identify the main gap in the research topic, which is the effectiveness of evaluation practice in supporting evidence-based policymaking in regional development planning. This section describes the stages of this systematic scoping review and summarises its results.

Section 2.4 is devoted to the main focus of this research, which is the evaluation practice in planning. It explores six aspects: concepts, potential roles, major approaches, utilisation, actors, and influencing factors. These are the elements of evaluations that are highly relevant to the topic of this research and will be explored in the research.

Section 2.5 addresses the first sub-question of this research which is about defining effective evaluation practice in the regional development planning context. It is divided into three sub-sections: the concept of effectiveness, effectiveness criteria, and evaluation standards.

Finally, Section 2.6 delves into realistic evaluation, especially its three main concepts: context, mechanism, and outcome. These concepts are used in following chapters to support the data analysis in this research.

2.1 The relationships between policies, plans, programs, and projects

Prior to the discussion of this research, it is useful to become familiar with the differences and the relationships between policies, plans, programs, and projects. These are four common concepts, but with different instruments or tools or interventions utilised by public sector agencies in achieving the proposed goals, where often programs or projects are the enactment of plans that are developed from a policy. Dye (2011: 1) states that “public policy is whatever governments choose to do or not to do”. De Coning (2006) states that policy interprets the value of society and is usually manifested in the management of related projects and programs. Another perspective is conceptualised by Pawson and Tilley (1997b) who believe that social and public policies are delivered through active programs to active subjects. Conyers (1982) indicates that a policy process involves decision making to direct the changes that must occur, while planning is the process of making decisions about which actions should be undertaken to produce this change. Therefore, in this case, policy precedes planning. In this research, these types of policies are called overarching policies. Moreover, planning is a practical translation of these types of policies. This understanding is in line with Brynard (1996: 132) who states that “planning is a way of enabling policies to be translated into practical action programs which produce visible results”. Plans are one of the forms of government tools, which play an important role in planning. They are considered as the key currency of the planning profession (Guyadeen and Seasons, 2018). According to Faludi (2000: 203), “plan is supposed to be an unambiguous guide to action, so its adoption implies closure of image of the future”. Plans could be broader than policies and programs considering they contain a vision for development in the future and are accompanied by relevant goals, facts, and policies that translate the vision into spatial elements (Berke, Godschalk, *et al.*, 2006).

Another government intervention can be in the form of programs or projects. Programs are usually utilised to actualise general and abstract policies, they are constituted of cluster activities

developed in order to achieve specific objectives (McDavid, 2006). According to Stockmann (2011: 20), programs are “derived from a political strategy, planned and implemented in individual implementation steps, and as a rule endowed with financial support for a limited period of time, in order to achieve desired effects”. Meanwhile, Moore and Spires (2000: 227) defined a program as “a group of interventions often linked together over time and encompassing more than one project” and a project as “a single intervention or discrete, one off, form of activity”. Based on this definition, it can be concluded that a program may consist of more than one project. Therefore, project may be referred to as a component of a program. A project is a planned intervention to achieve one or more goals, which includes a series of interrelated activities carried out over a limited period of time, using certain financial, physical, and human resources (Dale, 2004).

Based on the definitions presented above, Figure 2.1 illustrates the relationships between policies, plans, programs, and projects that will be discussed further in this research. It describes the relationships between the overarching policies which direct the plans, the specific policies inside the plans which direct the programs, programs as operationalisations of plans, and projects as implementations of programs. From these relationships, it can be concluded that policymaking and planning processes in this research are different notions but are closely related.



Figure 2.1 Illustration of Policies, Plans, Programs, and Projects

An example of the relationship between overarching policy, plan, program, and project is demonstrated in Figure 2.2.

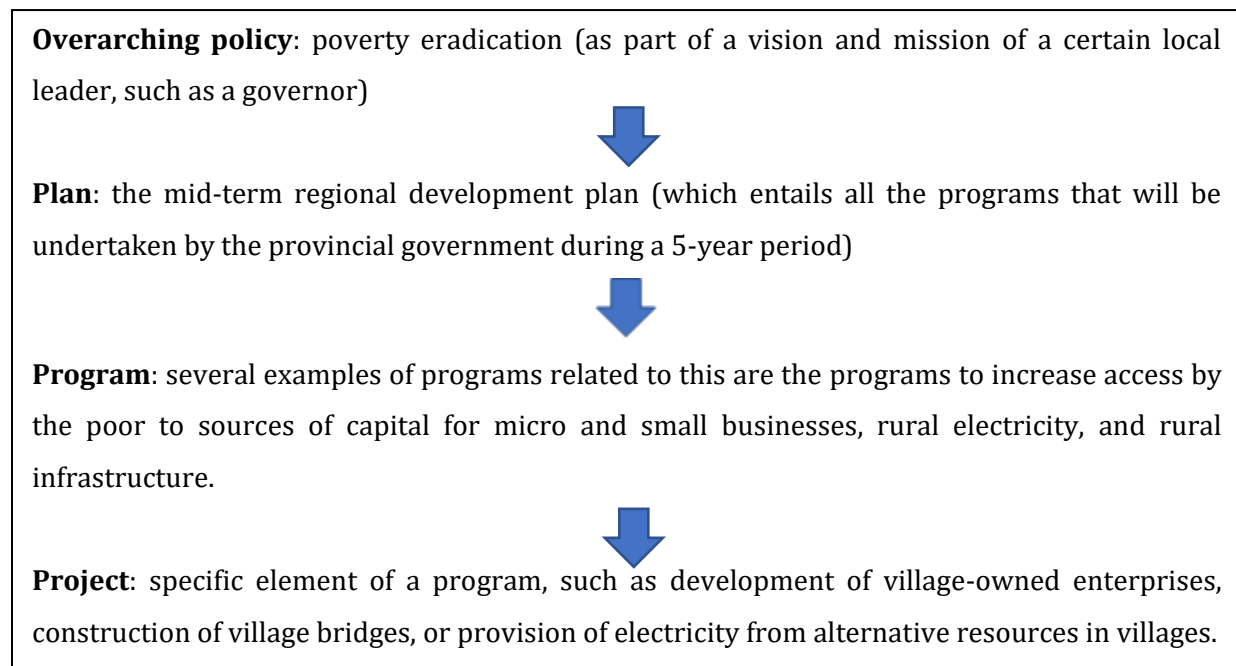


Figure 2.2 Illustration of Relationships between Policies, Plans, Programs, and Projects

It is important to understand that plans are an instrument that constantly evolve and undergo continuous updates and revisions to remain relevant to the changing knowledge and needs (Brody, 2003). Consequently, evaluation plays an important role in ensuring the relevance and applicability of policies, plans, programs, and projects. For the purpose of this study, the focus will be on evaluations of the plans within the development planning process. The next section will explore the conceptual aspects of development planning.

2.2 Development Planning

In this section, three elements will be addressed in relation to development planning, namely definitions, approaches, and types of plans.

2.2.1 Definition

There are a variety of planning types, but in this study we shall particularly focus on development planning, which began in earnest in the latter part of the 20th century. It is the kind of planning which is employed by governments or other public agencies to bring about general social and economic change (Conyers, 1982). It is a process that includes decision making to select from a range of alternatives for using resources in order to achieve future goals (Conyers and Hills, 1984). One of the seminal authors in the field of development planning, Lewis (1966) described development planning as a collection of policies and development programs to encourage people and the private sector to use available resources more productively. However, Bryant and White (1982: 246) defined development planning as “an effort by the government to enlarge its capacity to make choices to consider and select among its alternatives”. More recently, Yağci and Ardiani (2018: 215) have argued that:

“development planning in many emerging economies has been an essential part of policy making in many respects, not just for economic concerns. These plans have been utilized as “roadmaps” for realizing economic, social, and political objectives in many countries, and these plans give a very good understanding of priorities and strategies of policy makers in different contexts”.

In conclusion, development planning over its evolution has come to mean comprehensive planning which has greater social and political significance and requires decision making by the government or public sector agencies to select the most appropriate strategies for using available resources to achieve the desired objectives in the society.

2.2.2 Approaches

In addition to defining development planning, this subsection describes the various approaches to development planning. Both Todaro (1971) and Muhammad (2017) differentiated these different approaches by the scope of the planning model:

- a. The aggregate model is the simplest type of development planning. This model is related to the economy as a whole and involves aggregate components such as consumption, production, investment, savings, exports, and imports.
- b. Sectoral planning is carried out with a sector-based approach. For example, Health Sector

planning or Education Sector planning are common approaches. A sector is a collection of activities or programs that have similar characteristics and objectives. Although this approach determines certain activities by certain agencies in certain locations, the location is of less importance, and only seen as locations of activities. This approach is different from other planning approaches which emphasise the location of activities.

- c. The model of the inter-industry approach. This approach considers that economic activities in the main industrial sectors are always interconnected with one another in a simultaneous set which eventually shows the production process or technology used in each industrial sector.

However, the approaches which are more frequently discussed in development planning processes in developing countries, especially in Indonesia, are identified by Muhammad (2017) :

- a. The political approach, where the process of the development plan is based on the elaboration of the vision and mission of the nation, region or local leader.
- b. The technocratic approach, where development planning is carried out using scientific methods and frameworks by institutions or work units that are functionally assigned to it.
- c. The participatory approach, which involves all levels of society in the development planning process.
- d. The top-down approach, in which the initiative of the "top" organisation/unit/ institution is followed up or translated down to the structures or levels below.
- e. The bottom-up approach, in which initiatives of the lower organisation/ unit/institution are followed up (translated upwards) to higher levels.

Considering the case study locations of this research are in Indonesia, all of the above development planning approaches are relevant to the current study (especially the political and technocratic approach) and will be further discussed in the analysis and discussion chapters of this thesis.

2.2.3 Types of Development Plans

Dale (2004) distinguished development planning by its functional level and the type of area. By the type of area, he categorised development planning according to the administrative level at which it is undertaken or the type of area to which it relates: urban development planning and rural development planning. By functional level, he categorised development planning into national, regional and local community planning. This research focuses on the categorisation of development planning by functional level considering the emphasis of the study is on evaluation practices in supporting regional development planning.

Todaro (1994: 566) described national development planning as “a deliberate governmental attempt to coordinate economic decision making over the long run and to influence, direct, and in some cases even control the level and growth of a nation’s principal economic variables (income, consumption, employment, investment, saving, exports, imports, etc) to achieve a predetermined set of development objectives”. Turner and Hulme (1997: 139) mentioned that this type of planning “can contribute to the more effective programming of publicly financed development projects”. Dale (2004: 50) stated that “it could be undertaken entirely at the national level or it be deconcentrated to larger or lesser extents. In the latter case, units of the central government at one or more subnational levels, such as provinces and/or districts, are responsible for at least some of the planning confined to their geographical areas, within frameworks of rules and regulations provided by the head office of the respective national agencies.”

Regional development planning is carried out by provincial governments and/or the federal governments which usually focus on socioeconomic development with the aim of improving the well-being of people (Webster and Robinson, 1985). Similar to national development planning, regional development planning can be carried out at a certain regional level, such as province or a district, or it may be deconcentrated to units below that level (Dale, 2004). This type of planning can be a means of disaggregating the national development plan and the sectoral components into comprehensive regional plans or vice versa, it can also be an instrument to aggregate local plans, programs and projects into comprehensive regional development plans (Weissmann, 1969).

For the purpose of this research, the focus will be on the regional development planning considering this research is investigating evaluation practices in supporting regional development

planning at the provincial level. However, there will be some discussion of national development planning as well.

2.3 The Systematic Scoping Review

In the previous section, the definitions, approaches, and types of development planning were reviewed. In the next sections, the focus shifts more specifically to the evaluation practices in development planning. Six elements of evaluation in planning are discussed: the concepts, potential roles, major approaches, utilisation, actors, and influencing factors. However, before the six elements of evaluation in planning were reviewed, a systematic scoping review was conducted in the early stages of this research to gauge the scope of the existing relevant literature and to identify the gaps in the research topic: effectiveness of evaluation practice in supporting evidence-based policymaking in regional development planning. This review was important to formulate more specific questions and hypotheses involving this research.

2.3.1 Review of the Stages

The scoping review in this research was divided into six main stages. It followed the practical guide for carrying out systematic reviews in the social sciences as suggested by Petticrew and Roberts (2008) with a slight adjustment based on the needs of this research. The last stage of the guide, which is to “disseminate the findings of the review” will be done later as part of the full research dissemination). The six main stages in this review are:

1. Framing the review question
2. Deciding inclusion/exclusion criteria
3. Searching the literature
4. Screening the search results
5. Assessing critically the included studies
6. Synthesizing the studies

The broad topic of this review is the **effectiveness of evaluation practice in supporting evidence-based policymaking in regional development planning.**

The initial review questions for this scoping review were:

- How have studies on evaluation practice defined effectiveness?
- How is this definition related to the potential support of evaluation practice in evidence-based policymaking in regional development planning?

After framing the review questions, the initial search keywords and some inclusion/exclusion criteria were determined. The main inclusion criteria for publications in the review were that they be: i) From a database which was recommended by Flinders University Library for ‘Politics and Public Policy’ platform relevant to ‘Politics and Public Policy’ for accessing peer reviewed academic publications; ii) Dated 1980 onwards; iii) High quality, peer reviewed journal articles (ranked by SCImago as being in the top two quartiles); iv) In English language; v) Available in full text through Flinders University Library (physical, electronic subscriptions or document delivery/interlibrary loan); or free of charge via the internet.

First, a set of selected keywords was formulated based on a preliminary search of several sets of related keywords. The sets of selected keywords were:

```
((("Evaluation practice" OR "evaluation effectiveness" OR "effective evaluation" OR "evaluation activity" OR "evaluation result") NEAR/5 (policy* OR plan* OR develop*)) AND (evidence OR "Evidence-based" OR "Evidence informed") AND ((regional OR municipal* OR local) NEAR/5 (policy* OR plan* OR develop*))) NOT ((history OR educat* OR medic* OR health OR psych* OR environment))
```

Considering the different Boolean operators, truncation, wildcards, word limit of keywords, and proximity searching in databases, the keywords used in each database are slightly different.

Next, appropriate databases were identified by following recommendation from the search smart guide for searching articles in the Politics and Public Policy section provided by Flinders University Library (http://flinders.libguides.com/politics_publicpolicy/articles). Ten databases are provided in this guide, including APAFT - Australian Public Affairs Full Text (Informit), Cambridge Core, Emerald Fulltext, Expanded Academic ASAP, Google Scholar, JSTOR, Oxford University Press Journals, Proquest, SAGE Journals, and Scopus. After using the selected keywords in each database, some indications showed that not all of the databases were relevant to the research topic. The indicators were the absence of “advanced search” (which led to an

abundance of low relevance results for the review topic) and minimum search results from initial search trials. Therefore, four databases (Informit, Cambridge Core, Emerald Fulltext, Oxford University Press Journals) were eliminated.

Table 2.1 Database Search Results (June 3, 2017)

No.	Treatment	Database searching results						TOTAL
		Expanded Academic ASAP	Google Scholar	JSTOR	ProQuest	SAGE Journals	Scopus	
1	Using advanced search	5	47	9	153	55	14	283
2	Using advanced search and inclusion/exclusion criteria	0	5	8	20	30	7	70

The number of database search results for each database is shown in Table 2.1. ASAP database had “0” result; therefore, this database was eliminated from further review. For each database, the review involved identifying all articles from peer reviewed journals and then determining whether the journal was of a higher quality. The criterion of quality was categorised by SCImago as ‘Development’ journals or ‘Geography, Planning and Development’ journals or 'Public Administration' journals or 'Sociology and Politics' journals and ranked by SCImago as being in the top two quartiles for that category. The SCImago Journal Rank is a publicly available portal including journals and country indicators developed from the information in the Scopus® database (www.scimagojr.com). This resulted in 56 results. The summary of all of the stages in this systematic scoping review is illustrated in Figure 2.3.



Figure 2.3 The Systematic Scoping Review Process

The 56 final results were then screened for relevance, initially through a reading of the titles and abstracts and then more thoroughly with a close reading of the full text of the article. The criteria of relevance used were:

1. The abstract indicates the study includes evaluation as a type of evidence in policymaking.
2. Closer read confirms the study discusses evaluation practice and its effectiveness in policymaking.
3. The study elaborates policymaking in regional development planning or regional planning or local planning.

Through a reading of titles and abstracts, 48 of the studies were assessed as not having covered the required indication with the relevance criteria. As they did not meet this relevance criterion they were excluded, leaving 8 positive returns.

2.3.2 Results

The eight final results were categorised into three major themes. The first theme is evaluation in policymaking, which included four studies: Martin and Sanderson (1999), Saunders (2012), Worthen and Schmitz (1997), and Widmer and Neuenschwander (2004).. The second theme is evaluation and evidence-based policy. The second theme is evaluation and evidence-based policy, including two studies: Sanderson (2000) and Hayton (2015). The third theme is effective evaluation in policymaking. This is the most relevant theme to this research topic but only two studies, Seasons (2003) and Olejniczak (2013), belong in this category.

Studies on evaluation in policymaking

All of the eight final results discuss evaluation in policymaking; however, four of these focus more on key concepts of this review namely “evaluation effectiveness”, “evidence-based policy”, and “regional development planning”.

The most common concept or issue considered in almost all eight studies is “usefulness” of evaluation or evaluation utilisation. Most of the studies mention that scholars and evaluation practitioners have long discussed the evaluation utilisation issue and tried to formulate different solutions. This issue was highlighted by Saunders (2012) and Widmer and Neuenschwander (2004). Saunders (2012) introduced the different interpretation of “use” and “usability”, where “use” emphasizes the context and organisational capacity where evaluation results are used; and “usability” focuses on how the evaluation design itself encourages the use of its results in the widest sense. Widmer and Neuenschwander (2004) examined the circumstances of evaluation use based on investigation of several Swiss federal administration offices.

On the other hand, studies by Worthen and Schmitz (1997) and Martin and Sanderson (1999) discussed different evaluation approaches for specific programs or projects. These two studies have low relevance to the topic of this research. Worthen and Schmitz (1997) defined and discussed cluster evaluation as an approach in evaluating large scale initiatives with multiple

project sites that need multi-levels of evaluation. They described its key features, position the concept in the evaluation field, distinguishing it from other forms of multi-site evaluation, and identified challenges of this approach in order to broaden its applicability and success. Martin and Sanderson (1999) discussed the different kinds of evaluation support and roles of evaluators in evaluating governmental pilot programs. They suggested some alternative social-constructivist evaluation approaches and argued that evaluators need to take on the role of change agents in combining summative evaluation of outputs and impacts with additional formative approaches, emphasising the development of a deeper understanding of processes.

Studies on evaluation and evidence-based policy

Both Sanderson (2000) and Hayton (2015) aimed to improve evaluation support as evidence in informing policy development for policy makers. They highlighted the complex reality in which evaluations are conducted. However, they have different emphases in their studies. Sanderson (2000) focused on the theoretical foundations of policy evaluation, whereas Hayton (2015) emphasised the evaluation methodologies.

Sanderson (2000) argued that complexity theory has important influences on policy evaluation. He suggested practitioners acknowledge evaluation as ‘practice’ instead of ‘technique’, an exercise in applying a variety of appropriate methods to particular conditions to provide understanding of the relevance of policy initiatives. He introduced the notion of ‘institutional conditions of communicative competence’, which the writer believed is more important to develop than the traditional ‘technical competence’ in order to harness evaluation more effectively in supporting evidence-based policymaking.

On the other hand, Hayton (2015) focused on the factors considered important in realistic evaluation which are: mechanisms, context, and the outcome configuration. Hayton suggested that these factors should be recognised and accepted by both policy makers and evaluators in order to improve evaluation practice. This study also argued for a higher level of transparency and honesty in evaluation practice, and higher awareness of its limitations.

Studies on effective evaluation practice in policymaking

Both Seasons (2003) and Olejniczak (2013) identified key factors which determine an effective evaluation but neither explicitly defined what an effective evaluation practice is. Seasons (2003) saw evaluation practice simply as evaluation activities in the planning process. Olejniczak (2013) portrayed evaluation practices as part of a wider system of regional policy decision making. His study also argued that evaluation practices are evaluation activities or studies “to produce knowledge based on evidence from performed interventions and to feed this knowledge into key strategic discussion or everyday managerial practices” (Olejniczak, 2013: 1661).

Olejniczak (2013) focused on the evaluation system as a whole, where evaluation practice was considered part of this system. He highlighted the criticism of evaluation practice in regional policy in Central and Eastern European countries that stresses its function in terms of accountability and formal reporting on progress rather than its role in policy or institutional learning. It was a research using systems thinking which focused on the case of regional policy in Poland, especially the evaluation practices conducted between 1999 and 2010. This study identified key mechanisms that determine the main function of evaluation as a learning tool that produces and utilises knowledge in the decision-making process. This study identified four groups of mechanisms as well as the factors that facilitated the development of the evaluation system, and utilisation of evaluation as a learning tool in the policymaking process. These mechanisms are: growth in EU funding, stability of institutions, incentives, and architecture of the evaluation system. The last mechanism (architecture of the evaluation system) has two aspects: the network of evaluation units; and the loose connection and relatively isolated position of evaluation units in relation to strategic and operational functions.

Seasons (2003) explored the existing condition of monitoring and evaluation activities and the main aspects that facilitate or hinder monitoring and evaluation practices in the department of planning at municipal level in Ontario, Canada. Seasons’ study comprised exploratory research using interviews with senior staff from the planning departments of the 11 regional municipalities in Ontario, a literature review, and a document analysis of municipal plans and policies. Findings from this study show that there is a significant gap between the normative ideal for monitoring and evaluation based on literature and the reality for planning departments in Ontario regional municipalities. The study found that the ideal monitoring and evaluation model is often unrealistic

and rarely attainable in its entirety. Factors affecting the potential for monitoring and evaluation practice identified from this study are: resources (time, money, and expertise), evaluation methods, appropriate indicators, causality (linking goals and outcomes), political realities, organisational culture (support from senior management, willingness to improve, tolerance of risk, acceptance to failure), evaluability of plans, justification for evaluation, clear and regular communication. In addition, this study argued that there are several aspects of political and organisational contexts which can determine the nature of monitoring and evaluation practice. In the political context, this study found that the planners' goals of efficiency and effectiveness seemed less important than political exigencies. For the organisational context, the results of this study suggest that the organisation conducting evaluation must have the right culture and the ability to learn by doing.

Both studies are part of the key literature in this research due to their relevance to this research topic and have managed to identify influencing factors for effective evaluation. Seasons' (2003) qualitative study used interview and document analysis as the research methods, whereas Olejniczak (2013) employed a mixed-method approach using systems thinking which involved quantitative data as well as interviews. However, neither study examined the relationships among the identified influencing factors for effective evaluation or which factors are the most influential. Therefore, a realist approach is expected to overcome these weaknesses. Nevertheless, considering these studies did not use the realist approach, the information from these studies needed to be analysed further to fit into the analysis and discussions of this thesis. Additionally, as explained in Chapter 1, the regional context is important. Therefore, the different contexts of Canada as a developed country in North America, Poland as a developed country in East Europe, and Indonesia as a developing country in the Asian region should be considered. For example, the importance of evaluation theory and practice have been recognised longer in North America and Europe than they have in Asia. One of the indications for this is that the establishment of evaluation societies in Canada and Europe have occurred in the 1980s and the 1990s, whereas the establishment of evaluation societies in Asia only begun in recent years (e.g., the Asia Pacific Evaluation Association was established in 2012).

The results from the scoping review show that there is no explicit definition of effectiveness of evaluation practice in supporting evidence-based policymaking in regional development planning. However, drawing on Seasons (2003), it can be concluded that evaluation practice would be

effective if the obstacles when developing the evaluation process are addressed. The identified impeding factors in the Canadian municipal planning context were resource constraints, which included an organisational culture that avoided criticism, little support for risk taking, and had reliance on limited evaluation results. Further impeding factors Seasons (2003) found were a heavy dependence on quantitative research methods; indicators with unclear meaning and application; inadequate justification for evaluation; plan contents that were difficult to evaluate (e.g., vaguely worded); and insufficient communication of the purpose of evaluation and approach used in publications.

Olejniczak (2013) found four groups of mechanisms and factors that facilitated the development of the system and utilisation of evaluation as a learning tool in the context of the Polish policymaking process. The first mechanism Olejniczak identified was growth in EU funding. The second was stability of institutions. The third was incentives. The final mechanism that shaped the system of evaluation was its architecture which had two aspects: the network of evaluation units, and the loose connection and relatively isolated position of evaluation units in relation to strategic and operational functions. Some of these identified mechanisms are relevant to this research, including the stability of institutions, incentives, and the network of evaluation units. These will be discussed in more detail in Chapter 6 and 7 of this thesis.

The initial scoping review gave a comprehensive overview of what the existing literature discusses in relation to the effectiveness of evaluation practice in supporting evidence-based policymaking in regional development planning. This overview confirms the initial hypotheses of this research that almost no study examines the effectiveness of evaluation practice in supporting regional development planning in the context of Asian developing countries which tends to have stronger culture-determined value systems. Therefore, there is a clear gap in the literature on what occurs in Asian regional or local government agencies with regard to evaluation practice and its effectiveness in supporting regional development planning processes.

2.4 Evaluation Practices in Development Planning

In the previous section, the definitions, approaches, and types of development planning were reviewed. In this section, six elements of evaluation in planning that are considered important for

this study will be discussed: the concepts, potential roles, major approaches, utilisation, actors, and influencing factors.

2.4.1 The Concepts

Evaluation is a broad concept that can be examined in more detail in terms of evaluation practice, evaluation theory, or evaluation policy. In this research, the emphasis is on the evaluation practice.

There are multiple definitions of evaluation in the literature. Patton (2008: 38) defined evaluation as “the systematic collection of information about the activities, characteristics, and results of programs to make judgments about the program, improve or further develop program effectiveness, inform decisions about future programming, and/or increase our understanding.” Similarly, Weiss (1998: 4) described evaluation as, “a systematic assessment of the operation and/or the outcomes of a program or policy, compared to a set of explicit or implicit standards, as a means of contributing to the improvement of the program or policy”. Pawson (2013: 115) defined evaluation as “the systematic determination of the merit, worth, and significance of programs”. In summary, evaluation can be defined as the systematic activities to assess, make judgements, or determine the value of program or policy, or plan, in order to improve further practice.

Definitions of evaluation practice vary between contexts. However, when the notion is used by evaluators, they usually refer to the everyday work of doing evaluation, such as developing the evaluation plans, collecting data, communicating findings, and so on (Stern, 2006; Kallemeyn *et al.*, 2015). Stern (2006: 293) stated that “evaluation as practice is an ‘open system,’ shaped by many particular societal, institutional and global contexts”. This perspective takes into account the contextual factors which shape what constitutes evaluation and how it is undertaken. Aligned with this, Kallemeyn *et al.* (2015) suggested that evaluation practice is influenced by various things such as evaluation theories and different levels of contexts. Additionally, O'Sullivan (2004: 2) observed that “the range of evaluation practice has often mirrored the personalities of evaluators and their accompanying preferences for making sense of the world, which in turn influences how they conduct evaluations and practice their evaluation”.

Evaluation practice in planning or plan evaluation is defined as the “systematic assessment of plans, planning processes, and outcomes compared with explicit standards or indicators” (Laurian

et al., 2010: 741). Khakee (1998) argued that, from both theoretical and practical points of view, planning and evaluation are inseparable concepts. Oliveira and Pinho (2010b) strongly contended that planning should be evaluated, and that evaluation should constitute a cyclical process developed together with the planning process cycle, that it should focus on the different aspects of planning (e.g., policies, plans, programs, processes, results) and that it should provide information that can be used to promote an effective planning dynamic.

Another notion that often relates to evaluation is ‘effectiveness’ which, on the whole, conveys the degree of achieving the goals (Green and South, 2006). Related to this, Patton (2002: 11) states, “when one examines and judges accomplishments and effectiveness, one is engaged in evaluation.” The Development Assistance Committee of the Organisation for Economic Cooperation and Development (OECD) defined effectiveness as:

“the extent to which the development intervention’s objectives were achieved, or are expected to be achieved, taking into account their relative importance. Also used as an aggregate measure of (or judgment about) the merit or worth of an activity, i.e. the extent to which an intervention has attained, or is expected to attain, its major relevant objectives efficiently in a sustainable fashion and with a positive institutional development impact”. (OECD, 2002: 20-21).

However, these definitions of effectiveness only focus on determining the extent to which the intervention met its goals, and not the needs of its recipients. Therefore, this research will extend existing definitions in the literature by formulating a more comprehensive yet specific definition of the effectiveness of evaluation practices in the regional development planning context.

2.4.2 Potential Roles

The importance of research, including evaluation, for policy making has received a considerable amount of attention in recent years (Sanderson, 2002; Head, 2016). A synthesis of a number of key works in the literature suggests that the main rationale for conducting evaluation in planning is to further support the planning process considering that evaluation has two primary roles: the accountability role and the learning or improvement role (Scriven, 1991; Chen, 1996; Meer and Edelenbos, 2006; Patton, 2008; Cousins *et al.*, 2014).

Evaluation supports accountability by assessing to what extent the objectives of an intervention are being or have been achieved. Evaluation “is a tool to be used to promote good governance, modern management practices, innovation and reforms, and better accountability” (Kusak and Rist, 2004: 60). Evaluation is considered to make a significant contribution to good governance, especially transparency, although it is context-specific (Dahler-Larsen and Boodhoo, 2019). Seasons (2002b: 52) stated that “evaluations can be conducted as a way to keep program stakeholders informed and as a means to ensure accountability by an organisation.” Evaluation can provide information about the progress, the achievement, and the value of government interventions (Leeuw and Furubo, 2008; Weiss *et al.*, 2008). According to Molas-Gallart (2012: 589), “it focuses on the direct audit of how resources are spent. In the case of research policy, the controlling purpose will typically focus on the analysis of inputs and the audit of direct research outputs, and will fit with traditional bureaucratic models of administration”. In the evaluation context, accountability serves as a means to secure legitimacy. Evaluation is now considered a legitimizing function and good governance practice (Chouinard, 2013). One of the greatest benefits of evaluation is that it holds planners and those involved in plan preparation accountable; this can be a way to legitimize the field of planning (Brody and Highfield, 2005; Laurian *et al.*, 2010; Oliveira and Pinho, 2010a).

On the other hand, the learning role focuses on deriving lessons from the past experiences with the intention of informing related stakeholders in order to improve future practice. This role therefore relies on the existence of feedback mechanisms and the operational flexibility needed to function as a learning organisation (Molas-Gallart, 2012). Olejniczak *et al.* (2016) stated that one of the main functions of evaluation in policymaking is to build a stronger knowledge base. Torres and Preskill (2001) and Cousins and Bourgeois (2014) also discussed about how evaluation relates to organisational learning. Evaluation is considered to be one of the supporting evidences to inform key decision makers in evidence-based policymaking (Sanderson, 2002; Segone, 2008; Krizek *et al.*, 2009; Davies, 2012; Head, 2016). Evaluation helps in supporting capacity building to better design, implement and value planning interventions; and in determining lessons which can be utilised to guide future planning (Faludi, 2000; Patton, 2002; Roberts, 2006). In summary, evaluation fosters continuous learning in planning, from promoting assessment of plans, and supporting the achievement of intended outcomes (Seasons, 2003; Head, 2016), to providing

constant support for improvement in the planning profession (Oliveira and Pinho, 2010b, 2010a, 2011).

2.4.3 Major Approaches

This review is not intended to include an exhaustive discussion of evaluation approaches. Rather, it draws together those elements which appear most appropriate as factors influencing the effectiveness of evaluation practice in the regional development planning process. This section discusses briefly the major approaches of evaluations in the planning process. This will help to identify and categorise the evaluation practices in the case study locations of this research.

Alexander (2016) identified two categories of evaluation in planning: by object and by timing. Based on the objects of evaluation, he distinguishes four kinds of evaluation which are applied in spatial planning: neighbourhood plans, city plans, regional plans, and strategic developmental and infrastructure plans; the latter can be at the local, regional, national, or multinational scale. Using the time dimension, he distinguishes between three kinds of evaluation: *a priori*, in progress, and *ex post*. Likewise, Oliveira and Pinho (2010a), stated that these evaluation types correspond to particular stages in the planning process. *Ex ante* evaluation occurs in the beginning of the planning process to promote a strategy or solution that best addresses the planning issues, plan goals and objectives from among alternative proposals (Khakee, 2003; Oliveira and Pinho, 2010a). “*Ex ante* evaluations are used as part of the planning process to refine program details” (Seasons, 2002b: 54). Common approaches in this category are planning balance sheet analysis (PBSA), goals-achievement analysis or goals achievement matrix (GAM), and cost-benefit analysis (CBA). *Ongoing* evaluation “takes place during plan implementation, and its conclusions can lead to shifts in the planning process (Oliveira and Pinho, 2010a: 347), or in other words, “judging the merits of the implementation process” (Chen, 1996: 124). Seasons (2002b) used the term ‘formative evaluation’ and states that it is usually undertaken in mid-plan cycle as a way to assess and modify delivery. Chen (2005) suggested the use of qualitative data especially through interviews and observations as they can provide rich and in-depth data about the implementation process. In addition, he classifies process evaluation into two types which are the identification of

implementation issue (also known as the ‘development-oriented process evaluation’) and examination of efficiency (otherwise known as ‘assessment-oriented process evaluation’).

Ex post evaluation, or usually referred to as summative evaluation, “occurs at the end of the plan implementation process and focuses on the impacts of the plan” (Oliveira and Pinho, 2010a: 347). *Ex post* evaluation is mostly undertaken when the plan is fully implemented to check whether the stated objectives and goals have been achieved, and the knowledge from this evaluation is usually used to improve further planning (Khakee, 2003; Laurian *et al.*, 2010). *Ex post* evaluation approaches include the conformance and performance-based approaches.

A *conformance-based* evaluation judges whether the objectives of the plan have been realised. Laurian *et al.* (2004) explained that the focus of this kind of evaluation is on the plan outcomes that are usually examined from the linkages between plans and the actual development. In contrast, the *performance-based* approach focuses on the planning process and does not consider the plan to be a blueprint but only acts as a guideline for practice (Mastop and Faludi, 1997; Faludi, 2000; Alexander, 2016). Plans are considered successful if the policy maker utilises them as a guide on a regular basis (Mastop and Faludi, 1997; Hull *et al.*, 2012).

2.4.4 Utilisation

There is a large body of literature on evaluation utilisation. It can be considered as one of the main indicators of a good evaluation. Somehow evaluation is valuable only if it is intensely utilized and its success is measured by the extent to which the findings are used to make changes (Pearsall, 2013). Evaluation utilisation is assessed by the extent to which appropriate data were evaluated and used to inform decision-making and resource allocation (Bamberger *et al.*, 2011). Effective use of evaluation in development agencies is one of the most crucial management facets which contribute towards performance of development programs (Kusak and Rist, 2004).

Since the 1970s, there have been concerns regarding the actual utilisation of evaluation findings (Weiss, 1972; Alkin *et al.*, 1979). Several studies tried to synthesise literature on evaluation utilisation, where Shulha and Cousins (1997) published a review of publications on this topic from 1986 until 1997 and Johnson *et al.* (2009) reviewed literature from 1986 to 2005. Recent publications such as Ramírez and Brodhead (2013), Cousins and Bourgeois (2014), Højlund

(2014), Loud and Mayne (2014), and Eberli (2016) indicate there is still strong interest in utilisation of evaluation because adequate use of monitoring and evaluation processes and findings is still lacking, and more often than not ends up gathering dust on shelves (Forss *et al.*, 2002) or disappearing in bureau drawers (Patton, 2008). This is also the case in the planning area, where Alexander (2016) stated that evaluation has remained relatively under-used and overlooked in planning practice, although there is evidence of increasing interest in the subject.

2.4.5 Actors

As discussed in the previous subsections, evaluation roles and its utilisation are highly influenced or even shaped by the key actors or stakeholders involved in evaluation practices. Therefore, it is important to identify the relevant actors with respect to evaluation practices in regional development planning.

There are varied actors of evaluation practices in regional development planning. As Scharpf (1997: 43) stated, “Actors are characterised by their orientations (perceptions and preferences) and by their capabilities”. Actors usually orient their actions based on a common set of norms and values or shared systems based on knowledge and belief (Scott, 2014). Scharpf (1997) further distinguished actors by whether they are individual or composite. Individual actors, the author argues, are expected “to follow the rules adopted by central decision processes and the hierarchical directives of the leadership” (Scharpf, 1997: 56). Furthermore, composite actors were designated separately as collective actors or corporate actors. Collective actors, such as clubs, coalitions, movements, and associations, are expected to follow the guidelines and preferences of their members, whereas corporate actors possess a high degree of autonomy in decision-making due to their independence from stakeholder oversight. Corporate “activities are carried out by staff members whose own private preferences are supposed to be neutralised by employment contracts” (Scharpf, 1997: 54). Corporate actors generally comprise a group of people who are employed, formally organised, and typically “top down” in their management style, while also being structured by formal rules and capable of purposeful action. Therefore, since strategy choices are decoupled from individual preferences, corporate actions are undertaken mainly for the purpose of benefiting the organisation.

Based on the descriptions above, it is suggested that the main actors in this research are the corporate actors such as agencies, evaluation units, research organisations. Seasons (2002b) identified the main actors in planning and evaluation at the local level as being the internal and external evaluators, and also the internal and external users. He further identified the evaluation planning department staff as the internal evaluators and the professional staff as the external evaluators. For the users, he considered the program unit (head, managers, and staffs), politicians, senior administrative staff, other departments within the organisation, and other units, agencies, or departments elsewhere in government as the internal users. Meanwhile, the residents were considered as the external users. These designations are in line with other studies which have identified that the evaluation department or the evaluation unit as the main actor (Loud and Mayne, 2014; Olejniczak *et al.*, 2016). Olejniczak *et al.* (2016) further explained the extended role of these evaluation units as knowledge brokers. However, the program managers, policy makers, politicians, and even researchers are also considered as important actors since they often drive evaluation (Visser *et al.*, 2014) and are engaged in the knowledge production process (Olejniczak *et al.*, 2016).

From the information in this section, it can be concluded that there are a variety of different actors of evaluation practices in regional development planning. Therefore, this research aims to involve the identified main actors in the process of gathering data from the fieldwork, which will be explained in more detail in Chapter 3.

2.4.6 The Influencing Factors

As noted in the results section of the systematic scoping review, there is no study in the current literature that has defined explicitly what an effective evaluation in planning is. However, several studies have identified factors that influence an effective or good or successful evaluation. Drawing on these studies, it is proposed that evaluation practice can be effective if the obstacles when developing, conducting, and communicating the evaluation are well addressed. Based on the iterative literature review process from the initial scoping review (described in the beginning of Chapter 2) until the end of this study (including the discussion in Chapter 7), the influencing factors which are discussed are resources, organisational culture, political factors, indicators, assessability of plan contents, stakeholder involvement, commitment, incentives, awareness,

appreciation, compliance, and understanding. The following paragraphs present the influencing factors identified in the literature review results. Considering this is a realist study, these factors are categorised into elements of context and mechanisms. Resources, organisational culture, availability of indicators, and assessability of plan contents are categorised as elements of context. Stakeholder involvement, commitment, incentives, awareness, appreciation, compliance, and understanding are categorised as mechanisms. Furthermore, this research categorised political factors into two different categories: supportive political system as part of the context and supportive political will as part of the mechanisms.

One of the most discussed influential factors in the literature is resources. For example, Seasons (2002b: 45) argued that “to be effective, evaluation processes must have sufficient resources, including properly trained staff, financial resources, and technical support for evaluation”, while Olejniczak (2013: 1662) mentioned that the volume of available resources, the number of institutions involved, and the level of public interest created a significant support for development of an effective evaluation system. In addition, Waldner (2004) stated that evaluations are costly and besides the actual financial costs, there are also opportunity costs of competing priorities and projects. She further pointed out that planners are often involved in the “front-loading” of resources where most of the resources are allocated towards the plan development while the evaluation of plans only receive the remaining resources. In this study, these resources are categorised as part of the context which influence effective evaluation practice.

Another factor is the technical obstacle, where there is a lack of a well-developed body of evaluation theory and methodology for planners so they would neither know how to conduct the evaluation nor learn how to place it in an appropriate context (Waldner, 2004). Especially with respect to evaluation of plan outcomes, there are limited guidelines on how to measure the success of a plan, both in terms of outcomes or implementation (Berke, Backhurst, *et al.*, 2006; Oliveira and Pinho, 2010a).

Another factor is the organisational culture. Organisational culture is “the collection of relatively uniform and enduring beliefs, values, customs, traditions, and practices shared by the organisations’ members and transmitted from one generation of employees to another” (Marcic, 1992: 184). In the planning context, it refers to the level of support from senior managers and politicians and also the attitudes of staff towards planning evaluation (Seasons, 2002b, 2003;

Oliveira and Pinho, 2010a). The organisational culture needs to value the benefits of evaluating plans and their outcomes in order for evaluation to be recognised as important functions in planning agencies (Guyadeen and Seasons, 2018). Moreover, planners need to be familiar with and value evaluation in order for it to play its role optimally. However, as Seasons (2003) noted, evaluation in planning agencies has often been considered discretionary due to the change-averse nature of public sector organisations. To overcome this issue, Oliveira and Pinho (2010a) suggested the promotion of an evaluation culture within the planning agencies or similar departments which undertake planning activities. Therefore, analysis and discussion of organisational culture in this study is strongly linked to evaluation culture.

Indicators or causal attribution is another influencing factor based on the literature. Evaluation is expected to establish the causal links between a plan's inputs, goals, outputs, and outcomes. Indicators play an important role in assessing these links. Laurian *et al.* (2010) stated that for evaluations to be effective, empirical evidence is needed including the selection of appropriate indicators that link the objectives of the plan with the outcomes. However, Seasons (2003: 435) noted that “a gap seemed to exist between the wealth of knowledge about indicators and their use in many regional planning departments”.

Assessability of plan contents is also considered as a significant influencing factor. Waldner (2004) mentioned that some of the more subjective plan goals might be difficult to evaluate. Based on study results, Seasons (2003: 435) also found that “the municipalities' plans were not readily conducive to evaluation because the goals, objectives, and policies were vaguely worded”. In this study, this factor is considered as part of the context.

Political factors are also considered significant. “The creation of plans is inherently a political process because politicians use plans to garner public support and, more importantly, elected officials are usually the decision-makers in planning” (Guyadeen and Seasons, 2018: 10). Politicians and planners may be concerned that evaluation will reveal shortcomings or failures that reflect poorly on their competency and political decisions (Waldner, 2004; Laurian *et al.*, 2010). Political factors also have a strong effect on the stability of public institutions. This stability is needed for evaluations to be effective as it “allows the accumulation of experience on the demand side (that is institutions that contract out studies), it increases the methodological quality of studies and, over time, the institutional capacity for strategic thinking” (Olejniczak, 2013:

1662). As previously mentioned, this research classified political factors into two different categories: supportive political system as part of the context and supportive political will as part of the mechanisms.

Stakeholder involvement is another influential factor that has been widely discussed in the evaluation literature. Johnson *et al.* (2009), Patton (2008), Mertens (1999), Fetterman (1994), and Deane *et al.* (2020) have all discussed the importance of stakeholder involvement or stakeholder participation, although their discussion of evaluations was more in general and not specific to the regional development context. Other literature has discussed stakeholder involvement although often not specifically using that exact term. For example, Sanderson (2000) showed his concern for stakeholder involvement using the term: stakeholder participation. Sanderson (2000) discussed two types of stakeholder involvement. The first, involves a more or less explicit value stance in favour of using evaluation processes to empower participants, to encourage the development of democratic values and to build capacity for participant self-evaluation. In the second, the author states that evaluations should identify and elaborate the values of all stakeholders in a policy or program and identify the implications and effects for each. Although Sanderson (2000) did not discuss evaluations in the regional planning area, evaluation in the context of complex social policy systems is discussed where the regional planning context can be considered as one of the complex social policy systems. This complexity is due to the potentially very wide range of stakeholders involved (especially in cross-sectoral initiatives) and the considerable scope for value conflict in relation to the underlying issues.

The next influencing factor is commitment. Several studies in the literature, such as Seasons (2003) and Widmer and Neuenschwander (2004) discussed commitment although they did not use the exact term specifically and did not discuss it in detail, therefore this factor was not identified in the initial literature review. Seasons (2003) briefly discussed commitment using the term demonstrable support from senior management and politicians. Widmer and Neuenschwander (2004) also discussed it very briefly by stating that decision makers should be committed in acknowledging that defining the purpose as well as the intended utilisation of an evaluation is essential.

Olejniczak (2013) revealed the importance of incentives as an enabling factor, stating that where low incentives existed there was a significant negative effect on the quality of products and low

utilisation of evaluation results. This aligns with Waldner (2004), who mentioned that there is often a lack of economic incentive to evaluate the plans. In this study, incentive and disincentive is perceived as part of the political will to support evaluations.

Literature also shows that awareness about evaluation is needed. Stockmann *et al.* (2020) noted that low awareness of evaluation and its potential benefits might be one of the factors that hinders regular utilisation of evaluation in the decision-making process. Additionally, Saunders (2012) and Hayton (2015) also briefly mentioned that awareness of or concern for evaluation and implications of the evaluation are needed although they did not provide further explanations about this.

Appreciation is considered as one of the first requisites to useful evaluations (Chelimsky, 1994). Additionally, Weiss (1991) as cited in Alkin (2013) states that an appreciation for evaluation is a political activity. The author recognises that the process of conducting an evaluation affects a political situation in which there are vested interests, negotiations, constituents, and critics.

Likewise, compliance is perceived as one of the important factors to be considered. McDonald *et al.* (2003) stated that making evaluation mandatory could promote a culture of token compliance, but voluntary adoption is much slower to take effect. Furthermore, Davies (1999) stated that evaluation cannot be forced on people because attempting to impose it will likely lead to goal displacement, unreliable information and an increase in the risk that programme relevance will be diminished rather than augmented.

Understanding is identified as one of the important factors in the fieldwork for this study. However, existing studies did not discuss this term specifically. Seasons (2002a) used the word ‘understand’ or ‘understanding’ in his argument that it is important to understand the integration of quantitative and qualitative research, as well as triangulation of research methods as one of the facilitating factors for effective monitoring and evaluation in municipal and urban planning. Seasons also stated that a sufficient understanding of how to articulate or communicate the purpose of, and approaches to evaluation in publications is needed for evaluations to be effective.

It is important to note that not all identified influencing factors from the empirical data in this study were discussed in detail by the authors in the existing literature. For example, regarding consensus as a factor, Widmer and Neuenschwander (2004: 404) only stated “Evaluation activities

in the federal administration are very varied. There is no consensus about what evaluation is or should do.” There is no further explanation regarding consensus in their article which is identified as one of the mechanisms in this study. Although Hayton (2015) mentioned the importance of acceptance by both evaluators and policy makers that evaluation is very inexact, he did not discuss consensus clearly and in detail, and not in the same context as this study.

Synergy with experts is also one of the identified influencing factors from the fieldwork of this study. However, existing studies did not use the same terms and did not discuss this factor specifically in relation to evaluations in regional planning. Lasker *et al.* (2001) discussed partnership synergy in the health sciences discipline while Carden (2018) discussed the importance of relationship building for interventions in the policy arena, but not specifically in relation to evaluation.

Additionally, there was very limited literature in the policy and planning area which discussed confidence of evaluators. Therefore, this study had to find literature from other disciplines, especially psychology such as Luttrell *et al.* (2013) and Grimaldi *et al.* (2015), to obtain a relevant literature to discuss this factor and compare it with the empirical data from the fieldwork of this study.

From the above descriptions, it can be concluded also that various studies have identified a different range of influencing factors for an effective evaluation. However, only one study by Seasons (2003) investigated the evaluations in regional planning context. The focus of other studies is more on the national or project context, or evaluations in general. Additionally, none of the identified studies have determined the relationships among the different influencing factors and they all highlight different factors as particularly important. Hayton (2015) and Carden (2018) used the realist approach but their focus was not specifically on evaluations in regional planning. Hayton’s focus was more on the extent to which concepts such as realistic evaluation can inform both evaluators and policy makers by evaluating specific interventions administered by one of the public sector agencies in Scotland, while Carden’s focus was on interventions in general in the policy arena.

In summary, existing evaluation practice studies have identified several significant factors which influence evaluations. However, examination of the wider public policy literature, which has been discussed in this chapter, suggests that other factors may be important. In order to identify such

factors, this study needed a more open-ended approach that allowed for a wide range of issues to be raised. Therefore, it was decided that a realist methodology to investigate all significant influencing factors, referred to as contexts and mechanisms in a realist perspective, was required to explore and identify the underlying reasoning and resources involved in producing effective evaluation practice. Additionally, Hayton (2015) suggested that affecting factors highlighted by realistic evaluation, including mechanisms, context, and the outcome configuration should be recognised and accepted by both policy makers and evaluators to improve evaluation practice. Therefore, this research will fill the gaps in the literature and improve evaluation practice by exploring the mechanisms, context, and outcome configurations which influence evaluation practices in planning. A more detailed discussion of the literature related to each of the influencing factors, especially the ones which were categorised as mechanisms in this study, is provided in Chapter 7.

2.5 Effective Evaluation

Based on the systematic scoping review, it was concluded that effective evaluation in the regional development planning context has not been clearly defined. In order to define effective evaluation, it is important to identify the existing concepts of effectiveness, effectiveness criteria, and evaluation standards. These themes are described in the following subsections.

2.5.1 The concept of effectiveness

To understand how to enhance effectiveness of evaluation practice, the term ‘effectiveness’ must first be clearly defined. Effectiveness in general could simply be defined as ‘the degree to which something is successful in producing a desired result’ (Oxford online dictionary, see <https://www.lexico.com/en/definition/effectiveness>, accessed 14 June 2019). Similarly, effectiveness is defined as the degree of achieving the goals (Green and South, 2006) or the extent to which the output of an intervention contributes to the outcome (IOB Evaluation, 2009). The Development Assistance Committee of the Organisation for Economic Cooperation and Development (OECD) also defined effectiveness as “the extent to which the development intervention’s objectives were achieved, or are expected to be achieved, taking into account their

relative importance.” (OECD, 2002: 20-21). From all of these definitions, effectiveness can simply be said to be a measure to determine whether an intervention is achieving its expected objectives. However, based on a literature search across various disciplines, it is evident that effectiveness can be more complex depending on the related context. From the environmental assessment literature, four categories of effectiveness can be identified:

1. Procedural effectiveness. This considers the principles and practice of governing impact assessment processes (Sadler, 1996). Chanchitpricha and Bond (2013) identified several influencing factors for this category of effectiveness including policy framework, availability of resources, political context, public participation, and knowledge of the professionals.
2. Substantive effectiveness. Theophilou et al. (2010) stated that substantive effectiveness is relevant to performance and argued that it is demonstrated when changes are made to the policy, plan, or program being assessed. Chanchitpricha and Bond (2013) also identified factors that influence substantive effectiveness: mechanism in decision-making, regulatory framework, quality of the report, and public participation.
3. Transactive effectiveness. This category of effectiveness is usually achieved when resources are invested at the minimum level to achieve the set of objectives (Sadler, 1996). It means that proficiency in using resources to achieve the objectives should be examined to assess the effectiveness of the intervention (Baker and McLelland, 2003; Theophilou *et al.*, 2010).
4. Normative effectiveness. This is related to the achievement of normative goals (Baker and McLelland, 2003). These normative goals could be changes in organisations, institutions, culture and science that could affect decision-making (Cashmore et al., 2004).

These effectiveness categories will be adapted and explained in greater length in Chapter 3.

2.5.2 Effectiveness Criteria

In the literature, various criteria or elements are used to evaluate effectiveness. In the environmental assessment literature, there are comprehensive criteria under four dimensions of effectiveness: procedural, substantive, transactive, and normative (Chanchitpricha *et al.*, 2011).

For evaluating effectiveness of multilateral organisations, the OECD uses the following criteria: achieving development objectives and expected results, cross-cutting themes, sustainability of results/benefits, relevance of interventions, efficiency, and using evaluation and monitoring to improve development effectiveness (OECD/DAC, 2012). Martz (2013) proposes four other criteria for evaluating the effectiveness of organisations: purposeful, adaptable, sustainable, and harm minimisation. For evaluating effectiveness of public participation, Pagatpatan and Ward (2017) propose: public influence, consensus, increased understanding, improved quality of decision, and increased trust.

Those effectiveness criteria from various disciplines have some commonalities, although some differ significantly due to the divergent objectives of the intervention. Those commonalities are: relevant, efficient, achieve expected results, increased understanding, support decision making, participation of stakeholders. Another important point to highlight is that context is very influential in determining those effectiveness criteria. However, this research is mainly focusing on the effectiveness criteria from the environmental assessment (EA) literature in view of the similarities in nature between EA and evaluation practices in regional development planning context which are the focus of this research. In effect, EA is also a form of evaluation practice but specific to the environmental planning and evaluation context. In Chapter 3, there is a detailed explanation of the effectiveness framework in the EA literature which has been adapted for this study.

2.5.3 Evaluation standards

In the evaluation literature, Patton (2002: 11) stated, “when one examines and judges accomplishments and effectiveness, one is engaged in evaluation”. In the domain of evaluation in development planning, effectiveness cannot be easily described, identified, and measured, as it is complex, multidimensional and subjective. Therefore, it is important to take into account the range of stakeholders, objectives, interests, and issues involved. This research will aim to formulate a comprehensive yet specific definition of the effectiveness of evaluation practices in the regional development planning context by incorporating the perspective of stakeholders.

In the literature on evaluation, many standards are already established. These basically identify how the quality of an evaluation will be judged. One of the most popular standards is proposed by

the Joint Committee on Standards for Educational Evaluation (JCSEE) which were originally developed for the evaluation of educational programs but now are used more widely. The JCSEE standards are continuously developed by the Evaluation Checklists Project at the Western Michigan University Evaluation Centre (Stufflebeam, 2001). The standards include these groups of criteria: appropriateness of evaluation content, clarity of purposes, completeness and relevance, organisations, clarity of writing, reference and sources. Many organisations worldwide have adapted the JSCEE standards to suit their own context, such as the African Evaluation Association (AfrEA), the Uganda Evaluation Association (UEA), and the International Program for Development Evaluation Training (IPDET). Standards have also been developed by the American Evaluation Society, the Australasian Evaluation Society, and other organisations. However, these standards usually focus on the evaluators' ethical standards while conducting evaluation, not on the evaluation per se. Therefore, these standards are not described further in this research.

Although a variety of evaluation standards have been developed globally, effectiveness of evaluation has not been well explored. However, some elements of those standards can be adopted in defining effectiveness of evaluation. These are: relevant, useful, valid methods, and appropriate communication. For example, based on the JSCEE standard, relevant means that all content is pertinent to what users need to do to complete the tasks (Stufflebeam, 2001). These elements can be utilised in developing the definition of effective evaluation that will be explored more in detail in the following subsections.

2.6 Realistic Evaluation

Realism is a transdisciplinary approach. Pawson (2018) has illustrated the transdisciplinary nature of realism by constructing a realist family tree which can be seen in Figure 2.4. This tree illustration presents some of the main contributors to realist thinking where Pawson divided them into four areas: philosophers, sociologists, evaluators, and other disciplines. This research draws substantially on the evaluators area, especially the work of the evaluation researchers Pawson and Tilley (1997b). However, as can be seen in Figure 2.4, Pawson is categorised as both sociologist and evaluator. This indicates that these four branches or areas are not totally different, they are connected with each other. Therefore, although this research mainly based on the work of Pawson

and Tilley (1997b), contributions from other realist authors from those four areas are referred as well.

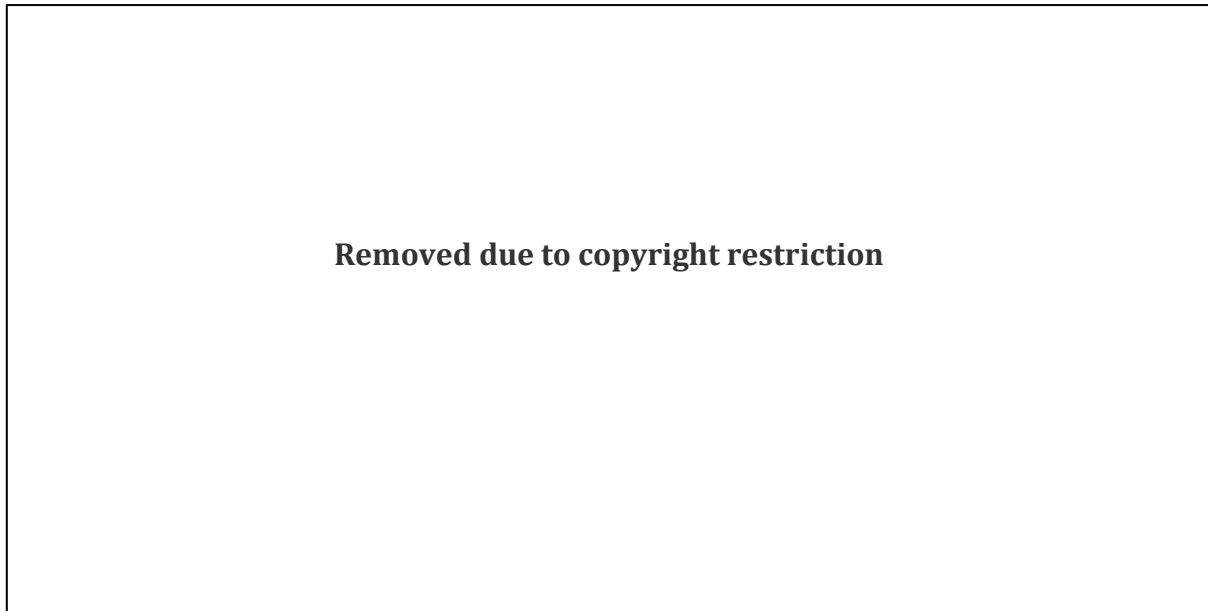


Figure 2.4 Realist Family Tree

Source: Pawson (2018: 207)

The most central characteristic of realism relates to the distinction between ontology and epistemology (Maxwell, 2012). Ontology refers to theories of what actually exists and the nature of reality, whereas epistemology refers to theories of how we gain knowledge of what exists. Realism combines a realist ontology (the belief that there is a real world that exists independently of our beliefs and constructions) with a constructivist epistemology which believes that our knowledge of this world is inevitably of our own construction, and that there is no possibility of our achieving a purely objective account that is independent of all particular perspectives (Maxwell, 2012). Therefore, realism has sought to position itself as a model of scientific explanation that avoids the traditional epistemological poles of positivism and relativism (Pawson and Tilley, 1997a). Realism's key features are its stress on the mechanics of explanation and its attempt to show that certain explanatory strategies can lead to a progressive body of scientific knowledge (Pawson and Tilley, 1997a).

A stronger statement is made by Westhorp (2018) who holds that realist evaluation is not simply a type of evaluation, but a type of applied realism underpinned by an understanding of how the world is and how it works (ontology), and an understanding of the nature of knowledge (epistemology). Therefore, it is not just another tool in the toolkit but rather a fundamentally unique set of assumptions with real and persistent implications for how programs or interventions work.

2.6.1 The Main Concepts

A realist approach differs from typical evaluation research as it not only relies on the standard approach of assessment but rather focuses more on its connection with the subject and context and therefore elaborates on the importance of examining the mechanisms (Sayer, 2000). Therefore, a realist research explores the outcome of intervention, its relationship with the underlying mechanisms and elements of context that would activate the mechanisms. As a result, the Context-Mechanism-Outcome (CMO) configuration was constructed based on the identified elements of context which are considered to activate the specific mechanisms to operate and produce the outcome. Pawson and Tilley (1997b) have argued that realist research provides this CMO configuration to improve social interventions.

Figure 2.5 demonstrates the link between the CMO configurations as proposed by Pawson and Tilley (1997b). The diagram conveys the idea that an intervention is delivered through a specified context and that within that intervention are the mechanisms. The illustration suggests that certain context(s) trigger(s) and activate(s) the mechanism(s), hence producing certain outcome(s). Pawson and Tilley (1997b) describe this as a 'generative causation'. They also provide this simple illustration formula: $\text{Outcomes} = \text{Mechanisms} + \text{Context}$.

Context (C)

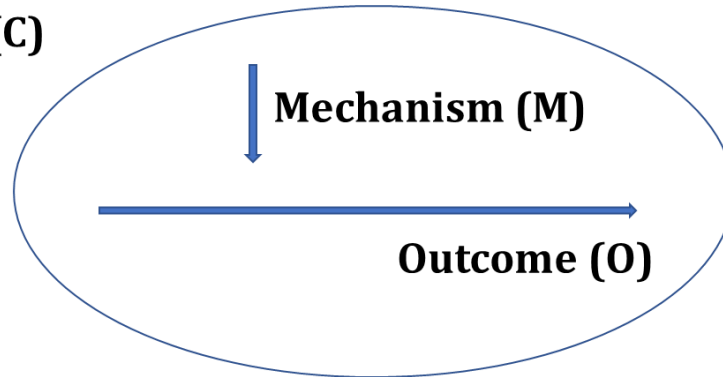


Figure 2.5 The Realist Causation Explanation of the CMO Configuration

Source: Adopted from Pawson and Tilley (1997b: 72)

In the most recent literature, Pawson (2018) clarifies that in a realist view, things are not C, M, or O, rather they function as C, M, or O in a particular part of the analysis. Pawson (2018) further gives an example referring to Westhorp’s explanation in Chapter 3 of the same book:

“I use the example of self-esteem. You might have a program which intends to raise students’ self-esteem. In that case, self-esteem is an **outcome**. Or you might have a program which intends to increase the proportion of disadvantaged students who complete high school, and works in part by raising their self-esteem. In that case, self-esteem is a **mechanism**. Or you might have a program that works best for students who already had high (or low) self-esteem. In that case, self-esteem is a **context**. And, of course, you can have a program which intends to raise self-esteem (**interim outcome and mechanism**) in order to increase school retention (**higher level outcome**) and which works best for students with moderate levels of self-esteem (**context**).” (Pawson, 2018: 210)

2.6.1.1 The Context

In a realistic evaluation, Pawson and Tilley (1997b: 216) defined context as “spatial and institutional locations of social situations together, crucially, with the norms, values, and interrelationships found in them”. Understanding the context in a realist approach means looking at the conditions within which mechanisms are activated and successfully achieved. This implies that realistic evaluation requires an understanding and explanation of “*for whom and in what circumstances* a programme works through the study of contextual conditioning” (Pawson and

Tilley, 1997b: 216). Recently, Pawson (2018: 212) stated that “contexts are most definitely not limited to location” and further explained that context may refer to any characteristic of:

- The individuals who partake in the programme;
- The interrelationships between stakeholders;
- The institutional arrangement into which the programme is embedded;
- The infrastructure, including the wider societal, economic and cultural setting of the programme.

2.6.1.2 The Mechanism

According to Pawson (2013: 115), “Mechanisms are agents of change. They describe how the resources embedded in a programme influence the reasoning and ultimately the behaviour of programme subjects”. According to Astbury and Leeuw (2010), mechanism is not demonstrating intervention activities and that it is sometimes not observable. Mechanism can be inferred from the “causal relationship”. To further explain, Astbury and Leeuw (2010: 368) depict mechanisms as the “underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest”. Astbury and Leeuw (2010) argue that mechanisms possess three elemental signs:

- 1) Mechanisms are usually hidden;
- 2) Mechanisms are sensitive to variations in context; and
- 3) Mechanisms generate outcomes

Additionally, Westhorp (2018: 45) argued that mechanisms are usually not observable considering:

- 1) They operate at different levels of the system than the outcome they generate;
- 2) They operate at different timescales than the outcome of interest; and
- 3) They necessarily depend on relationships and interactions between components, some of which can be observed but others cannot (or not with currently available instruments). For example, gravity can be observed through certain instruments but peer pressure cannot be easily observed.

To provide a clearer overview of the constructs of mechanism, Westhorp (2018) created a summary of examples that can be seen in Table 2.2.

Table 2.2 Constructs of Mechanism at Levels of Systems

Construct of mechanism	Material	Psychological/cognitive	Social-group	Social-institutional
Powers and liabilities	Trees can grow: e.g. photosynthesis	Humans can learn: e.g. sensitisation; motivation	Groups can make agreements	States can make laws
Forces	Gravity	Love	Peer pressure	Laws, regulations
Interactions	Gunpowder explosion	Reasoning and resources	Contracts	New technologies and market systems
Feedback or feedforward processes	Genetic inheritance	Developing attachment style	Negotiation	Stock market crash
Reasoning and resources	Neurons firing: electrical signals	Logic-in-use; affective response	Group think	Cultural assumptions

Source: Westhorp (2018: 49)

As shown in Table 2.2, constructs of mechanism vary. Mechanism can be in the construct of powers and liabilities, forces, interactions, feedback or feedforward processes, reasoning and resources, or another relevant construct. For example, forces (in the second row of the table) push or pull or otherwise exert pressure. Furthermore, Westhorp (2018) uses the table to explain and give examples of constructs of mechanism and how they may not be observable. For example, love, which is one of the constructs of forces mechanism (in the second row of the table), pushes and pulls humans to do extraordinary things; however, love is not the sole determinant of outcomes. This too is typical of all mechanisms and therefore all examples in the table. Mechanisms operate only when the circumstances are suitable, and always in concert and in competition with other mechanisms. This information is useful as guidance in the analysis in this research, since it can help explain the complexity and interaction of mechanisms in producing effective evaluation practice.

2.6.1.3 The Outcome

In a realist perspective, outcomes are the results from the mechanisms set in motion in different contexts. Pawson and Tilley (1997b) stated that an evaluator needs to identify the outcomes and understand how it had resulted in such a way. Outcomes can be a short-term or long-term consequence. By understanding the outcomes, an evaluator can suggest further improvement and modification of an intervention. The outcomes can be produced through certain mechanisms and appropriate contexts. Therefore, the CMO configuration is set in such a way as shown in Figure 2.5.

2.7 Conclusion

The systematic scoping review has given a reasonably comprehensive overview of the existing literature and the discourse about the effectiveness of evaluation practice in supporting regional development planning. The initial intervention theory that was formulated in this research was based on the results of this systematic review. The review concludes that there is no study that discusses and examines the effectiveness of evaluation practice in supporting regional development planning in the context of Asian developing countries. Considering that context is very influential in determining effectiveness of evaluation, filling the gap in the literature for the context of Asian developing countries is important. As can be seen in previous sections, literature about evaluation in planning has been growing significantly but there are still areas that are not well explored.

Guyadeen and Seasons (2016) stated that there is a need to further explore the institutional and political frameworks within which planning operates to identify the factors that support and inhibit plan evaluations. Likewise, Head (2016) also stated that there are major gaps in the literature on what occurs in public agencies in relation to evaluation activities. Therefore, the author highlights the importance of incorporating evaluation as part of research-based evidence into policy advice and regulatory activities. The review of the literature has also determined that prior studies have only identified the factors influencing evaluations, but have not taken into account or determined what are the underlying mechanisms and elements of context, and the interrelations between them. In addition, there is no clear identification in these studies of which factors have greater influence

and no explicit definition of effective evaluation practice. Moreover, clear information on institutional design, especially in the developing country contexts, is lacking. On the other hand, Hayton (2015) suggested that mechanisms, context, and the outcome configuration, as highlighted by realistic evaluation, should be recognised and accepted by both policy makers and evaluators in order to improve evaluation practice. Therefore, to improve evaluation effectiveness, this research will explore the contexts, mechanisms, and outcome configurations which influence evaluation practices in planning.

In this study, realistic evaluation theory is mainly incorporated in the analysis process. Realistic evaluation has been chosen in consideration of its principles in understanding the realities and relationships of interventions. Since this research is intended to examine the causal effect behind the effectiveness of evaluation practice, this evaluation method is argued to be the most appropriate approach for addressing the research questions because it can support analysis of "how, what, and for whom" questions that is important to arriving at conclusions.

CHAPTER 3 RESEARCH METHODOLOGY

This chapter presents the research methodology through a discussion of the realist-informed research, the intervention theory or CMO configuration that is developed and analysed further in Chapter 6 and 7. It also provides an overview of the adaptation of the effectiveness framework, which is used in Chapter 6, the case study selection, and the research methods for data collection and analysis employed in this research.

A realist case study design, which is heavily influenced by the realistic evaluation methodology developed by Pawson and Tilley (1997b), was selected to answer the questions of this research and better understand causal mechanisms within a given context that influence the effectiveness of evaluation practices in regional development planning. This is because what may work with some people in certain circumstances may not work in different circumstances or with different people (Kazi, 2003: 160). It was also particularly important that an appropriate methodology was selected that was sensitive to the contextual factors which influence effectiveness of evaluations in development planning and also appropriate to uncover the underlying mechanisms, as development planning is known to be very complex. Consequently, a methodology focussing on outcomes alone is likely to be a poor predictor of what would be effective in another context and would not highlight what processes might be effective in which types of contexts, whereas a realist approach would. In addition, the realist approach is also useful in identifying unobservable yet real mechanisms underlying an intervention and determine under which conditions they are triggered or fail to operate. Therefore, the realist approach is needed in order for the findings of this research to be used not just to identify the effectiveness or ineffectiveness of evaluation practices, but to develop the intervention theory to meet the needs of different contexts or circumstances and explain why the evaluation practice is effective (or not) and for whom. Furthermore, the realist approach adds value to the research in the form of generalisability, while the case study approach strengthens the realist approach by delineating the studied cases so that the initial intervention theory can be tested in all its dimensions within those cases. The overall research design of this research can be seen in Figure 3.1.

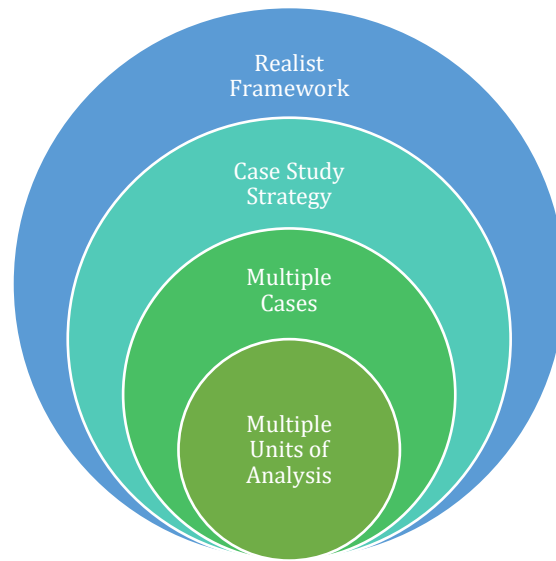


Figure 3.1 Research Design

Although the case study approach has its own benefits, it also suffers from limitations in establishing a causal link between intervention and outcomes (Christie et al., 2000). Therefore, combining the case study approach with a realist approach not only enhances the usefulness of case study, but also assists in addressing its limitations in terms of external validity. By making use of the Context-Mechanism-Outcome configuration, critical case studies can inform theory building (Christie et al., 2000).

3.1 Realist-Informed Research

Based on the conceptual framework and the literature review, realistic evaluation played a major part when assessing both case studies in this research. However, it is important to note that this study is a realist-informed research rather than a pure realistic evaluation, considering that in this research some aspects were done slightly differently compared to the features of realistic evaluation. This is based on the consideration that several modifications were needed in order to adapt the methods for achieving the purpose of this study. First, this research did not conduct a realist review but conducted a systematic scoping review. Usually, a realist review is conducted at the beginning of a realistic evaluation. However, this type of review is resource-intensive where several experts are needed to be involved and usually requires a minimum of six months full time.

Given that the main aim of the review was to inform this study about the gap in the literature regarding the topic of this research, then a systematic scoping review was perceived to be adequate to fulfil this purpose. Additionally, usually both quantitative and qualitative data are collected in a realistic evaluation. Considering the research question of this study, qualitative data collection was selected as the best option in order to get more detail information on context, mechanisms, and outcomes. Therefore, this research was only focusing on qualitative data and analysis, and not collecting quantitative data. The qualitative methods were essential in identifying the elements of context and mechanisms explained by the participants in the study. It is an exploratory study of intervention theories from literature and stakeholders, which is explored and expanded through qualitative methods of data collection and analysis.

Although it is a realist-informed study, this research is heavily influenced by realistic evaluation methodology developed by Pawson and Tilley (1997b). This methodology has a focus on generative causal explanation and also emphasises that evaluation research had a different cause from other social sciences, which is to have realistic ambition to inform real-world policy and practice (Pawson, 2013). Adopting methodology from (Pawson and Tilley, 1997b; Pawson, 2006; Pawson, 2013, 2018), three operational stages in this research have been identified, as shown in Table 3.1.

Table 3.1 Operational Stages

Stages	Operational Tasks	Purpose
Identifying intervention theory	Formulate theory on what is expected to work, in what circumstances to produce what outcomes based on existing literature and regulations.	To set the initial intervention theory
Testing the intervention theory	Data collection on contexts, mechanisms, and outcomes through interviews, focus groups, and relevant document gathering.	To interrogate the initial intervention theory
Refining the intervention theory	Analyse and interpret data: <ul style="list-style-type: none"> • Development of case study specific CMOs • Cross-case analysis to refine intervention theory 	To provide a refined intervention theory of what works, for whom, and in what circumstances

The first stage was development of initial intervention theory, as explained in detail in Chapter 1, and illustrated in Table 1.1. The second stage was testing the initial intervention theory by comparing it with the data from fieldwork as explained in detail in Chapter 6 where intervention theories for each case study location are presented under the headings of the three categories of effectiveness: procedural, transactive, and substantive. The third stage was refining the intervention theory, as described in detail in Chapter 7. Figure 7.1 illustrates the proposed comprehensive intervention theory, while Figure 7.2 illustrates the proposed intervention theory for the procedural effectiveness category. Table 7.3 explains the proposed intervention theory for the transactive effectiveness category, and Table 7.4 describes the proposed intervention theory for the substantive effectiveness category.

3.2 Intervention Theory or CMO Configuration

This research commenced with a systematic scoping review the results of which outlined the program or intervention theory. Considering the focus of this research is not a program, ‘intervention theory’ will be the preferable term in this research. In a realist approach, an initial intervention theory should be developed, and then used as hypotheses to focus the research question, select relevant data collection methods, and then guide the data analysis.

Rogers (2000) defined intervention theory as an explicit model of a program or an intervention which provides details about the mechanisms that are believed to contribute to the intended outcome. It can be analytically and empirically powerful and can lead to better research questions, answers, and interventions. Similarly, White (2009) and Masterson-Algar *et al.* (2014) described intervention theory as a set of assumptions which explain how the researcher foresees the intervention to achieve its expected outcomes.

As described in Chapter 1, the initial intervention theory in this research was developed in the form of sets of CMO configuration as can be seen in the Table 1.1. It is proposed that evaluation practice would be effective in supporting regional development planning if all the identified key mechanisms and elements of context are sufficiently present.

3.3 Effectiveness Framework Adaptation

In order to examine ‘what works and for whom’, it is first necessary to define what is meant to be ‘working’ or what constitutes an effective evaluation practice. This study adapts the effectiveness framework from an environmental assessment context into evaluation practice in regional development planning. It begins with some considerations in adapting this effectiveness framework and is followed by descriptions of effectiveness categorisations and their related criteria.

This research develops a framework drawing on the Environmental Assessment (EA) literature which categorises effectiveness into procedural, substantive, transactive, and normative effectiveness in order to achieve a better understanding of effectiveness of evaluation practice in regional development planning. I propose that adaptation of this framework is appropriate based on the following considerations:

1. A range of effectiveness frameworks do exist in the literature including the development effectiveness, organisation effectiveness, and Environmental Assessment (including Environmental Impact Assessment/EIA, Strategic Environmental Assessment/SEA and Social Impact Assessment/SIA) effectiveness. However, this research is mainly adapting the effectiveness framework from the Environmental Assessment (EA) literature in view of the similarities in nature between EA and evaluation practices in regional development planning context which are the focus of this research. In effect, EA is also a form of evaluation practice; however, it is specific to the environmental planning and evaluation context.
2. In the domain of evaluation in development planning, there is no effectiveness framework available. Additionally, this study argues that effectiveness of evaluation in development planning cannot be easily described, identified, or measured, as it is complex, multidimensional and subjective. Therefore, it is important to take into account the range of stakeholders’ perspectives and interests. It is also important to note that different actors will have different values regarding desirable outcomes. Consequently, effectiveness of evaluation depends on whose perspective is being considered and what this perspective entails. Therefore, it is essential to ensure this research incorporates these diverse perspectives.

3. As a consequence of the realist approach, understanding of the contextual factors is central to identifying the effectiveness of evaluation practice. Each effectiveness criterion is influenced by layers of contextual factors that could aid a more complete understanding of the effectiveness of evaluation.

The following sections discuss the four categories of effectiveness that incorporate the corresponding specific criteria of effectiveness that were derived from the evaluation literature and data from the fieldwork. The four dimensions of effectiveness are mainly adapted from Baker and McLelland (2003) which was built upon Sadler's (1996) three dimensions of effectiveness by introducing a fourth dimension, normative effectiveness. As mentioned by Loomis and Dziejczak (2018), it should be noted that those four dimensions of effectiveness are not mutually exclusive. All four effectiveness categorisations will be described in the following paragraphs. After describing all four effectiveness categories or dimensions, the criteria for each categorisation are formulated. Besides adapting the criteria from literature, most of the criteria are also confirmed by the field-level data.

First, procedural effectiveness: In EA literature, procedural effectiveness is defined as “examination of the practice involves finding out how the policy was applied or what procedures were used” (Baker and McLelland, 2003: 585). It will address the question about what procedures or principles were used and how they were implemented (Chanchitpricha and Bond, 2013). Therefore, in this research this category of effectiveness refers to the principles governing evaluation process, what and how the procedures are implemented. To measure this dimension of effectiveness, this study includes relevant procedures; integration of evaluation into the planning process; involvement of stakeholders in the process; delivery of results to relevant stakeholders; objectivity, clarity and understandability of evaluation results as the criteria under this effectiveness dimension.

Second, substantive effectiveness: In EA literature, substantive effectiveness is described as an examination of performance. It involves finding out what objectives were met as a result of the practice (Baker and McLelland, 2003). It also relates to the question of whether integrated decision-making is achieved or not (Theophilou *et al.*, 2010). Therefore, in this research, substantive effectiveness refers to the achievement of set objectives that include utilising the

evaluation to support accountability and informed decision-making, considering that these are the main roles of evaluation in planning as discussed in Chapter 2 of this thesis.

Third, transactive effectiveness: A number of studies currently available in the Impact Assessment literature look at transactive effectiveness from efficiency point of view. They define it as an “examination of proficiency involves finding out how resources were used in achieving objectives” (Baker and McLelland, 2003: 586) and the criteria are based on the efficiency concept. In this research, this category of effectiveness essentially means the delivery of the outcomes at minimum cost or resource use. It means that this dimension of effectiveness should be measured based on proficiency in resources use and time consumed during the process (Chanchitpricha and Bond, 2013). Therefore, this study includes resource management and specification of roles as the criteria under this effectiveness dimension.

Fourth, normative effectiveness: In EA literature, normative effectiveness means an “examination of the purpose involves finding out what normative goals are realised” (Baker and McLelland, 2003: 586), where normative goals are those which are derived from a combination of social and individual norms (Bond *et al.*, 2013). Normative goals are related to incremental changes in organisations, science, philosophy, and culture (Cashmore *et al.*, 2004). Additionally, Chanchitpricha and Bond (2013: 69) consider normative effectiveness as “the perceptions or attitudes that lead people to react or to take action in impact assessment processes, such that they can learn from the experience”. Therefore, in this research this category of effectiveness means the achievement of normative goals of evaluation practice. To measure this dimension of effectiveness, this study includes improvement to the plan or policy and development or changes in relevant institutions as the criteria under this effectiveness dimension.

In summary, the categorisations and criteria for evaluation practice in regional development planning context can be seen in Figure 3.2. This figure illustrates the effectiveness framework for evaluation practice in regional development planning context which is specifically developed and proposed in this study.

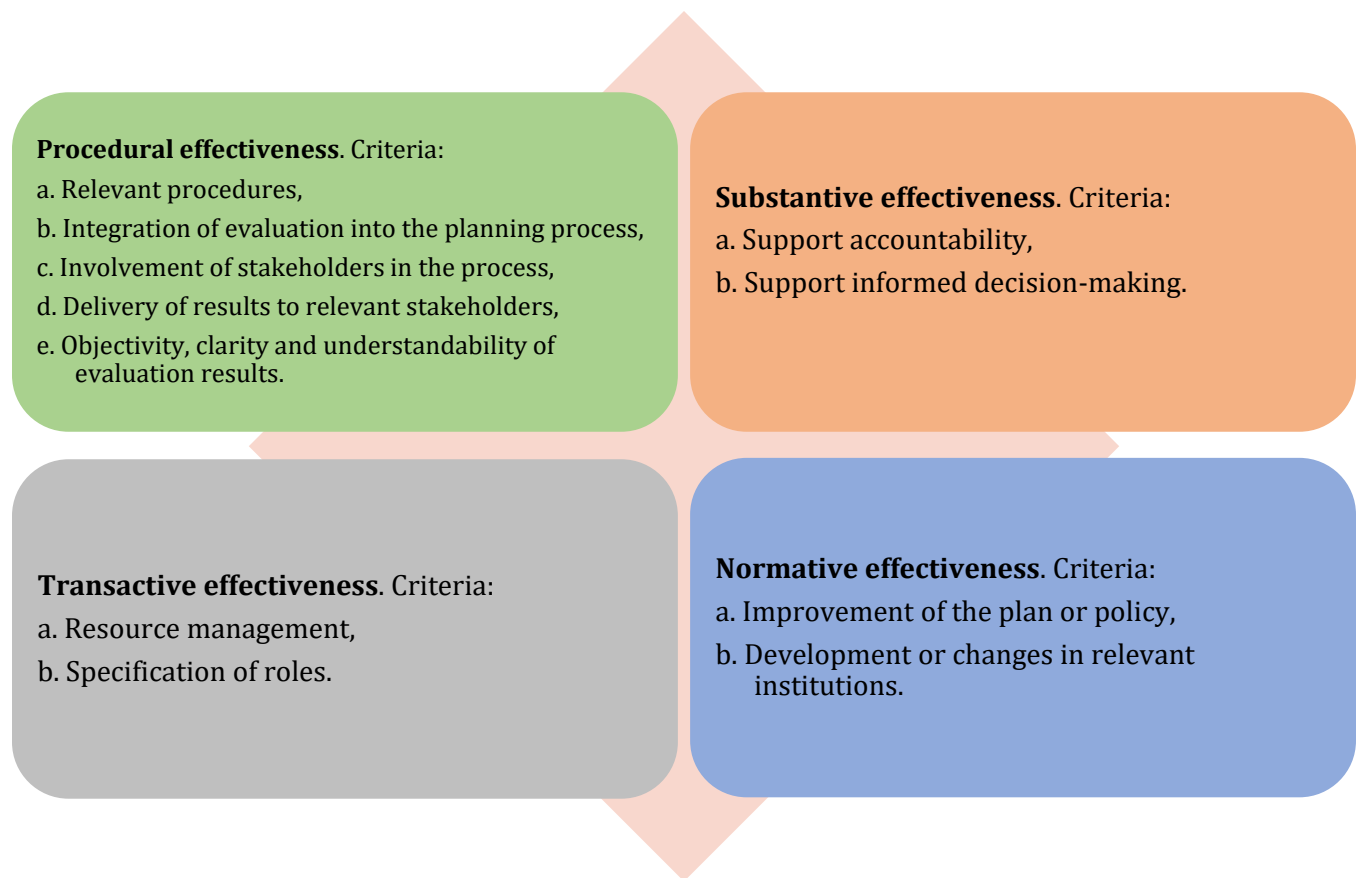


Figure 3.2 The Developed Effectiveness Framework for Evaluation Practice in Regional Development Planning Context

There are several criteria from the Impact Assessment literature that are not included in the Evaluation Practice Effectiveness Framework. This is due to the consideration that some of the criteria are redundant, and not all criteria are considered to be relevant in analysing effectiveness of evaluation practice in regional development planning context. For example, IA literature include “Identification of financial funds for SEA/IA practice” as one of the criteria under the Procedural Effectiveness dimension. However, this research considers that the identification of financial resources is expected to be included in “resource management” criterion under the Transactive Effectiveness dimension. Another example, “close collaboration” criterion, is considered to be one of the criteria under the Substantive Effectiveness dimension in the IA literature. However, in the present study this criterion has been covered in the “involvement of stakeholders in the process” which is under the Procedural Effectiveness dimension.

The effectiveness framework formulated in this study aimed to provide a comprehensive perspective to evaluate effectiveness of evaluation practice in regional development planning context. It is expected to be helpful in identifying causal factors which influence the effectiveness of evaluation practice by including multi-stakeholder perspectives (a more detailed analysis about this will be explained in the next sections). However, it is important to note that this study focuses mainly on procedural, substantive, and transactive effectiveness, considering that more extensive data and evidence are needed to analyse the normative effectiveness. However, it is not possible to collect those data in the timeframe of this study. Nevertheless, this framework is the first one to be developed in analysing effectiveness of evaluation practice in regional development planning context. Therefore, this study offers a starting point to stimulate further debate on effectiveness of evaluation practice, and invite other researchers to test and develop the framework further.

3.4 Case Study Selection

The case study approach was used because this research needs an extensive and in-depth description to explain the phenomenon chosen for investigation. It is an approach which is common in many fields of study, such as political science, sociology, public administration, anthropology, and urban planning (Gomm *et al.*, 2000; Yin, 2014).

To answer the research sub questions, two RDPAs were selected as cases in this research based on the following criteria:

1. Representing two categories in planning practice are best and middle-rank planning categories. The region with the best planning practice is considered to have higher compliance in fulfilling the requirements obligated by the regulations related with evaluation practice. In other words, this region is expected to have better evaluation practices in policymaking. Therefore, it is assumed that this region will provide valuable information on the supporting factors of an effective evaluation practice in planning. Because, despite the existing regulations stipulate that all regions must conduct evaluation of their own development plans, some regions in Indonesia still have not fully followed these rules. On the other hand, the region with the middle-rank planning practice is expected to provide valuable information on the factors impeding effective evaluation

practices in planning. This research will not include the poorest planning practice considering that the socioeconomic conditions at the provincial level differ considerably between the provinces with poorest planning practice and the provinces with best planning practice.

2. Representing one level of government is the provincial level. During the decentralisation era, the roles of provincial governments have changed several times. Under the Law No. 22 Year 1999 on Local Governments, the national government delegated authority to local government. Nevertheless, under the new law, the Law No. 32 Year 2004, the provincial government regained its role of prominence as the representative of the national government at the regional level. With regard to evaluation practice, the provincial governments also have a larger role where, in addition to evaluating their regional development plans, they are also required to evaluate the regional development plans of the local governments within their jurisdictions. Additionally, based on APN Award, better planning practices are produced at the provincial level than at the municipality level. Therefore, it is assumed this level of government will provide more valuable information on evaluation practices in planning compared to the local level, where capacities and resources for evaluations remain limited. Therefore, the provincial level was perceived to be more useful as case studies in this research.
3. The two selected regions have a similar socioeconomic background, including the level of economic welfare and Human Development Index. This is important to focus more on identifying the main factors, besides socioeconomic conditions, which influence the effectiveness of evaluation practice in supporting regional development planning.
4. The two selected regions were perceived as the best options for case study locations to answer the research questions of this study. This is based on the consideration where generalisation is not the main focus of this study. This study is focusing more on in-depth exploration.

The planning rank is based on Anugerah Pangripta Nusantara (APN) Award presented annually by the Ministry of National Development Planning/BAPPENAS of Republic of Indonesia. This award is given to the provinces/regencies/cities that excel in planning their local development. It is assessed on every regions' annual development plan, which is called RKPD (Regional

Government Work Plan), using 12 criteria, namely (A) linkage; (B) consistency; (C) completeness and depth; (D) measurement; (E) policy innovation; (F) bottom-up planning process; (G) top down planning process; (H) the technocratic planning process; (I) political planning process; (J) innovation of regional processes and programs; (K) display and presentation materials; and (L) presentation and mastery of the material (BAPPENAS, 2017). From these criteria, 16 indicators were developed. Two of these indicators are highly related with evaluation:

1. The **consistency between the evaluation results of the implementation of RKPD with strategic issues.**
2. Availability of regional economic framework, funding, regional finance policy, and description of strategy and direction of economic growth policy **based on evaluation of previous year.**

Based on the case study selection criteria, the APN award information, and the overview of the social welfare background, DI Yogyakarta Province and West Sumatra Province have been selected as case studies in this research. DI Yogyakarta was selected to represent the best planning category. West Sumatra was chosen to represent the middle rank planning category. However, it is relevant to note that the two regions have a significant difference in the political system. DI Yogyakarta Province has a Special Status which means that the traditional cultural leader of the region, the Sultan of Yogyakarta, has been permitted to be the governor of the province for life. Therefore, the political system of this province combines informal features of a pre-colonial sultanate with a modern administration based on formal institutions, such as laws written after Indonesia achieved independence. This difference is expected to provide additional value and perspective in the findings of this research.

Figure 3.3 shows the representation of case studies and Figure 3.4 shows the location of selected case studies.

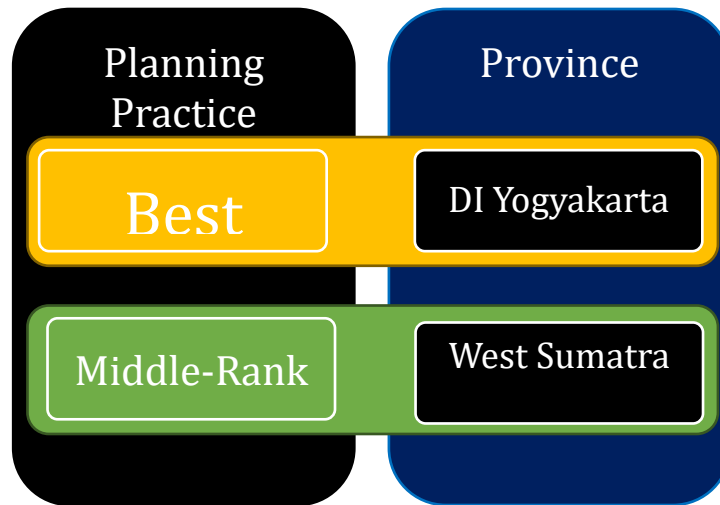


Figure 3.3 Representation of Case Studies

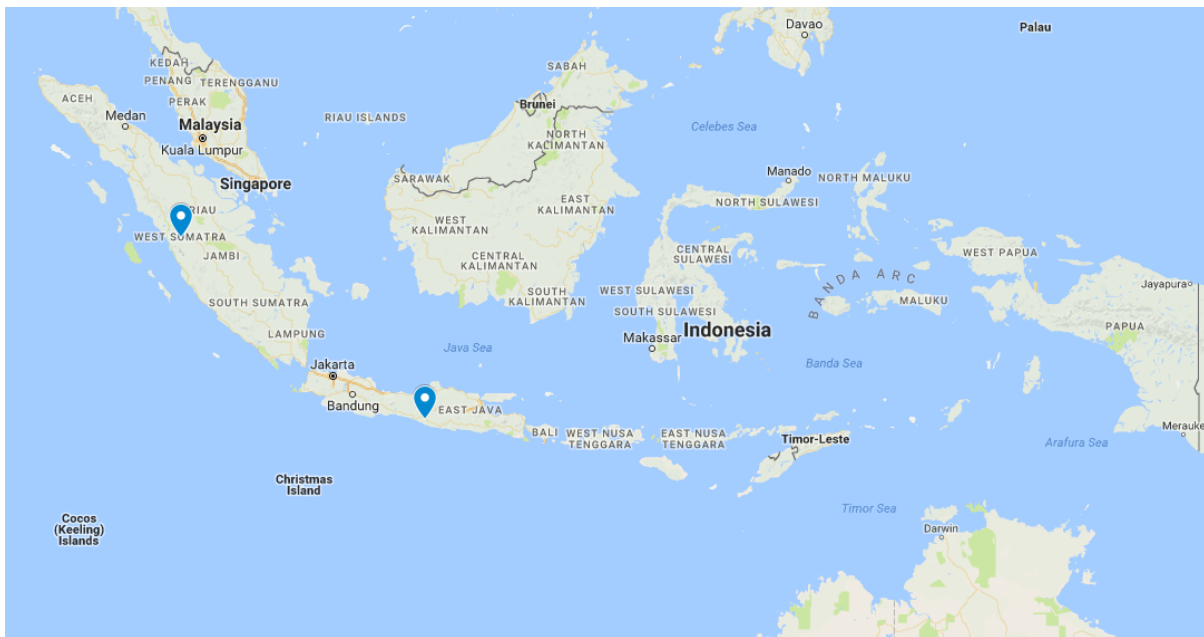


Figure 3.4 Location of Case Studies

Source: Modified from Google Maps, 2017

Multiple case studies improve the generalisation potential of the findings and provide insightful samples for the research (Christie *et al.*, 2000; Miles *et al.*, 2014). Inside the selected case studies, several units of analysis were examined. These units consist of different professional groups,

including officials from the top management level, and how they understood or viewed the intervention theory.

3.5 Data Collection

As both case study and realist approaches provide flexibility in selecting data collection methods, a mixture of methods was used. Four main data collection methods were employed to address the research questions: systematic scoping review, relevant document collection, semi-structured interviews, and focus groups. First, a systematic scoping review was carried out to strengthen the foundation of the research. Scientific publications, such as books and journals, were studied. Second, relevant documents, such as evaluation reports produced by RDPAs, policy, regulations, relevant official reports, and news articles, were collected to be analysed. Third, semi-structured interviews with key informants were conducted to obtain in-depth information. Finally, one focus group in each case study location was conducted involving key stakeholders to complement interview results.

Before conducting the fieldwork to collect the research data, ethics approval was obtained from the Social and Behavioural Research Ethics Committee (SBREC) at Flinders University. Then, a series of pilot interviews were conducted with respondents from different groups. Based on the results of these pilots, the interview guide was refined. Following this, the fieldwork was conducted in Jakarta to obtain the national level data, and then in West Sumatra Province and DI Yogyakarta Province to obtain the data from case study locations.

Interviews were conducted with the aim of obtaining detailed information on the effectiveness of evaluation practice from evaluation practitioners, policy makers, and experts. One particular interview technique employed was a semi-structured interview. This type of interview is a qualitative data collection method using a series of prearranged but open-ended questions (Ayres, 2008). It is useful for investigating complex perceptions and for collecting diverse experiences (Longhurst, 2009).

To use semi-structured interviews, it is important to develop an interview guide before the data collection (Ayres, 2008). In this research, the interview guide was made specific to the

respondent's category, with some probes. There were different interview questions for different categories of respondents due to different information needed from those sources. For example, from the evaluation practitioner category, the main information was related to current state of evaluation practices, institutional design, influencing factors, and challenges. From the policy maker category, the main information collected was related to institutional design, the role and utilisation of evaluation, influencing factors, and challenges. From the academic category, the main information was related to related theories, influencing factors, and challenges. From the local leader category, the main information was related to the role and utilisation of evaluation, influencing factors, and challenges.

The areas of question were classified into several main topics. These topics tended to differ from one interviewee to another, based on the background and position of the interviewee. The interview questions had a basic structure with similar subjects, but flexibility was applied here, so questions tended to be added or excluded based on the experience and response of the interviewee. Sometimes there were questions that needed to be added instantly following the flow of the interview where it seemed to be leading to more useful information for this research. However, the main broad topics of question that guided the interview were:

- The organisation context
- Personal experience related with evaluation practices
- Definition of effective evaluation practice
- The effectiveness of evaluation practices
- The elements of context
- The mechanisms
- Expectations and recommendations

The interviews ranged from one hour to two and a half hours in length. Before the start of the interview, the research information was introduced, including the consent form. Interviewees were informed that they could request to stop the interview at any time during the interview and that they could request the audio-recorder to be switched off any time they wanted to.

The interviews were conducted with key informants using purposive sampling. The NGO official participants in this study were selected based on their experience and knowledge related to

evaluation practices in Indonesia. Those selected at the national level, have either worked for international NGOs such as the World Bank or for national NGOs such as SMERU. Those selected in the case study locations, have had worked with local NGOs, and are familiar with the context and evaluation practices on the ground. In addition, there were officials from Central Bureau of Statistics (BPS) or the Regional Statistic Centre and academicians/experts from leading public universities in each selected case study area or surroundings with experience in the field of development evaluation for more than 10 years. In the fieldwork, the researcher discovered new information that the regional working units are one of the key stakeholders in evaluating the regional development planning. Additionally, from the fieldwork, the researcher also received information about another potential source of information: mass media. Therefore, representatives of several regional working units and mass media were also interviewed. The total representations of respondents can be seen in Table 3.2. Although the respondent categories in both case study locations were the same, the number of respondents for each category varied based on the process to reach the information saturation point.

Table 3.2 The Number and Representation of Interview Respondents

No.	Affiliation/Organisation	Respondents' Category	Total Number of Respondents
National Level			
1.	Ministry of National Development Planning	Intended users of evaluation & policy makers	6
2.	Ministry of Home Affairs	Intended users of evaluation & policy makers	3
3.	Academicians	Experts	3
4.	NGO	Other supporting stakeholders	4
Case Study 1 (West Sumatra)			
1.	Development Control and Evaluation Division in RDPAs	Evaluation practitioners	3
2.	Development Planning Division 1-3 in RDPAs	Intended users of evaluation	4
3.	Macro Planning Division in RDPAs	Intended users of evaluation	2
4.	Academicians from public universities in surrounding case study location	Experts	3
5.	Local leader	Other stakeholders	1

No.	Affiliation/Organisation	Respondents' Category	Total Number of Respondents
6.	Bureau of Statistics/Regional Statistic Centre	Other stakeholders	2
7.	NGO	Other stakeholders	2
8.	Mass Media	Other stakeholders	1
9.	Planning and Evaluation Division in Regional Working Units in RDPAs	Other stakeholders	3
10.	Local parliament members	Other stakeholders	2
Case Study 2 (DI Yogyakarta)			
1.	Development Control and Evaluation Division in RDPAs	Evaluation practitioners	2
2.	Development Planning Division 1-3 in RDPAs	Intended users of evaluation	3
3.	Macro Planning Division in RDPAs	Intended users of evaluation	2
4.	Academics from public universities in surrounding case study location	Experts	4
5.	Local leader	Other stakeholders	1
6.	Bureau of Statistics/Regional Statistic Centre	Other stakeholders	1
7.	NGO	Other stakeholders	1
8.	Mass Media	Other stakeholders	2
9.	Planning and Evaluation Division in Regional Working Units in RDPAs	Other stakeholders	4
10.	Local parliament members	Other stakeholders	1
TOTAL			60

In order to complement the interview results, one focus group was also conducted in each case study location. Focus groups are one method to collect qualitative data which involves a limited number of participants in a group discussion to express their views on a set of issues or a particular topic (Wilkinson, 2004; Gilbert, 2016). The focus group method can be used in conjunction with interviews (Wilkinson, 1998, 2004) to gather information on an underexplored topic where participants are encouraged to discuss their views (Parker and Tritter, 2006). In comparison to individual interviews, focus groups are usually more spontaneous and dynamic where group members engage and interact with each other in discussion about the subjects in focus (Alreck and Settle, 2004).

The duration of focus groups ranged from one hour to two hours. It was scheduled after individual interviews were conducted to follow-up or cross-check the interview results (Linhorst, 2002), to pursue an interesting finding or simply to add richness of the data (Wilkinson, 1998), and to avoid influence or bias in individual interviews. Open-ended questions were prepared for the discussion which were based on the research question, sub-questions, and initial findings from the interviews especially related with the intervention theory. Representative respondents from each respondent category that had been interviewed, except the local leaders and local parliament members, were invited to join the focus group. Due to the participants' busy schedule, the total number of participants of each focus group was around 2-5 participants. The participants were a mix representation of evaluation practitioners, policy makers, and academics.

3.6 Data Analysis

In this research, data analysis is categorised into separate techniques, which are document analysis, case study analysis, and realistic evaluation analysis. The secondary data collection, such as evaluation reports produced by RDPAs, policy, regulations, relevant official reports, articles, and scientific publications, such as journals and books, were the subject of document analysis. All of the relevant information was conceptualised, coded, and categorised accordingly.

For the case studies, the analysis of data commenced after recordings of interviews and focus groups were fully transcribed. The transcriptions, the researcher's notes, collected documents and regulations from the fieldwork formed the data set for this research. These data were classified into different themes and common ideas, as well as similarities and differences. Bias was dealt with by conducting a triangulation of interview data with FGD data and there was also a coding discussion with supervisors.

Individual case analysis and comparative analysis between case studies were conducted in this study. First, the individual case analysis was conducted between the interview data, to find patterns for a specific theme based on the interview question. Then, the similarities and differences of an individual case were identified. Second, a comparative analysis was conducted in which data from each case study was compared to find similarities and differences. A comparative or cross-case

analysis can enhance generalisability or transferability to other contexts (Miles *et al.*, 2014). Although generalisability is not the goal for this research, it is still important to know the relevance or applicability of this research findings to other similar settings.

In the analysis, the stakeholders' perceptions of effectiveness of evaluation practices in regional development planning in each case study location were framed under this effectiveness framework. Therefore, the Effectiveness Framework developed in this research is useful in defining what is effective evaluation practice and also determining the state of evaluation practices in case study locations while the Context-Mechanism-Outcome Configuration is useful in unpacking the elements of context and mechanisms that influence the effectiveness of evaluation practice. The CMO configuration framework offers tactical features in evaluating case studies in this study. In addition, this will also facilitate analysis in finding correlations between context, mechanism and results. Another important feature of the CMO configuration is that it allows investigations to identify the right mechanism in the appropriate context and allows this research to determine for whom the delivery process is most suitable and who will benefit from the intervention.

For configuring the intervention theories, a combination of analysis of interviews and FGDs, and then document analysis of documents and regulations were used to review and refine the intervention theory. NVivo 12, qualitative data analysis software was used to manage and analyse the research data. Coding was created to classify recurring themes, topics, and relationships. This classification was analysed further to identify relevant mechanisms and causal chains. The preliminary codes that were being created in NVivo were based on the previously discussed literature review, the initial intervention theory, and also questions in interviews and focus groups.

Data obtained in the first coding phase were used to develop CMO configuration patterns. As elaborated in Chapter 2, the realistic evaluation uses the CMO configurations to assess an intervention. This framework justifies the contexts, mechanisms, and outcomes of an intervention through its exploration of relationships and causal effects. The emerging patterns helped outline how the evaluation practices were being undertaken, what were the influencing factors, and what were the challenges. The refined CMO configuration patterns would then be examined to determine whether they supported or refuted the initial intervention theory. Based on this result, a refined intervention theory was constructed and discussed.

3.7 Conclusion

This chapter has explained the research methodology. It has justified its qualitative design and its strategies through the case study method and the realistic evaluation (using the Context-Mechanism-Outcome configurations). In relation to the research methods, this chapter elaborated on the various approaches to carry out the investigation which involved using secondary data, performing focus group discussions, conducting pilot and primary interviews. The Effectiveness Framework developed in this research will be useful in framing and determining the state of evaluation practices in case study locations, while the Context-Mechanism-Outcome Configuration will be useful in unpacking the elements of context and mechanisms which influence the effectiveness of evaluation practice. Therefore, the Effectiveness Framework and the CMOc will complement each other.

This chapter has also discussed how this research is slightly different from a pure realistic evaluation. The main difference is that usually both quantitative and qualitative data are collected in a realistic evaluation. However, in view of the research question of this study, this research was focused only on qualitative data and analysis. The qualitative methods were considered more useful and sufficient in identifying the elements of context and mechanisms pointed out by the participants in the study. It is an exploratory study of intervention theories from literature and stakeholders, which has explored and expanded through qualitative methods of data collection and analysis.

Before introducing the findings of this study, development planning and evaluation in Indonesia are explained in Chapter 4. Following this, Chapter 5 presents an overview of the case study locations.

CHAPTER 4 DEVELOPMENT PLANNING AND EVALUATION PRACTICES IN INDONESIA

This chapter presents an overview of development planning and evaluation in Indonesia under decentralisation. This is divided into two main sections. The first section describes the development planning in Indonesia which consists of two subsections: national development planning and regional development planning. Section two introduces the monitoring and evaluation systems and practices in Indonesia, including the regulatory frameworks that govern it.

4.1 Development Planning in Indonesia

Development planning is a continuous process in formulating development policies to achieve goals and objectives, taking into account all its potentials and limitations. In other words, the development planning process will produce several alternative possible policies that must be prepared in the future, so that in turn the target can be achieved by looking at all aspects of the potential and obstacles that may be faced. Development planning in general is a way or technique to achieve development goals appropriately, directly, and efficiently in accordance with the conditions of the country or region concerned (Sjafrizal, 2016).

After the Proclamation of Independence on August 17, 1945, Indonesia laid the foundation for the ideals of the state formation in the Preamble of the 1945 Constitution, which affirmed that the Republic of Indonesia will protect the entire Indonesian nation and its peoples, and seek to advance the general welfare, educate the population, and participate in the world order based on the country's independence, lasting peace, and social justice. Therefore, development must be carried out while ensuring that the intended objectives can be achieved with the sustainable management of available resources (Sjafrizal, 2016). Based on the Planning Law in Indonesia, planning is defined as a process to determine appropriate future actions by listing options and taking into account available resources. Accordingly, the development planning is important to accomplish.

Indonesia has been formulating and implementing development planning since its independence in 1945. However, it is important to note that development planning in Indonesia has experienced

a number of transformations. In the early years of independence, development planning could not be managed effectively due to the unstable political situation and poor state of the country after the end of World War II and occupation by the Japanese. For some years, the newly formed nation struggled to defend its independence against former European colonisers and, at the same time, also faced internal rebellion across many regions of Indonesia's widespread archipelago (Subkhan, 2014). Hence, a national development plan and formal policies for planning were not established immediately.

During the New Order regime, which began in 1966, the concept of the Garis-Garis Besar Haluan Negara (GBHN), the Supreme Guideline for State Policy, was introduced and used as the only guideline for development planning at central and local levels. The People's Consultative Assembly (MPR) endorsed the GBHN for a period of five years and the president, as the owner of MPR's mandate, operationalised it (Wasono and Maulana, 2018).

The change of governance model, from parliamentary to presidential, impacted the relationship between MPR and the president, and the GBHN was at one point removed from the 1945 Constitution. To replace the GBHN, the government passed the Planning Law No. 25 of 2004 (UU SPPN). The law stipulates that the National Development Planning System (NDPS) refers to a unit of procedures of development planning to produce a long-term (20-year) development plan/RPJP, medium-term (five-year) development plan/RPJM, and short-term (one year) development plan/RKP conducted by government agencies at all levels (Wasono and Maulana, 2018).

There are a variety of legal instruments for development planning in Indonesia. Some of the key regulations are:

1. Law No. 25 of 2004 on National Development Planning System (UU SPPN/Planning Law).
2. Law No. 23 of 2014 on Local Government (UU Pemda/Local Government Law, replacing Law No. 32 of 2004).
3. Law No. 17 of 2007 on National Long-Term Development Plan 2005-2025.
4. Law No. 33 of 2004 on Fiscal Balance of Central and Sub-National Governments.

Based on this regulatory framework, development planning can be mapped into four stages: 1) Plan formulation; 2) Plan formalisation; 3) Control and implementation; and 4) Evaluation. Table

4.1 summarises each stage of planning in Indonesia that has its respective outputs and characteristics.

Table 4.1 Planning Stages in Indonesia

Stage	Descriptions	
	National	Sub national
Plan Formulation	National Long-Term Plan (RPJPN)	Regional Long-Term Development Plan (RPJPD)
	National Medium-Term Plan (RPJMN)	Regional Medium-Term Development Plan (RPJMD)
	Agency specific Strategic Plan (Renstra K/L)	Regional Working Unit Strategic Plan (Renstra SKPD)
	National Development Consultative Forum (Musrenbang)	Regional Development Consultative Forum (Musrenbang)
Plan Formalisation	Law on RPJPN	Local Regulation on RPJPD
	Presidential Regulation on RPJMN	Regional head regulations on RPJMD
	Presidential Regulation on RKP and Renstra K/L	Regional head regulations on RKPD and Renstra SKPD
	Renja K/L	Renja SKPD
Plan Implementation Control	By specific heads of K/L	By specific SKPD Leaders
	Bappenas gathers and analyses result of development plan implementation monitoring from respective K/L pursuant to their scope of duties and authority	Bappeda/RDPA gathers and analyses result of development plan implementation monitoring from respective SKPD pursuant to their scope of duties and authority
Plan Implementation Evaluation	K/L evaluates development plan implementation from previous period	SKPD evaluates development plan implementation from previous period
	Bappenas outlines development plan evaluation based on K/L evaluation	Bappeda/RDPA outlines development plan evaluation based on SKPD evaluation
	Evaluation is used to inform subsequent development planning	Evaluation is used to inform subsequent development planning

Source: Law No. 25 of 2004 on National Development Planning System (NDPS)

More details about Indonesia's national and regional development planning are described in the following sections.

4.1.1 National Development Planning

One of the most significant transformations in Indonesia occurred during the Asian Financial Crisis at the end of the 1990s with the resignation of President Soeharto in 1998, which marked the beginning of the democratisation and rapid decentralisation process in the country.

Development plans also reflected the influence of radical political change in the Republic of Indonesia which presented obstacles for the government in implementing a consistent development strategy and allocating its scarce resources (Kimura, 2005). The Government of Indonesia summarised the priority areas in the 2000–2004 development plan or National Development Program, also known as *Program Pembangunan Nasional (Propenas)*: guaranteeing equal opportunities to workers and business people and protecting consumer rights; emphasising economic growth along with social values of equity, quality of life, and environmental protection; establishing a just market mechanism with healthy competition; (Booth, 2005). Booth (2005) criticised the Propenas plan for having overoptimistic targets, lacking specific policy initiatives to meet the targets, and no evaluation of past development plans. However, Booth (2005) also acknowledged that after a problematic democratic transition process, and economic catastrophe after the Asian Financial Crisis, planners needed to be optimistic for the longer term. Based on these considerations, the Government of Indonesia enacted Law Number 25 of 2004 on the National Development Planning System (NDPS) to improve the credibility of development plans. This law states that the national development planning comprises integrated planning across government agencies and departments; and across levels of government (Government of Indonesia, 2004). Moreover, the NDPS aims to:

1. Support coordination between development actors
2. Ensure integration and regional, spatial and time synchronisation, as well as central and sub-national administration governance function alignment
3. Ensure the link and consistency between planning, budgeting, execution and monitoring
4. Optimise community participation
5. Ensure that resources are utilised efficiently and effectively, fairly and sustainably.

This Planning Law also outlines the scope of the national development plan: The Long-Term Development Plan (RPJP) that sets forth the development vision and mission for the next 20 years; the Medium-Term Development Plan (RPJMN) that breaks down the RPJP into five-year plans; the Annual Work Plan (RKP) that elaborates the government's annual plan based on the RPJMN. The Ministry of National Development Planning (Bappenas) plays a central role in formulating all three national development planning documents (Wasono and Maulana, 2018).

Further, the government enacted Law No. 17/2007 on Long-Term National Development Plan (*Rencana Pembangunan Jangka Panjang Nasional* or RPJPN) for the period of 2005–2025. Implementation of the RPJPN is operationalised through a 5-year Medium-Term National Development Plan (*Rencana Pembangunan Jangka Menengah Nasional* or RPJMN). RPJMN serves as a basis for ministries and government agencies in formulating their strategic and budget allocation plans. Both RPJPN and RPJMN also serve as two key documents shaping current Indonesia’s regional development policy (Yağci and Ardiani, 2018).

In formulating regional development plans (*Rencana Pembangunan Jangka Menengah Daerah* or RPJMD), subnational governments are required to take into account both RPJPN and RPJMN in order to ensure alignment between central and regional governments. Both RPJMN and RPJMD are further operationalised through the Annual Government Work Plan (*Rencana Kerja Pemerintah* or RKP) at the national level and RKPD at the local level that will serve as basis in formulating Draft Government Budget (*Rencana Anggaran Pendapatan dan Belanja Negara/RAPBN* at the national level or RAPBD at the local level) (Yağci and Ardiani, 2018). An overview of Indonesia’s national development planning system can be seen in Figure 4.1.



Figure 4.1 Overview of Indonesia’s National Development Planning System

Source: Arnaldo Pellini (2018: 70)

Implementation of RPJPN, RPJMN, and RPJMD relies substantially on coordination across ministries and agencies, between central and local governments, which until now has remained a profound challenge.

4.1.2 Regional Development Planning

Regional development planning refers to the practice of preparing the development of an area which involves systematic activities and covers various development processes including plan making, implementation, and the evaluation of policies, regulations, as well as programs. It is a development planning process that is intended to make changes for the better for a community, the government and its environment in a certain area or region, by utilising existing resources, and must have a comprehensive orientation, complete but still adhering to the principle of priority (Riyadi and Bratakusumah, 2003). In Indonesia, regional development planning is perceived as a process for determining future policies, through a sequence of choices, involving various elements of stakeholders, in order to utilize and allocate existing resources within a certain period of time in the Region. It aims to improve the social welfare of the region, to improve access and quality of public services, and to improve the regional competitiveness (Minister of Home Affairs, 2017).

Regional development planning in Indonesia is expected to be aligned with the national development planning system. This means that the development carried out in regions is inseparable from the concept of the national development plan. Therefore, regional development plans and programs should keep referring to the national development plan, in both long and medium-term development plans. This is the approach used in the preparation of regional development planning in accordance with the Government Regulation No. 8 of 2008, which uses a combination of political, technocratic, participatory, top-down and bottom-up approaches.

Like the national development plan, regional development plans consist of long-term, medium-term, and short-term annual plans. Based on the Planning Law, regional development plans comprise the following:

- The Regional Long-Term Development Plan (RPJPD). This planning document has a long-term perspective. This document includes the local Vision and Mission and the local development directions for the next twenty years. In its drafting, this document refers to

RPJPN.

- The Regional Medium-Term Development Plan (RPJMD). This planning document has a medium-term perspective. This document outlines the explanation of Vision, Missions, and Programs of the Local Leader (governor at the provincial level, and mayor/regent at the city/regency level). After decentralisation, the heads of local governments elected by the people are the representatives whose leadership vision is appreciated by the majority of the voters. This vision becomes a political contract between the elected candidate and their voters. This is why this leadership vision is very influential in the medium-term regional development planning documents. In its drafting process, the RPJMD is guided by the RPJPD and taking into account the RPJMN. From the perspective of its substance, the RPJMD includes local development strategies, general policies on local development, local financial policy directives, programs within and across Regional Working Units (SKPDs), and regional programs.
- Annual Local Government Working Plan (RKPD). This document has an annual dimension, which is an elaboration of the RPJMD and refers to the National RKP. Substantially, the RKPD includes a local macroeconomic framework, local development priorities, a work plan and its budgeting framework. In the drafting process, the RKPD should give attention to the direction of data and information derived from RTRW as well as from the SKPD strategic plan (Firman, 2009; Hill and Vidyattama, 2016; Holzacker *et al.*, 2016; Arnaldo Pellini, 2018; Talitha *et al.*, 2020).

The NDPS mandates that the central and local government development plan products be synergised. Regional development priorities are not only focused on the implementation of regional government carried out in stages, but also how to translate wisely the ways to achieve the RPJMD goals. The RPJMD target itself is an elaboration of the regional head's vision, mission, and program that includes the objectives, targets, strategies, policy directions, regional development, and regional finances as well as regional tool programs and cross-regional instruments accompanied by an indicative funding framework for a period of five years.

Under the decentralisation era, the role of local governments in regional development planning is to have the authority and ability to manage and implement regional development programs. Additionally, the role of local governments is to determine the success of the implementation

process of governance and development activities in the area through monitoring and evaluation. Both the Planning Law and the Local Government Law regulate the responsibilities of regional governments in terms of policy formulation, medium-term planning, and annual planning. Within six months after a head of regional government is elected and sworn in, a medium-term development plan needs to be formulated by local regulation.

However, as stated by (Nasution, 2017), the decentralisation process in Indonesia has had inadequate preparation and been carried out with little consideration of logical objectives. One of the reasons is that the capacity of local governments to manage their own planning was not increased. Because of the long tradition of centralisation before the National Development Planning System was initiated, local government had not built the capacity to carry out planning (and evaluation) and undertake initiatives. Since the local governments had mainly functioned as implementing agencies of national policies and programs, they were unable to undertake the new task of planning and evaluation.

4.2 Evaluation Systems and Practices in Indonesia

One of the important components in the planning process is evaluation. As described in Table 4.1, Figure 4.2 presents the stages of planning in Indonesia where evaluation is an integral part of the planning process. Evaluation involves activities and processes to get feedback about the consistency between planning and implementation, and about the appropriateness of planning with the objectives to be achieved. Evaluation can also be used to determine the enabling factors and inhibiting factors in plan implementation. In addition, evaluation can also be used to ensure synergy between the central and local government development plan products.

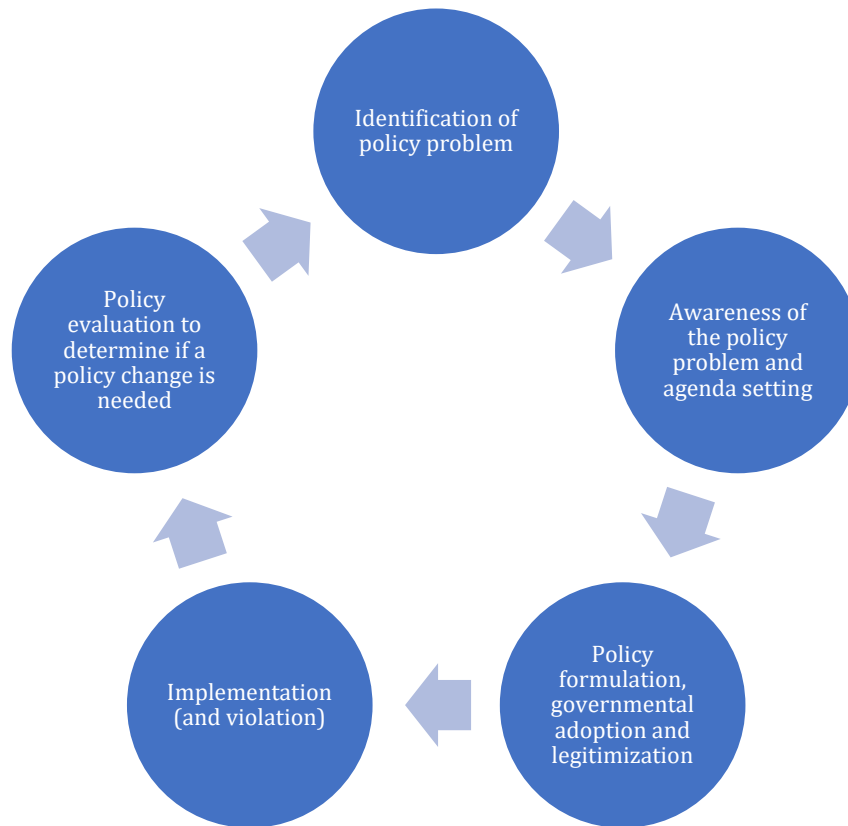


Figure 4.2 Stages of Indonesian Policy-making

Source: Shaxson (2016: 39)

There are several existing evaluation practices in the national development planning system in Indonesia. For the national level, the evaluations are mainly guided by the Regulation of the Minister of National Development Planning/Bappenas Number 1 of 2017 on Guidelines for Evaluation of National Development. For the regional level, the evaluations are mainly based on the Regulation of the Minister of Home Affairs Number 86 of 2017 on the Stages, Procedures, Monitoring and Evaluation of Regional Development Plan Implementation. However, these existing evaluations in Indonesia are quite complicated, inefficient and ineffective, overlapping, at times inaccurate, not timely, and to a certain degree unreliable (Stottele-Ishmi, 2017). Stottele-Ishmi (2017) identified that the parallelism duplication and overlapping of monitoring and evaluation applications in place is due partly to the widespread mistrust of the data reliability and underperformance of institutions. For a useful monitoring and evaluation system, the data must be reliable and accessible. Despite their importance, data reliability and accessibility, continue to be

an issue in Indonesia. Unfortunately, as noted by Stottele-Ishmi (2017), there is no mechanism in place to harmonise the data collected. Currently, several ministries, which are responsible for the substance of the data collected, collect data in parallel for the same indicator. Yet, the results from the data collection are not the same. There is no synchronisation of such data collected. There is a consensus that there should be no more than one Ministry collecting the same data and the results should be shared among all the Ministries interested in the results. For example, some of the regulations for existing online systems which need to be followed by the local governments in Indonesia are:

- Presidential Decree No 20 of 2015 regulates a system called TEPPRA (web based). Those in charge of this system are: KSP (*Kantor Staf Kepresidenan/Presidential Staff Office*), Ministry of Finance, LKPP (*Lembaga Kebijakan Pengadaan Barang Jasa Pemerintah/National Public Procurement Agency*), Bappenas, and Ministry of Home Affairs (MOHA). The authorities that must report include all Ministries and RDPA. The purpose of measurement is guarding the budget absorption at Central & subnational government.
- Government Regulation No. 39 of 2006 regulates a system known as e-money (web based). Bappenas is the responsible institution of this system. All ministries, RDPA (Bappeda), and working units receiving national budget must report under this regulation. The purpose of measurement is control and evaluation of development plans.
- Based on Government Regulation No. 60 of 2008 and Presidential Decree No. 192 of 2008, there is a system called SIMDA (database accountability management information systems). The authority in charge of this system is BPKP. Sub national government institutions are required to report under this system. SIMDA is an internal monitoring tool for measuring the work progress of subnational government.
- Based on Government Regulation No. 56 of 2005, there is a system known as SIPKD (web-based). The Ministry of Home Affairs (MOHA) is responsible for SIPKD and subnational governments are required to report under this system. The monitoring, controlling and evaluation of funding of decentralisation and subnational development budgets is the function of SIPKD.
- Government Regulation No. 70 of 2012 regulates a system named SIRUP (database). The responsible institution for this system is LKPP, while the authorities that must report are

ministries and RDPA. The SIRUP database is a measurement tool used to support RUP (the General Procurement Plan).

- Presidential Staff Office (KSP) Regulation No. 5 of 2015 regulates a system named SISMONTEP (web based). In charge of the SISMONTEP system is LKPP and those reporting are ministries and RDPA. The system monitors and evaluates Budget Absorption & Procurement in Goods/Services.
- Based on Government Regulation No. 60 Year 2008 and Decree of the Minister of Finance No. 95 Year 2012, there is a system named MONIKA. The Ministry of Finance is in charge of this system, while the Working Units of Ministry of Finance (1,085 in total) and RDPA are required to report to MONIKA. This system measures financial and asset performance.

Based on this regulations overview, it can be seen that there are many reporting systems involving many institutions, with the main purpose being to monitor only two fundamental aspects of government. Those key areas being physical performance and financial performance, such as budget realisation or absorption, which results in a great deal of data redundancy and overlapping of work activity. However, those systems are online so it is easier to check the compliance. Therefore, local governments (RDPA as the one in charge) tend to follow those regulations. In addition to those listed regulations, there are other regulations with the focus of measurement of local government performance where the local governments need to report their own performance. As a result, the local governments have many reporting obligations that they have to fulfill.

An umbrella policy of exchange is missing among the main actors, including Bappenas, Ministry of Finance, MOHA, LKPP, BPKP, and other Ministries, which makes it difficult for these actors to be connected through a mechanism for the purpose of data exchange. These government agencies or ministries have no authority over each other, except for KSP, that has the mandate and authority to request, coordinate, and elaborate the data from any of these institutions.

The absence of an umbrella policy leads to ineffective coordination among these ministries or agencies, which is resulting in the failure to produce outcomes as expected. Therefore, it is not easy to mend and merge the numerous existing evaluation systems, as it is most likely to be undermined by the poor coordination and the wide-spread understanding at government institutions that information is power. Consequently, many institutions are reluctant to share information which may be useful in retaining or increasing their relative power and influence. The

working culture built around this understanding does jeopardise any typically established coordination (Stottele-Ishmi, 2017).

4.2.1 Regulatory Framework

As previously explained, there are several monitoring and evaluation practices in Indonesia which are regulated by different ministries with different sets of regulations. Most of them are online reporting systems. The main government institutions that are in charge of those systems include Bappenas, Ministry of Home Affairs, Ministry of Finance, National Public Procurement Agency (LKPP), Ministry of Administrative and Bureaucratic Reform (Kemen PAN RB), and the Finance and Development Supervisory Board (BPKP). Those that need to report in the online systems also vary, including ministries, Local Government Working Units (SKPD), and the Regional Development Planning Agency (RDPA). The time frame of those reporting systems varies starting from monthly, quarterly and annually.

Most of the existing online systems only cover monitoring activities (some with limited evaluation aspects, such as e-money) where the focus is only on two aspects: physical performance, and financial performance (budget realisation or absorption). The main purpose of those systems is accountability to the central government institutions.

On the other hand, the focus of this study is evaluation practices which are expected to support regional development planning. Therefore, the main focus of this study is the evaluations which are regulated under the Planning Law (Law No. 25 of 2004) and Local Government Law (Law No. 23 of 2014). This is based on the fact that these are the two laws that are concerned with monitoring and evaluations to support regional development planning. Other regulations are focusing more on monitoring of public expenditure and the procurement process, and the main users are the central government institutions.

In the Planning Law, it is stated clearly that the RDPA is responsible for evaluations at the local level, and the evaluation results should become material for the preparation of the regional development plan in the next period. These are regulated in the Article 29 of this Law:

(Section 1) Heads of Ministries / Institutions evaluating the performance of the implementation of the development plans of the previous Ministries / Agencies;

(Section 2) The Leadership of the Regional Working Units conducted an evaluation of the performance of the implementation of the development plan for the previous period;

(Section 3) The Minister / Head of Regional Development Planning Agency prepares an evaluation of the development plan based on the results of the evaluation of the leadership of the Ministry / Agency as referred to in section (1) and the evaluation of the Regional Work Units as referred to in section (2);

(Section 4) The results of the evaluation as referred to in section (3) become material for the preparation of the National / Regional development plan for the next period.

There are numerous regulations under the Planning Law and the Local Government Law that are used as the basis for evaluations in regional development planning. Under the Planning Law, there is Government Regulation No. 39 Year 2006. In line with the Planning Law, this regulation explains that in the current planning system in Indonesia, the development planning stages consist of four stages including evaluation. Planning, implementing, controlling activities, and evaluation of the implementation of the plan are part of the function of management, which are interrelated and inseparable from one another. The four stages complement each other and each provides feedback and input to others. Good planning is needed to ensure a smooth implementation of the plan. Correspondingly, in order to improve the efficiency and effectiveness of resource allocation, as well increase transparency and accountability of program management development, it is necessary to control and evaluate efforts on the implementation of the development plan.

Government Regulation No. 39/2006 also explains that evaluation is done with the intention of being able to find out with certainty whether the achievement of results, progress, and obstacles encountered in implementing the development plan can be assessed and studied to improve the implementation of the development plan in future. The main focus of evaluation is directed at outputs, outcomes, and impacts in implementation of the development plan. Hence, transparent and accountable planning must be accompanied by preparation of performance indicators for the implementation of the plan, which at least include (i) input indicators, (ii) output indicators, and (iii) impact / benefit indicators. In its implementation, evaluation activities can be carried out at different stages, namely:

- Evaluation at the Planning Stage (ex-ante), namely evaluation carried out before the stipulation of the development plan with goals to select and determine the priority scale from various alternatives and possible ways to achieve the goals that have been previously formulated;
- Evaluation at the Implementation Stage (on-going), namely evaluation carried out at the time of implementing the development plan to determine the rate of progress of the implementation of the plan compared with a predetermined plan, and
- Evaluation at the Post-Implementation Stage (ex-post), namely evaluation which is implemented after the execution of the plan ends, that is directed to see if achievement (output / outcome / impact) of the program is able to solve the problem development needing to be solved. This evaluation is used to assess efficiency (outputs and outcomes compared with input), effectiveness (results and impact on objectives), or benefits (impact on needs) of a program.

Under the Local Government Law, there are several regulations which guide the monitoring and evaluation practices in regional development planning. These include the Government Regulation No. 8 of 2008 on Stages of Procedures for Compiling, Control, and Evaluation of the Implementation of the Regional Development Plan. The Article 47 in this Government Regulation states that “Evaluation as referred to in Article 46 includes evaluation of: a. regional development planning policies; b. implementation of regional development plans; and c. the results of regional development plans.” Additionally, the Article 48 states that

(Section 1) Evaluation by the governor, regent / mayor in its implementation is carried out by Regional Development Planning Agency for the overall regional development planning and by the Head of Regional Working Units for the performance achievement of the implementation of previous Regional Working Units programs and activities.

(Section 2) Evaluation by Regional Development Planning Agency includes: a. assessment of the implementation of the process of formulating regional development plan documents and implementing regional development programs and activities; and b. collect, analyse and compile the evaluation results of the Head of Regional Working Units in order to achieve regional development plans.

(Section 3) The results of the evaluation as referred to in section (1) and section (2) become material for the preparation of regional development plans for the next period.

Additionally, the Article 49 in this regulation states that “Governors, regents / mayors are obliged to provide information regarding the results of the evaluation of the implementation of regional development planning to the public.”

Under the GR No. 8/2008, there is the MOHA Regulation No. 86 of 2017 (replacing the MOHA Regulation No. 54 of 2010) on the Implementation of GR No. 8 of 2008 on the Stages, Procedures, Monitoring and Evaluation of Regional Development Plan Implementation. This is considered to be the regulation most important to the local governments. It regulates procedures for planning, control and evaluation of regional development (including ex-ante, on-going, and ex-post evaluation), procedures for evaluating Regional Regulation Draft (Draft PERDA) on Regional Long-Term Development Plans (RPJPD) and Regional Medium-Term Development Plans (RPJMD), and procedures for amendment to Long-Term Development Plans Regions (RPJPD), Regional Medium Term Development Plans (RPJMD), and Local Government Work Plans (RKPD).

At the provincial level, the local government is permitted to establish local regulations to support the existing national regulations. In West Sumatra Province, the latest local regulation about evaluation systems and practices is the Regulation of West Sumatra Governor no 4 of 2016. This Governor Regulation is about monitoring and evaluation instruments for controlling development within the West Sumatra area. The regulation stipulates guidelines for evaluation practitioners, and benchmarks for the task force to do development planning, optimisation for development performance, and good administration implementation.

Local regulations which are prominent in supporting evaluations of regional development planning in DI Yogyakarta Province include Governor Regulation No. 48 of 2009, Governor Regulation No. 73 of 2013, and Governor Regulation No. 45 of 2014. Governor Regulation No. 48 of 2009 is about technical instructions for regional development implementation. The aims are to synchronise the mindset and understanding of all stakeholder and practitioners to plan, implement, and evaluate all regional development programs. Governor Regulation No. 73 of 2013 is the recent update for Governor Regulation No. 48 of 2009. It is about the technical instructions for regional development implementation in DI Yogyakarta Province.

Governor Regulation No. 45 of 2014 is about evaluation development planning results procedures and control regional development implementation. The development planning documents to evaluate include the Regional Long Term Development Plans and Regional Medium Term Development Plans, Regional Working Units Strategic Development Plans, and all other development programs which are funded by the national budget and regional budget.

4.3 Conclusion

Although development planning has been a part of the governmental activity since independence in 1945, Indonesia has seen real progress involving planning and evaluation policies and procedures only after democratisation and decentralisation reforms of the late 1990. Since then, Indonesia has witnessed rapid establishment of a National Development Plan that governs policies at all levels of government. Based on the existing regulatory framework under decentralisation, development planning in Indonesia can be mapped into four stages: 1) Plan formulation; 2) Plan formalisation; 3) Control and implementation; and 4) Evaluation.

Law No. 25 of 2004 on the National Development Planning System (the Planning Law) iterates that national development planning comprised integrated planning across government departments and agencies, as well as across levels of government. Therefore, regional development planning in Indonesia is aligned with the national development planning system. This means that the development carried out in regions is inseparable from the concept of the national development plan. Therefore, regional development plans and programs should continually refer to the national development plan, and both long and medium-term development plans.

There are several existing evaluation practices in the national development planning system in Indonesia. However, these evaluations in Indonesia are quite complicated, inefficient and ineffective, overlapping, at times inaccurate, not timely, and to a certain degree unreliable. It can be concluded that in Indonesia there are too many reporting systems from too many institutions which have poor levels of cooperation and information sharing. The main function of the reporting systems is to monitor physical performance, and financial performance, such as budget realisation or absorption which creates data redundancy and overlap in work by these institutions. However,

those systems are available online which enables compliance monitoring. Therefore, the local governments (RDPA as the one in charge) tend to follow those regulations. In addition to those regulations, there are other regulations requiring local governments to report their own performance. As a result, the local governments have many reporting obligations that they must fulfill. Consequently, most of the attention and resources are used in conducting the monitoring activities and performance measurements, but not in conducting evaluation.

Considering evaluation practices which are expected to support regional development planning is the focus of this study, this study explores the evaluations which are regulated under the Planning Law (Law No. 25 of 2004) and Local Government Law (Law No. 23 of 2014). These are the two laws that are concerned with monitoring and evaluations to support regional development planning. These two laws regulate the responsibilities of regional governments in terms of planning and evaluations. It is stated clearly in the regulations that the RDPA is the authority in charge of evaluations at the local level, and the evaluation results should become material for the preparation of the regional development plan in the following period.

Under the decentralisation era, local governments have the authority and ability to manage and implement regional development programs. Additionally, their role is to determine the success of the implementation process of governance and development activities in the area through monitoring and evaluation. At the provincial level, the local government is permitted to make local regulations to support the existing national regulations. Both West Sumatra Province and DI Yogyakarta Province have their own local regulations in supporting evaluations of regional development planning.

CHAPTER 5 OVERVIEW OF CASE STUDY LOCATIONS

This chapter presents an overview of the case study locations of this research. The first section introduces the West Sumatra Province and the second section gives an overview of the DI Yogyakarta Province, with a particular focus on the geographical context, the socioeconomic conditions, the governance, the RDPA, and the types of evaluation practices found in the case study locations.

5.1 West Sumatra Province

West Sumatra Province is one of the 34 provinces in Indonesia and among the 10 provinces in Sumatra Island. This province has of seven cities (Padang, Solok, Sawahlunto, Padang Panjang, Bukittinggi, Payakumbuh, and Pariaman) and 12 regencies (Kepulauan Mentawai, Pesisir Selatan, Solok, Sijunjung, Tanah Datar, Padang Pariaman, Agam, Lima Puluh Kota, Pasaman, Solok Selatan, Dharmasraya, and Pasaman Barat) (Central Bureau of Statistics of West Sumatra Province, 2019).

5.1.1 Geographical Context

West Sumatra Province is geographically located between 98° 36'-101° 53' East Longitude and 0° 54' North Latitude to 3° 30' South Latitude. Seen from the geostrategic position, West Sumatra is located in the middle of the western part of Sumatra Island. The area of West Sumatra is \pm 42.297,30 km² or 2.27% of the total area of Indonesia (1,860,359.67 km²). West Sumatra has boundaries: the north is bordered by North Sumatra Province; east side is bordered by Riau Province and Jambi Province; the south is bordered by Bengkulu Province, and west side is bordered by the Indian Ocean (Central Bureau of Statistics of West Sumatra Province, 2019) as can be seen in Figure 5.1.



Figure 5.1 Administration Map of West Sumatra Province

Source: ANU College of Asia and the Pacific (2020)

5.1.2 Socioeconomic Conditions

The population of West Sumatra Province at the most recent census in 2018 was 5,382,027 with 72.56% being distributed in the regency (rural) areas and 27.43% in the city areas (Central Bureau of Statistics of West Sumatra Province, 2019). As shown in Table 5.1, Padang City, Agam Regency, Pesisir Selatan Regency, Regency West Pasaman and Padang Pariaman Regency are the five regions with the largest populations in West Sumatra Province, while Kota Padang Panjang, Kota Sawahlunto and Kota Solok are the areas with the least population.

Table 5.1 Population in West Sumatra Province 2015 – 2018

Regency /City		Number of Population			
		2015	2016	2017	2018
Regency		3,781,545	3,824,550	3,866,815	3,908,185
1	Kep. Mentawai	85,295	86,981	88,692	90,373
2	Pesisir Selatan	450,186	453,822	457,285	460,716
3	Solok	363,684	366,213	368,691	371,105
4	Sijunjung	222,512	226,300	230,104	233,810
5	Tanah Datar	344,828	345,706	346,578	347,407
6	Padang Pariaman	406,076	408,612	411,003	413,272
7	Agam	476,881	480,722	484,288	487,914
8	Lima Puluh Kota	368,985	372,568	376,072	379,514
9	Pasaman	269,883	272,804	275,728	278,480
10	Solok Selatan	159,796	162,724	165,603	168,411
11	Dharmasraya	223,112	229,313	235,476	241,571
12	Pasaman Barat	410,307	418,785	427,295	435,612
City		1,414,744	1,434,978	1,454,674	1,473,842
1	Padang	902,413	914,968	927,011	939,112
2	Solok	66,106	67,307	68,602	69,726
3	Sawahlunto	60,186	60,778	61,398	61,898
4	Padang Panjang	50,883	51,712	52,422	52,994
5	Bukit Tinggi	122,621	124,715	126,804	128,783
6	Payakumbuh	127,826	129,807	131,819	133,703
7	Pariaman	84,709	85,691	86,618	87,626
West Sumatra		5,196,289	5,259,528	5,321,489	5,382,027
National		255,461,700	25,870,500	261,890,900	265,015,300

Source : Central Bureau of Statistics of West Sumatra Province (2019)

The average population density in West Sumatra in 2017 was 126 people per km². The most densely populated region was in Bukittinggi Municipality with nearly 5,024 people per km², which is a significant contrast with Mentawai Islands Regency which consisted of only 15 people per km². In 2017, the total population of Indonesia was nearly 262 million with an average population density of 137 people per km², which shows that only 2.03% of the population of Indonesia live in West Sumatra (Central Bureau of Statistics of West Sumatra Province, 2019).

In terms of age structure, West Sumatra is dominated by the young population where the percentage of young people under 15 years is high. This young demographic represents 29.86

percent of the population, while only 5.67 percent of the population are in the elder age group of 65 years and above (Government of West Sumatra Province, 2019).

In August 2018 the main employment status of the nearly 2.5 million workers were labourers or employees amounting to nearly 817,000 people or around 33.20%, followed by self-employed and precarious workers amounting to nearly 516,000 people (20.97%) and 426,000 people (18.90%) (Government of West Sumatra Province, 2019). Workers in the two latter classifications are regarded as having the most insecure employment, and are often underpaid, vulnerable and unprotected people working in risky or dangerous jobs. These workers in Indonesia are frequently exploited and denied basic rights, as they lack legal status or live in financially destitute circumstances (Yasih, 2017).

Agriculture, forestry, and fisheries industries dominate in the regional economy; however, these key sectors declined somewhat during the 2015-2018 period. Average contribution to the GDP of these primary industries is an average of 22.67%. Likewise, the Processing Industry category, which was expected to be one of the drivers of the regional economy, also decreased slightly in the period 2015-2018 from 10.18% down to 8.37% in 2018 with an average contribution to GRDP per year of 8.73%.

Labour absorption until August 2018 was dominated by the working population of low education with junior high school or less education attainment. of the percentage of the population who graduated elementary school was 35.58% and junior high school was 17.52%. Meanwhile, the percentage of the working population with a secondary education, having graduated from high school was 20.59% and Vocational High School graduates was 10.50%. The working population who had a Diploma was 3.74% and the percentage of university educated was 12.07% (Government of West Sumatra Province, 2019).

The Human Development Index (HDI) is an important indicator to measure success in an effort to build the quality of human life, society, and population. The HDI measure explains the standards for the citizens to access development which results in obtaining income, health, education, and other necessities. The HDI is measured by three basic dimensions, namely long lifespan and healthy life , knowledge , and good standard of living (Government of West Sumatra Province, 2019). Table 5.2 compares the HDI data for each regency and city in West Sumatra.

**Table 5.2 Human Development Index (HDI) in West Sumatra Province
2018**

No.	Regency/City	HDI
1	Kep. Mentawai	60.28
2	Pesisir Selatan	69.40
3	Solok	68.60
4	Sijunjung	66.97
5	Tanah Datar	71.25
6	Padang Pariaman	69.71
7	Agam	71.70
8	Lima Puluh Kota	69.17
9	Pasaman	65.60
10	Solok Selatan	68.45
11	Dharmasraya	70.86
12	Pasaman Barat	67.43
13	Padang	82.25
14	Solok	77.89
15	Sawahlunto	71.72
16	Padang Panjang	77.30
17	Bukit Tinggi	80.11
18	Payakumbuh	78.23
19	Pariaman	76.26
West Sumatra		71.73

Source: Government of West Sumatra Province (2019)

The average of HDI in West Sumatra Province in 2018 was 71.73. In comparison to other provinces on the island of Sumatra, the HDI achievement of West Sumatra is in position three after Riau Islands and Riau provinces. In the national rankings, West Sumatra was in 9th place. Among the 34 provinces in Indonesia, DKI Jakarta was the province with the highest HDI in 2018 followed by DI Yogyakarta, Kalimantan Timur, and Kepulauan Riau (Government of West Sumatra Province, 2019).

As can be seen in Table 5.3, only the HDI performance of education showed better performance than the National average. In 2018, the average length of schooling for citizens in West Sumatra only reached 8.92 years, meaning that most of the population did not finished junior high school. This achievement is slightly better than the National Average length of schooling which reached 8.24 years. Among the factors that support the increase in education level attainment in West Sumatra are the reduced dropout rates that have been achieved through assistance with school fees

from the Ministry of Education, and support of other scholarship institutions, as well as increasing availability of school facilities and infrastructure in many regions (Government of West Sumatra Province, 2019).

**Table 5.3 Human Development Index (HDI) in West Sumatra Province
2016-2018**

No.	HDI Components	Unit	2016		2017		2018	
			West Sumatra	National	West Sumatra	National	West Sumatra	National
1	Life expectancy rate	Year	68.73	70.9	68.78	71.06	69.01	71.2
2	Expected years of schooling	Year	13.79	12.72	13.94	12.85	13.95	12.91
3	Mean of years schooling	Year	8.59	7.95	8.72	8.1	8.76	8.17
4	Gross income per capita (adjusted)	Thousand rupiahs/ people/ year	9,804	10,150	10,306	10,664	10,638	11,059
Human Development Index			70.73	70.18	71.24	70.81	71.73	71.39
HDI Growth			1.07	0.91	0.72	0.90	0.69	0.82

Source: Government of West Sumatra Province (2019)

In addition to the increased level of schooling, the expected years of schooling in West Sumatra in 2018 was 13.95 years which is better than the national average of 12.91 years in the same year. This suggests a level of optimism in the people of West Sumatra about future educational conditions. However, the achievement of life expectancy in West Sumatra in 2018 was 69.01 years which is lower than the national average of 71.2 years. This is caused by the low level of public awareness of health and the importance of healthy lifestyles. Likewise, the achievement of adjusted per capita expenditure in West Sumatra in 2018, which reached Rp. 10,638 million per capita / year, is also lower than national achievements of Rp. 11,059,000 per capita / year. To summarise the HDI position of West Sumatra nationally from 2013 to 2018, it was ranked ninth after DKI Jakarta, DI Yogyakarta, East Kalimantan, Islands Riau, Bali, Riau, North Sulawesi, and Banten. However, West Sumatra's HDI during the past few years has risen above the national level, having reached 71.24 in 2017 and 71.73 in 2018 it reaches, whereas the national HDI was only 71.39 (Government of West Sumatra Province, 2019).

The Gross Regional Domestic Product (GRDP) of West Sumatra at current prices in the year 2017 reached Rp214.585,2 billion which placed it 14th in national ranking. Per capita GRDP of West Sumatra at current prices in the same year reached Rp. 40.324,3 thousand. The National Gross Domestic Product per capita of Indonesia at current prices in 2017 was Rp51,887.3000, which shows that per capita income levels in West Sumatra were below the national average. Among the 34 provinces in Indonesia, DKI Jakarta was the province with the highest per capita GRDP in 2017 (Rp232,342.0, thousand) followed by Kalimantan Timur, Kalimantan Utara, and Kepulauan Riau, while West Sumatra Province was well down in the rankings at 20th.

Culture

In West Sumatra Province, 90 percent of the population are Minangkabau ethnic or usually referred as Minang culture. In contemporary West Sumatra, regional governments have been utilising claims that entrenched local customs and practices are based on Islamic religion and the *Holy Qur'an* (traditions often referred to in local language as *Adat Basandi Syarak, Syarak Basandi Kitabullah*, or shortened to ABS-SBK) to defend their backward approaches to governance (Kosasih, 2013; Nelmaya, 2018). This assertion is ubiquitous, appearing as justification for almost every political decision made at the provincial level and below. Moreover, every disaster, from the 1997 Asian Financial Crisis to natural disasters, such as floods, landslides, and earthquakes, has been attributed to the community's failure to meet its religious obligations. For example, after the 2004 tsunami that claimed thousands of lives in western and northern Sumatra, large billboards were displayed throughout the city of Padang proclaiming 'religiosity must be practised in order to prevent a tsunami' (Salim, 2015).

Salim (2015) also found that Minang culture is defined mainly by matrilineality, a social system in which inherited lands, houses, and wealth are transferred from mothers to daughters. Minangkabau men, on the other hand, are well known for their culture of circular out-migration (a custom known locally as *merantau*), a practice which has engendered a culture open to the exchange of ideas and dynamic social relations in West Sumatra.

5.1.3 Governance in West Sumatra Province

As one of the 34 provinces in Indonesia, West Sumatra Province has the typical provincial governance structure as can be seen in Figure 5.2. However, this province has a customary governmental structure for their smallest unit of administrative region which is called *Nagari*. This arrangement has existed since mid-14th century. In the modern era, *Nagari* is comparable to a village, the sub-ordinate of district in formal administrative structure, but it is different because it is not a formal administrative structure and it has more extensive authority than a village. In subsequent developments, the *Nagari* status can be upgraded to become a village, which means that the smallest unit of customary administration structure is the same as the formal administrative structure.

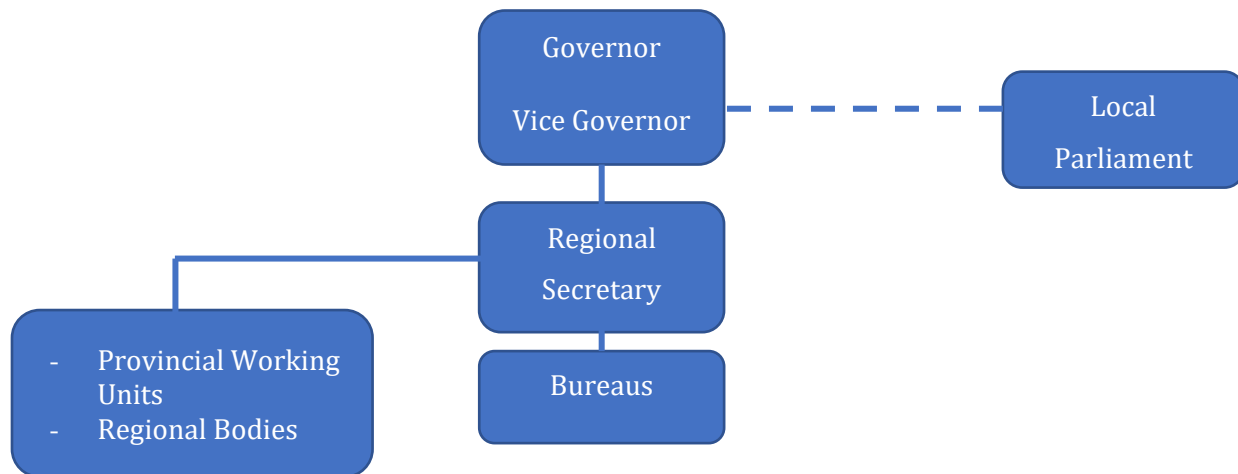


Figure 5.2 Organisational Structure of West Sumatra Provincial Government

Effective, honest, transparent and accountable governance is the basis for implementing bureaucratic reform in the Province of West Sumatra. Since the issuance of the Governor Regulation of the Province of West Sumatra Number 24 of 2011 concerning the guidelines for implementing the bureaucratic reform of the Province of West Sumatra, it has seen changes in the administration of government functions. This can be seen from the achievements of the Government of West Sumatra Province in 2014 related to the assessment of the Ministry of Administrative and Bureaucratic Reform regarding line ministries and province performance accountability which scored 70.52 with BB credit rating of 8th out of 34 provinces assessed. The

results of the ranking of provincial government performance accountability values can be seen in Table 5.4.

**Table 5.4. Provincial Government Accountability Performance Top Ten Ranks
2015**

No.	Provincial Government	Score (0-100)	Rating
1.	DI Yogyakarta	80.68	A
2.	East Java	80.04	A
3.	South Kalimantan	76.30	BB
4.	Bali	75.39	BB
5.	East Kalimantan	75.15	BB
6.	South Sumatra	75.11	BB
7.	Central Java	72.09	BB
8.	West Sumatra	70.52	BB
9.	West Java	70.06	BB
10.	Kepulauan Riau	68.62	B

Source: Ministry of Administrative and Bureaucratic Reform (2015)

This is an achievement for the provincial government in realising a high standard of governance and accountability. However, this achievement must be increased given the on-going problems in implementing bureaucratic reform in the Province of West Sumatra. To complete the overview of performance of the West Sumatra Provincial Government, this study examines other achievements, in particular the Indonesian Government Index (IGI). This index measures provincial governance performance based on a set of objective and measurable criteria. More specifically, it is an assessment of four governance arenas: government, bureaucracy, civil society, and economic society. Although the most recent measurement was done in 2012, nearly a decade ago, the assessment can still give a useful overview of provincial government performance in the case study locations of this research. The IGI for West Sumatra Province can be seen in Table 5.5.

**Table 5.5. Indonesia Government Index for West Sumatra Province (Range: 0-10)
2012**

Arena	Index per Arena	Participation	Fairness	Accountability	Transparency	Efficiency	Effectiveness
Government	5.00	5.93	3.19	6.32	3.74	7.23	5.56
Bureaucracy	5.54	2.63	7.53	6.62	2.34	7.99	5.60
Civil Society	6.40	6.40	6.40	6.40	6.40	6.40	6.40
Economic Society	6.13	6.40	6.40	6.40	6.40	6.40	4.66

Source: Kemitraan (2013)

Based on this index, West Sumatra ranks 20th and is categorised as Fair. The IGI is 5.70, which is the same as the national average of 5.70. Among the four arenas assessed, Civil Society (6.40), Economic Society (6.13), Government (5.00), and Bureaucracy (5.54) arenas are all in the category of Fair. Among the six principles assessed, Efficiency in Bureaucratic Arena (7.99) is categorised as Good, while Accountability in Bureaucracy Arena (6.62) and Effectiveness in Civil Society Arena (6.40) are categorised as Fairly Good. Meanwhile, the principles of Transparency (2.34 and Participation (2.63) in Bureaucracy Arena as well as Fairness in Government Arena (3.19) are in the Poor category. Although the Law No. 14/2008 on Public Information Disclosure demands the openness of public information, it does not necessarily compel the bureaucracy to be transparent. For example, almost all financial documents (RKA, PPKD, SKPDs DPA Summary, Summary PPKD, DPA) as well as the regulations on particular investments in the province are unavailable to the public. In addition, the low budget commitments to education and health sectors contribute to the low index measurement for Fairness in Government Arena. Moreover, the limited outreach of the Local Parliament in development monitoring and control of discrimination has contributed to a low score. Based on this condition, significant efforts still needed to be carried out by the provincial government to enable an equitable, transparent, and accountable form of government to be realised.

5.1.4 The Regional Development Planning Agency (RDPA)

Regional development planning in West Sumatra is governed by Law No. 23 of 2014 on Regional Government, Regulations Region Number 8 of 2016 concerning the Formation and Composition of Regional Apparatus West Sumatra Province, and by Governor Regulation Number 79 2016 concerning the position, organisational structure, duties and functions as well as work procedures. In addition, the Regional Agency and Governor of West Sumatra Province Regulation Number 55 of 2017 concerns job descriptions, principles and functions of the Development Planning Agency of West Sumatra Province. In combination, these laws state that the main task of the West Sumatra Province *Bappeda*, which is the title for the Regional Development Planning Agency, is to assist the Governor in carrying out the government affairs that become regional authority in the field of development planning.

To carry out these main tasks, the West Sumatra Provincial *Bappeda* has the following functions:

1. Formulation of technical policies in the field of regional development planning;
2. Carrying out government affairs and public services in the planning sector regional development;
3. Guidance and facilitation in the field of regional development planning province and district/city;
4. Provision of a regional development planning database;
5. Implementation of *Bappeda* secretariat;
6. Implementation of tasks in the field of human, government and social resources cultural, economic and regional development, macro planning and funding development and control, evaluation and planning information data regional development;
7. Implementation of other duties assigned by the Governor in accordance with the duties and function.

According to Governor Regulation Number 90 of 2017 concerning Amendments to Governor Regulation Number 79 of 2016 concerning Position, Organisational Structure, Duties and Functions and Work Procedures of Regional Bodies and Governor Regulation Number 55 of 2017 concerning Details of Main Duties, Functions and Work Procedures of the Planning Agency Regional Development of West Sumatra Province, the organisational structure of the Planning

Agency Regional Development of West Sumatra Province consists of the Head of the Agency, the Secretariat, and five fields. These five fields are Macro- Planning and Development Funding, Sector Socio-Culture and Government, Field of Control, Evaluation and Data Information Development Planning, Economic Sector and Regional Development, and Functional Group.

The Secretariat is led by a secretary from each field led by a Head of Division and a group of Functional positions, who are located below and responsible directly to the Head of the *Bappeda* of West Sumatra Province. The main tasks and functions of each field are as follows:

1. Secretariat. The Secretariat has the main task of planning, implementing, coordinating, and controlling general administration activities; staffing; equipment; public relations; protocols; drafting programs; and finance.
2. Macro-Planning and Development Funding Sector. This division has the principal task of planning, implementing, coordinating, facilitating, controlling, and reporting activities, which include macro-planning, development funding, and planning innovation systems.
3. Socio-Culture and Government Sector. These sectors have the main task of carrying out coordination of development planning in the fields of Education, Health, Social, Population Control and Family Planning, Population Administration and Civil Registration, Women's Empowerment and Child Protection, Culture, Youth and Sports, Manpower, Transmigration, Archives, Library, Staffing and Human Resources Development, *Trantibum Linmas* (policing and public safety), Supervision, Secretariat of DPRD, Secretariat Regions, Regional Liaison Bodies, Community and Village Empowerment, Research and development, Communication, coding, statistics, functions support, and other functions.
4. Economic Sector and Regional Development. The main task of this group is to carry out coordination of development planning in the field of public works and spatial planning, public housing and residential areas, land, disaster, food, environment, transportation, cooperatives, and small businesses medium, investment, marine and fisheries, tourism, agriculture, forestry affairs, energy and mineral resources affairs, trade affairs, industrial affairs, as well as supporting functions of financial government affairs.
5. Evaluation and Data Information Development Planning Division. This division has the

main task of carrying out control and evaluation development sourced from the APBN and APBD as well as control and evaluation of planning documents.

6. Group of Functional Positions. The Functional Group is in charge of preparing, performing and completing planning activities.

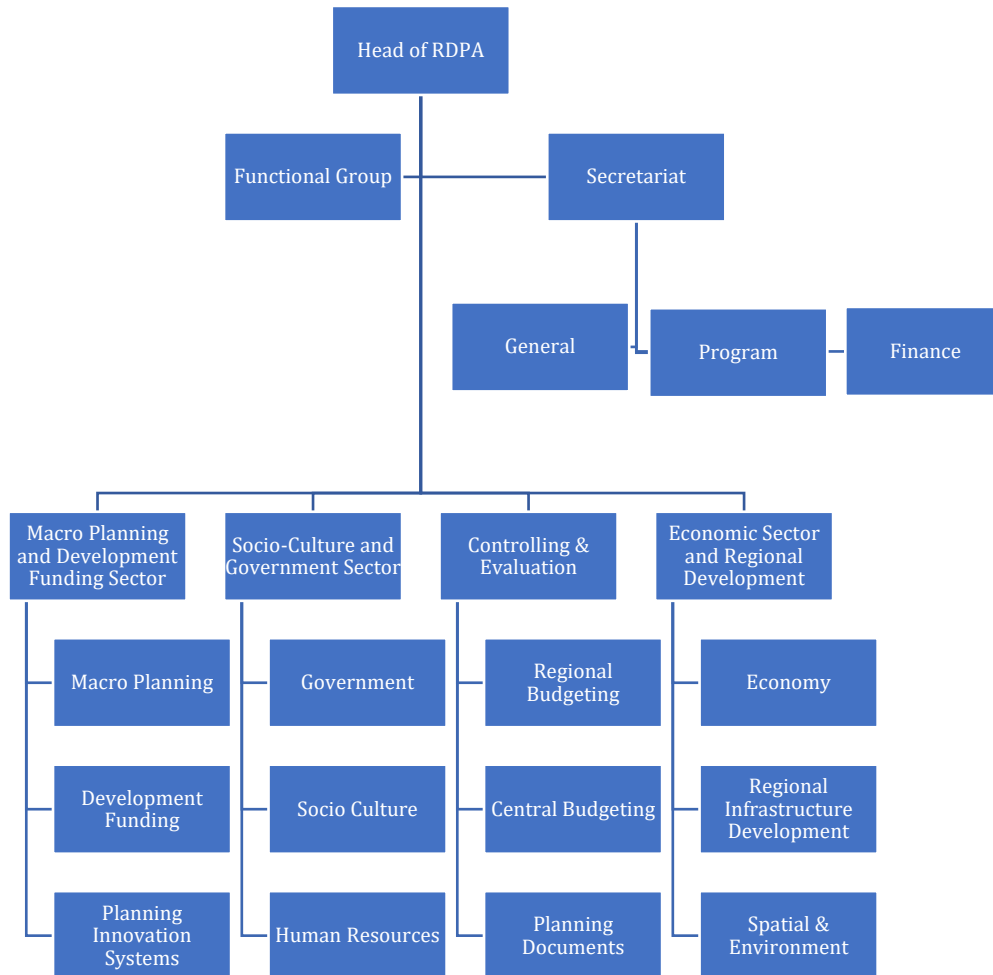


Figure 5.3 Organisational Structure of the RDPA in West Sumatra Province

Source: Bappeda of West Sumatra Province (2019)

Academic qualifications for civil servants within the West Sumatra Provincial Bappeda are consistent with the policy requirements or quantitatively in accordance with the demands of the job. Out of 94 civil servants, 23 people have an educational background of Strata 2 and another

31 people with Strata 1 level qualifications. Table 5.6 summarises the detailed information on human resources (HR) composition in the Bappeda of West Sumatra Province. As can be seen from the table, the largest proportion of the staff (37) are male with high school level education. However, a significant number (54) of the remainder are university educated. Therefore, it can be concluded that the overall level of education is relatively high with greater than half the employees having a tertiary qualification. However, none of them is specialised in evaluations. This fact is explained further in Chapter 6. A further conclusion to be derived from the data is that males dominate the work force at nearly double the number of females.

Table 5.6. Human Resources Composition in the Bappeda of West Sumatra Province based on the Level of Education and Gender 2018

No.	Level of education	Male	Female	Total
1	Doctoral Degree	0	0	0
2	Master Degree	14	9	23
3	Bachelor Degree	15	16	31
4	Diploma	1	1	2
5	High School	30	7	37
6	Junior High School	0	0	0
7	Primary School	1	0	1
	Total	61	33	94

Source: Bappeda of West Sumatra Province (2019)

Based on the Bappeda of West Sumatra Province (2019), it can be noted that the Bappeda has identified the problems that they face in carrying out the main tasks and functions of the organisation, including:

1. The unavailability of an accurate and valid planning data base has caused difficulty in transfer of data affairs to SKPD *Kominfo* as mandated in Article 11 of Law 23 of 2014. This condition has presented obstacles for Bappeda as an institution with main duties and functions in the Planning Sector.
2. Limited planning resources at the Provincial Bappeda with differing levels of expertise according to regional needs, including limited availability and quality of HR as executors of regional development planning.
3. The synergy of development plans among regencies and cities is not yet maximal, between provinces and regencies and cities as well as between provinces and the central

government.

4. Indicators as a measure of development success are necessary to be concrete and become a reference in planning.
5. Increasing the use of control and evaluation results in preparing upcoming development plans.

Point number five in the list above shows that the Bappeda itself has identified the under-utilised issue of evaluation results in the regional development planning process. This point will be discussed further in the following chapters of this thesis.

5.1.5 Types of Evaluation Practices

As described in the previous chapter, evaluation practices are regulated both by the national government and local governments in Indonesia. However, due to the limitation of resources, evaluation practitioners in the RDPA reported during interviews for this research that they could not conduct all evaluations (explained further in Chapter 6). Mainly, they have three groups of evaluation practices: evaluations of planning documents, evaluations of local budget implementation, and evaluations of state budget in the region.

For evaluations of planning documents, this includes evaluations of regional long-term planning documents, regional medium-term planning documents, regency and city planning documents (long-term and annual planning documents), and also Changes of RPJMD (medium-term planning document). In addition, there is the annual plan (*Renja*) and long-term plan (*Renstra*) of OPD (Provincial Working Units). For the Regional Long-Term Planning Document (RPJPD), the evaluation was taking place during the conduct of fieldwork of this research. This was the first evaluation ever conducted for the long-term planning document. For the regional medium-term planning document, there was a mid-term evaluation; however, but it had only been conducted once previously in this province. Moreover, the evaluation practitioner mentioned that there had been no direction from the senior management for conducting this type of evaluation again in the near future. The evaluation practitioner in RDPA reported that it was done a few years ago by the Research and Development Division but this division no longer existed when the fieldwork for this research was conducted. This was because the Evaluation and Controlling Division in RDPA

of West Sumatra Province had just been established in 2017. For the evaluations of regency and city planning documents, the evaluation practitioners reported that they could not evaluate optimally because the regency city governments have not reported regularly to them.

The evaluations for local budget implementation are an on-going evaluation and considered to be the main focus of the provincial government due to the high correlations to the budget. This type of evaluation is usually known as the Performance Evaluation of Programs/Activities as Direct Expenditures of Provincial Government Budget (*Evaluasi Capaian Kinerja Pelaksanaan Program/Kegiatan Belanja Langsung Urusan APBD Provinsi*) or can be called *Evaluasi Triwulanan* or quarterly evaluation. It requires every working unit to evaluate their achievements of financial realisation and absorption and physical achievements each quarter and report to the RDPA for the information to be compiled. This evaluation is resource-intensive and extensive in coverage, as can be seen in the sample of the report from October 2017 shown in Figure 5.4.



Figure 5.4 Quarterly Evaluation Report of West Sumatra Province

The state budget evaluation is mainly done for the purpose of monitoring and evaluation for activities or programs financed by the State Budget or the deconcentration funds in the region. There are several reporting systems for this evaluation where the local government needs to report to several line ministries.

Figure 5.5 shows a collection of evaluation reports that had been done recently by the RDPA. The evaluation practitioners stated that there are many evaluations that they must do, but most are conducted for reporting purposes only and then stored on shelves with little further use.



Figure 5.5 Evaluation Reports in the RDPA of West Sumatra Province

This section has given an overview of the first case study location, West Sumatra Province. This province is geographically located in Sumatra Island, has seven cities and 12 regencies. Based on the socioeconomic conditions in 2017-2018, this province was categorised in the top fifteen performing provinces in Indonesia. West Sumatra Province holds the 9th place in the national rankings based on the HDI achievement, has a higher expected year of schooling than the national average, and holds the 14th in national rankings based on the GDRP. Most of the population in this province are Minangkabau ethnic which is the world's largest matrilineal adherents that adhere to traditional customs and Islamic values in determining important matters and legal issues, including political decisions. Based on the government's performance, West Sumatra Province holds the 8th in the national ranking based on the Provincial Government Accountability Performance. However, based on the Indonesia Government Index, this province ranks 20th out of 34 provinces in Indonesia. Now that the overview of West Sumatra has been described, the following section will describe DI Yogyakarta Province as the second case study location of this research.

5.2 DI Yogyakarta Province

DI Yogyakarta Province, often shortened to DIY, has special status and autonomy as a *Daerah Istimewa* (special region) in the administration of government affairs within the framework of the Republic of Indonesia. This province has one city (Kota Yogyakarta) and four regencies (Gunung Kidul, Kulon Progo, Sleman, and Bantul).

5.2.1 Geographical Context

The province of DIY is geographically located between 7°33'-8°12' South Latitude and 110°00'-110°50' East Longitude. Meanwhile, seen from the geostrategic position, DIY is situated coastally in the middle of the southern part of the island of Java (Figure 5.6). The area of DIY is 3,185.80 km² or 0.17% of the total area of Indonesia (1,860,359.67 km²) and is the second smallest province in Indonesia, but one of the most culturally and economically important due to the income from tourism (Central Bureau of Statistics of DI Yogyakarta Province, 2018). To the north, the province is bordered by Magelang Regency and Boyolali Regency in Central Java Province; the east side is bordered by Klaten Regency and Wonogiri Regency, Central Java Province; the south is bordered by the Indian Ocean, and west side is bordered by Purworejo Regency, Central Java Province.



Figure 5.6 Administration Map of DI Yogyakarta Province

Source: Government of DI Yogyakarta Province (2019)

5.2.2 Socioeconomic Condition

The total population of DIY in 2017 was 3,762,200 people. The population proportion of females has been larger than the population of males, as shown by a sex ratio above 97%, for the last five years (Central Bureau of Statistics of DI Yogyakarta Province, 2018).

**Table 5.7. Population in DI Yogyakarta Province by Gender
2013-2017**

Gender	Year				
	2013	2014	2015	2016	2017
Male	1,775,872	1,797,389	1,818,765	1,839,951	1,860,900
Female	1,818,982	1,839,727	1,860,411	1,880,961	1,901,300
Total	3,594,854	3,637,116	3,679,176	3,720,912	3,762,200
Sex Ratio	97.63	97.7	97.76	97.82	97.88

Source : Central Bureau of Statistics of DI Yogyakarta Province (2018)

Sleman Regency has the largest proportion of Yogyakarta's population with 31.73% of the total population (1,180,479 people). The population in Bantul Regency is 983,527 inhabitants, in Gunung Kidul Regency 722,479 people, in Kulon Progo Regency there are 416,683 people, and

the Kota Yogyakarta with a population of 417,744 people. One factor that affects these figures is the popularity of Sleman and Bantul regencies as centres of study and business which are able to attract people to migrate to these two regencies. DIY Province is considered as a major centre of higher education in Indonesia (Huang, 2009). There are 11 public universities and 106 private higher education institutions in this province (Central Bureau of Statistics of DI Yogyakarta Province, 2018).

Yogyakarta's population composition is dominated by the younger generations or productive age group with 306,180 people between 20-24 years and 314,708 people within the age range of 25-29 years (Central Bureau of Statistics of DI Yogyakarta Province, 2018). Although they could potentially become productive human capital for development in DIY, it is also a challenge to create an employment field that can accommodate this workforce.

During 2013 to 2017, the population of working age adults in DIY was dominated by young people under 35 years. Based on level of education, the number of Vocational High School and SMK (*Sekolah Menengah Kejuruan* or middle school) graduates reached 20.07% with 538,356 people, while university graduates reached 23.65% totalling 284,927 people (Government of DI Yogyakarta Province, 2018). These statistics illustrate that the young, working age population is well-educated. Increased welfare and awareness in the DIY community of the importance of education has helped develop an improved workforce in DIY.

In 2017, the workforce age group nearing retirement at 60 years and over occupies a proportion that is classified as high at 15.77% with a total of 333,849 people. The high number of elderly workers is due to the increasing life expectancy, standard of living, and high education levels that enables people to keep working. However, in rural areas, those that are still working in old age may be forced to do so to fulfill their daily needs with the majority of jobs being in agriculture as unskilled labour (Government of DI Yogyakarta Province, 2018).

Economy

According to the formation of Regional GDP DIY during 2013-2017, the manufacturing sector made the largest contribution to the economy of more than 13%. The agricultural, forestry, and fisheries sectors contributed in the range of 10% to more than 11%, while provision of

accommodation and food and drink sectors ranged from 9% up to more than 10% (Central Bureau of Statistics of DI Yogyakarta Province, 2018).

The industrial sector in DIY is largely supported by small and medium industries (IKM - *Industri Kecil Menengah*). In 2017, the number of IKM in DIY was reported to be as many as 91,214 business units and with an increase of 0.34% when compared to 2016, which was as much as 90,906 business units. These small and medium industries include the Food Industry, Clothing and Leather, Chemicals and Building Materials, Metals and Electronics, and the Craft Industry. When viewed from the number of the business unit, Food Industry and Craft Industry is the IKM with the largest number. Industrial sector in DIY is able to absorb a large workforce with as many as 342,667 workers in 2017 (Central Bureau of Statistics of DI Yogyakarta Province, 2018).

Under the Accommodation and Food and Drink Provision Sector, tourism is one of the subsectors in the DIY economy which deserves attention because of the size of its quite significant contribution. This sector ranked third in the formation of the DIY Regional GDP structure in 2017, which was equal to 10.23% (Central Bureau of Statistics of DI Yogyakarta Province, 2018). This productivity was driven by increases in tourist visits and the number of activities in DIY throughout 2017, including Meeting activities, Incentive, Conference, Exhibition (MICE). Table 5.8 shows the upward trend in tourism numbers and activities during the period 2013 – 2017. The development of tourism in an area has an impact on acceleration of regional economic growth. This often happens because tourism creates demand, both consumption demand as well as investment demand, therefore creating various multiplier effects in the form of goods production activities and services, including consumer goods and capital goods.

**Table 5.8. Number of Tourist in DI Yogyakarta Province
2013-2017**

Year	Tourist		Total Tourist	MICE event number	Length of stay tourist (Days)	
	International	Domestic			International	Domestic
2013	235,888	2,602,074	2,837,962	13,695	1.90	1.59
2014	254,213	3,091,967	3,346,180	12,829	1.95	1.58
2015	308,485	3,813,720	4,122,205	11,337	2.07	1.85
2016	355,313	4,194,261	4,549,574	12,547	2.00	1.95
2017	397,951	4,831,247	5,229,198	13,874	2.13	1.98

Source : Central Bureau of Statistics of DI Yogyakarta Province (2018)

As shown in Table 5.8, the number of tourists in 2017 reached 5,229,298 people or an increase of 14.94% compared to 2016, with 4,831,347 domestic tourists (an increase of 637,086 people or 15.18%) compared to 2016 which was 4,194,261. The number of foreign tourists in 2017 amounted to 397,951 people or an increase of 12.01% compared to 2016 amounting to 355,313 people (Central Bureau of Statistics of DI Yogyakarta Province, 2018).

From 2013 to 2017 the Human Development Index of D.I. Yogyakarta Province was second highest of the 34 provinces in Indonesia, below the DKI Jakarta Province and above East Kalimantan Province. In September 2017, DI Yogyakarta was ranked as the 12th province with the largest percentage of poor people, which amounted to 12.36 percent (Central Bureau of Statistics of DI Yogyakarta Province, 2018).

In 2017, the growth rate of Regional GDP based on the constant market price of DI Yogyakarta was in the 20th position of 34 provinces in Indonesia. When compared with other provinces in Java, DI Yogyakarta had the lowest GDP growth rate of 5.26 percent. In contrast, DKI Jakarta was the highest with 6.22 percent (Central Bureau of Statistics of DI Yogyakarta Province, 2018).

Culture

In DI Yogyakarta Province, 96% of the population are ethnic Javanese and the region is often regarded as the heartland of Javanese culture. The Javanese customs and cultural values have foundations in pre-Islamic Buddhist and Hindu dharmic civilisations; the temples and monuments of their ancestors are still prominent features in many parts of the province. Even today, Javanese culture guides daily life and society. Some of the significant Javanese cultural values include *rukun* (get along well) and *nrimo* (accepting and grateful) (Pambudi, 2016). *Rukun* value is a way of action that is based on the continuous efforts of all individuals to interact peacefully with one another and eliminate elements that are potentially destructive (Pambudi, 2016). In addition, the value of *rukun* is also defined as an attitude of mutual respect and adjustment in order to create a harmonious and serene relationship (Karina, 2014). Whereas the value of *nrimo* refers to an attitude of accepting everything that happens with calmness, without rejection, along with psychological or spiritual awareness and assuming that everything is a provision or gift from God that can be acquired through hard work and effort. In Javanese culture, *nrimo* is a way of thinking

and acting based on mutual respect, courtesy, patience, indirectness, and emotional restraint (Djakababa, 2010).

Those cultural values affect the socio-political and socio-economic developments in Yogyakarta. This is especially true with the special status of this province and its relative autonomy, and where the traditional cultural leader of the region, the Sultan of Yogyakarta, has been appointed to be the governor of the province for life (this point is discussed further in the next section). Even though the special status law was only passed in 2012, for the people of Yogyakarta, the Sultan and the royal family have ruled the province in one form or another for centuries. More importantly, perhaps, even before democratisation of Indonesia began in 1998, the Sultan had always been the governor, so Yogyakartaans simply had no experience with a non-monarchical, political style of leader. While a small minority of activists sought to install a democratically elected political leader, the majority of citizens preferred to be ruled by a hereditary monarch, as their ancestors have been for thousands of years. Rather than mass protests against the continuation of a seemingly anachronistic monarchy, there were protests in Yogyakarta against plans to abolish the Sultan of Yogyakarta's dominant role as both hereditary ruler and political leader of the province. As Harsono (2018) observed, many people have apparently internalised the traditional and cultural Javanese values represented by the sultanate and royal ruling family. Consequently, there is the view among people in the province that these values should be embodied in any modern leader who should not simply fulfil ceremonial functions and be the spiritual guide for the people, but should also rule the province as its governor. Thus, the notion that 'the sultan is the governor and the governor is the sultan' does not appear to be unusual at all for many Yogyakartaans (Harsono, 2018).

5.2.3 Special Status of DI Yogyakarta Province

DI Yogyakarta Province gained the status as a special region immediately after the end of colonial period when its contribution to the struggle for independence was recognised in the constitution and in the Law Number 3 Year 1950. In essence, special status meant that the traditional cultural leader of the region, the Sultan of Yogyakarta, Sultan Hamengkubuwono X, became the governor. In the national political context, the existence of local level monarchies is not problematic because sub-national administrative entities have little autonomy compared to the central government.

Although the beginning of democratisation in 1998 led to legal challenges from the central government, all attempts to reduce the sultan to be a purely ceremonial leader failed and the status quo was eventually reconfirmed in a new law passed in 2012, the Law Number 13 Year 2012 on the special status of Yogyakarta.

Therefore, currently the main difference between DI Yogyakarta Province and other provinces in Indonesia is that the governor of DI Yogyakarta Province is not directly elected by the people but is a hereditary leader with political power stemming from both the traditional role as ruler and from direct appointment by the central government. This kind of appointment is not open to anyone who might be favoured by the central government but is limited to the Sultan of Yogyakarta and his hereditary line.

The success of the Sultan's administration in Yogyakarta has achieved a considerable degree of performance legitimacy in the eyes of the people who live in the city and its rural outskirts. Urban people are generally considered to be more rational in their political choices compared to rural people. In the case of Yogyakarta, the Sultan has always had a strong following among urban communities because of the combination of political stability and economic development attributed to the sultan's leadership (Efendi, 2012). Currently, Yogyakarta is one of the most prosperous and advanced regions in Indonesia. Accordingly, there is little impetus for Yogyakartaans to demand changes to the existing system of government. Rather than demanding elections, many people prefer to keep the Sultan's administration as it is as it has proven able to improve the living standards and socio-economic conditions of Yogyakarta. Imposing electoral competition on the city, by contrast, may jeopardise the institutional arrangement that made Yogyakarta what it is today (Harsono, 2018).

The Sultan has taken advantage of both cultural and performance legitimacies in the region since cultural legitimacy comes from the Sultan's charismatic position as the keeper of Yogyakarta's cultural traditions. This is reflected in the paternalistic behavioural patterns within the provincial bureaucracy where civil servants recognise the Sultan's position as the cultural symbol of society.

Arguably, the only groups that are highly critical of the Sultan are NGOs or community organisations who advocate the end of people's suffering due to the government's pro-business policies. In recent years, these community organisations or NGOs have increasingly utilised social media as a platform to express their criticisms and concerns, but overall their influence on

government policy is extremely limited. The mainstream media also are ineffective in shaping public opinion or in acting as guardians of civil society, since most TV stations and newspapers tend to report favourably about the Sultan for fear of intervention from the palace if they are too critical. National networks' media outlets tend to be more critical, but they too often censor contents for business reasons (Harsono, 2018).

The Sultan utilises interlocking roles to connect with the needs of various interests among stakeholders in Yogyakarta. The Sultan's multiple roles enable avoidance of some of the negative issues affecting other local politicians, even though there are indications of paternalistic and clientelistic relations between the Sultan and several stakeholders to cement the governor's power and influence. Overall, the Sultan enjoys a high degree of legitimacy, despite the absence of a public electoral mandate (Harsono, 2018).

5.2.4 Governance in DI Yogyakarta Province

As a result of the special status of Yogyakarta, the political system of this province combines informal features of a pre-colonial sultanate with a modern administration based on formal institutions, such as laws written after Indonesia achieved independence. Figure 5.7 illustrates that the Sultan sits at the top of both the provincial government and the Sultanate and can, therefore, draw on a diversity of power sources.

The provincial government consists of the various working units of the provincial administration and the provincial parliament. All of these working units are meant to work together to develop the province. However, cultural connection may create conflicting interests among those individuals working for the provincial government in performing their duties and responsibilities (Harsono, 2018).

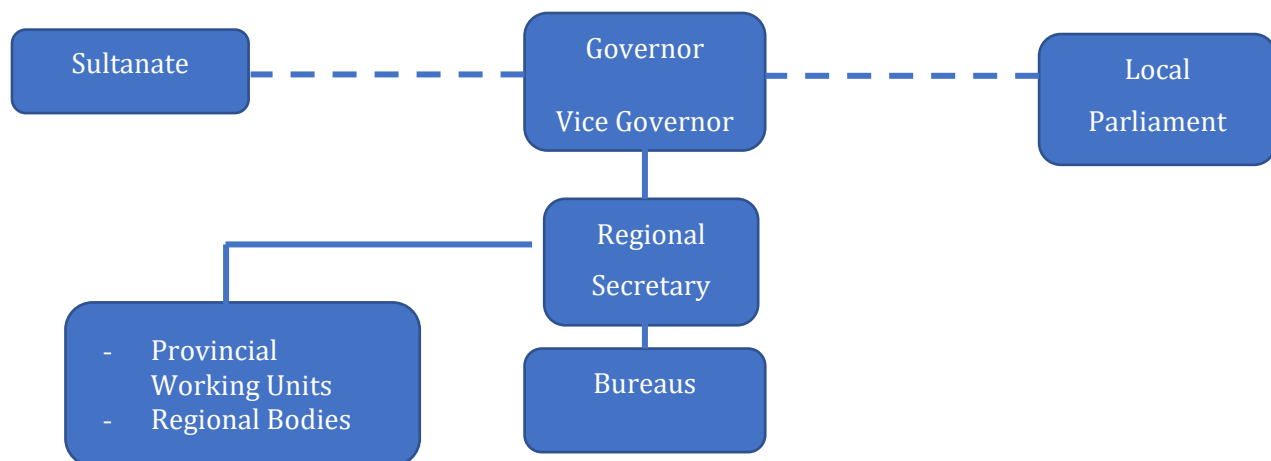


Figure 5.7 Organisational Structure of DI Yogyakarta Provincial Government

As can be seen in Table 5.4, DI Yogyakarta Provincial Government has a good performance in accountability which scored 80.68 with A rating or 1st out of 34 provinces assessed by the Ministry of Administrative and Bureaucratic Reform. For the Indonesia Government Index, DI Yogyakarta Province, as can be seen in Table 5.9, also has a higher index compared to West Sumatra Province.

**Table 5.9. Indonesia Government Index for DI Yogyakarta Province (Range: 0-10)
2012**

Arena	Index per Arena	Participation	Fairness	Accountability	Transparency	Efficiency	Effectiveness
Government	6.52	6.40	2.94	8.37	7.97	6.70	5.88
Bureaucracy	7.46	9.55	7.38	7.73	9.09	5.42	5.87
Civil Society	6.72	7.64	6.40	6.40	6.40	6.40	7.03
Economic Society	6.12	6.40	6.40	6.40	6.40	6.40	4.61

Source: Kemitraan (2013)

Based on this index, DI Yogyakarta ranks 1st and categorised as Fairly Good. Their index is 6.80, above the national average of 5.70. Among the four arenas assessed, Bureaucracy Arena (7.46) is categorised as Good, Civil Society Arena (6.72) and Government Arena (6.52) are in Fairly Good category, while Economic Society (6.12) is Fair. Among the six principles assessed, the significant principles contributing DIY on the top rank are Participation (9.55) and Transparency (9.09) in

the Bureaucracy Arena which categorised as Very Good, and Accountability of the Government Arena (8.37) which is in Good category. However, the principal Fairness of the Government Arena (2.94) is categorised as Poor, Effectiveness of the Economic Society Arena (4.61) is Fairly Poor, and Efficiency of the Bureaucracy Arena (5.42) is in Fair category. The Poor category of the Fairness of the Government Arena is due to the lack of commitment in budget allocation (APBD) for health. The budget allocation for health per capita per year is only IDR 5,807. This number is calculated from the total provincial health budget minus operational cost and divided by the total population. This allocation is the lowest compared to other provinces in Indonesia. The IGI indicator shows how budget allocation to bureaucracy expense was unfair to the public. Despite using their authority in determining budget allocation (program and overhead costs) for bureaucracy, IGI results show the ratio of personnel expenditure budget (direct and indirect) to total budget reached 96%. Therefore, Table 5.9 shows how DIY bureaucracy is inefficient. The result of Fairly Poor of Effectiveness in Economic Society Arena was, among other factors, contributed to by the low employment rate.

5.2.5 The Regional Development Planning Agency (RDPA)

The current Bappeda of DI Yogyakarta Province (Bappeda DIY) was formed based on a Special Region Regulation DI Yogyakarta Number 1 of 2015 concerning Government Institutions DI Yogyakarta. Bappeda DIY is an implementing element of government administration regions with the organisational structure shown in Figure 5.8.

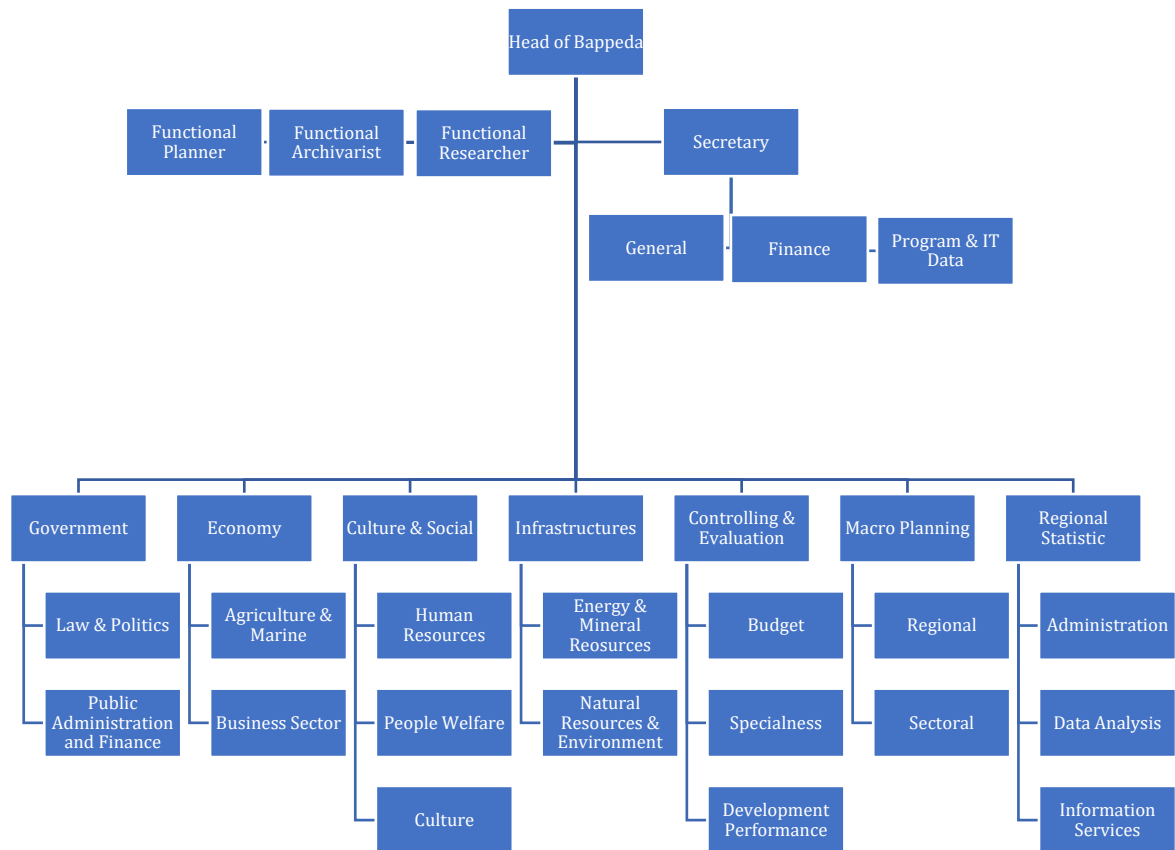


Figure 5.8 Organisational Structure of Bappeda DI Yogyakarta Province 2018

Source: Bappeda of DI Yogyakarta Province (2019)

As can be seen in Figure 5.8, Bappeda DIY has six main divisions and one statistics centre. The six divisions are: Government Planning Division, Economic Planning Division, Culture and Social Planning Division, infrastructures Planning Division, Controlling and Evaluation Division, and Macro-Planning Division. To carry out the duties and functions of Bappeda DIY, competent and qualified human resources are needed. In Bappeda DIY, as of December 31 2018, there were 105 employees consisting of 27 structural officials, 24 certain functional officials, and 54 general functional officials. These personnel were divided in the Main Agency (Bappeda DIY) as 96 employees consisting of 24 structural officers, 24 specific functional officers (20 functional planners, 3 functional researchers and 1 archivist), and 48 general functional people. Meanwhile, nine people support the implementation of duties and functions at the Regional Statistics Center (BSD DIY), consisting of three Structural Officials and six General Functional people (Bappeda of DI Yogyakarta Province, 2019).

Based on the DIY Governor Regulation Number 85 of 2017 concerning the Second Amendment to the Regulation of the Governor of the Special Region of Yogyakarta Number 114 of 2015 concerning Types and Needs of Functional Positions within the DIY Regional Government and Governor Regulation Number 122 of 2015 concerning Qualifications of Executive Position, to carry out the duties and functions of Bappeda , ideal human resources are needed in the main agency (Bappeda). These personnel comprise 25 structural officials, 31 certain functional officials (25 functional planners, 4 functional researchers, 1 archivist), and 55 general functional officials, while in BSD four structural and 33 functional officials general are required. Thus, up to the end of 2018, data from Bappeda DIY shows there is still a shortage of human resources. With the existing conditions of human resources, Bappeda must carry out their duties and functions optimally with the support of experts and non-civil servants to balance the workload that is the task of Bappeda. Table 5.10 summarises the human resources profiles in Bappeda DIY where there is a shortage of the ideal number of human resources (55 people) compared to the existing condition of 48 people (Bappeda of DI Yogyakarta Province, 2019).

Table 5.10. Human Resources Profiles in Bappeda DI Yogyakarta Province 2018

Position in Bappeda	The condition of human resources at Bappeda DIY as of 31 December 2018	Ideal human resources based at Bappeda DIY on DIY Governor Regulation number 85 of 2017
Structural Officers	24	25
Specific Functional Officers		
Functional Planners	20	25
Functional Researchers	3	4
Archivist	1	1
General Functional Officers	48	55

Source: Bappeda of DI Yogyakarta Province (2019)

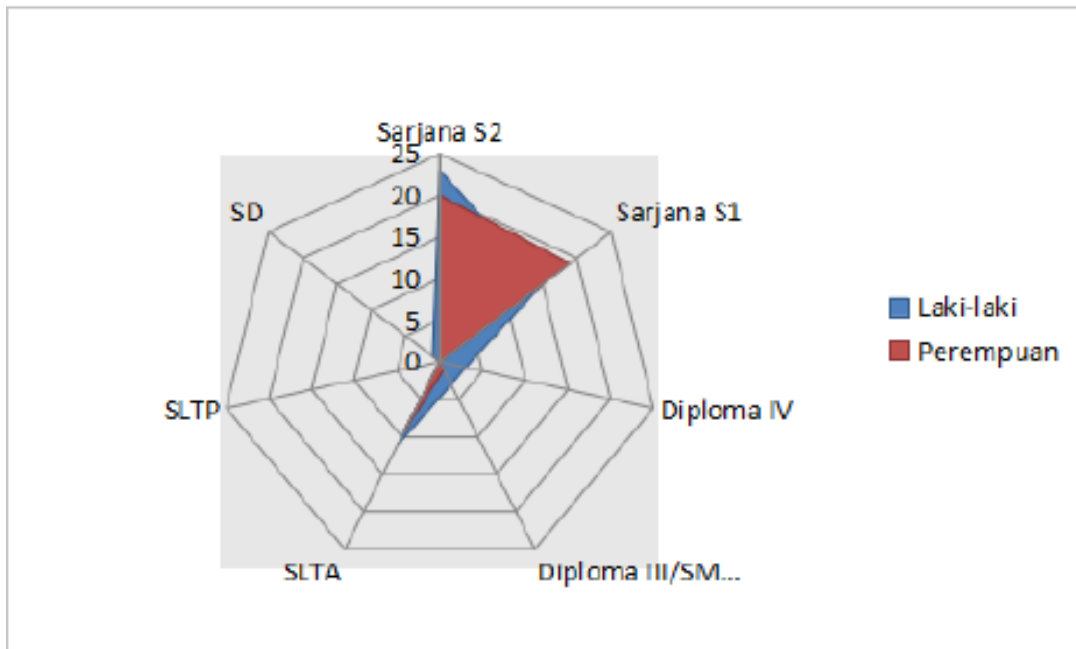


Figure 5.9 Human Resources Profiles in Bappeda DI Yogyakarta Province based on Education Level and Gender (Laki-laki=male, Perempuan=female) 2018

Source: Bappeda of DI Yogyakarta Province (2019)

As can be seen in Figure 5.9, when viewed from the level of education, the qualification status of employees at Bappeda DIY is dominated by undergraduate (S1) and postgraduate degrees (S2) related to functional planners and research positions (Bappeda of DI Yogyakarta Province, 2019). Therefore, it can be concluded that on average the level of education of the human resources in Bappeda DIY is highly educated.

Bappeda DIY implements Governor Regulation Number 51 of 2015 concerning Details of the Duties and Functions of the RDPA of the DI Yogyakarta Province; however, there are many factors that influence the success of Bappeda DIY in carrying out its duties and functions. In general, planning and implementation of regional development programs and activities is strongly influenced by two main factors. The first factor is the parties involved in the planning and implementation of programs and activities, and the second is how the planning and implementation processes of programs and activities proceeds (Bappeda of DI Yogyakarta Province, 2019).

Bappeda of DI Yogyakarta Province (2019) pointed out that the most influential parties involved in the planning and implementation of regional development programs and activities currently include the government and the community (including the Local Parliament). Meanwhile, the dynamic that most influences the planning and implementation of regional development programs and activities is the political process (rather than the technocratic process). This point will be discussed further in following chapters.

5.2.6 Types of Evaluation Practices

Similar to the West Sumatra Province, the practitioners in DI Yogyakarta reported that they mainly have three types of evaluation practices: evaluations of planning documents, evaluations of local budget implementation, and evaluations of state budget in the region. For evaluations of planning documents, this includes evaluations of regional long-term planning documents, regional medium-term planning document, regency and city planning documents (long-term and annual planning documents), and also Changes of RPJMD (medium-term planning document). Then also annual plan (*Renja*) and long-term plan (*Renstra*) of OPD (Provincial Working Units). The evaluation practitioners in RDPA reported that they conduct all the evaluations regularly. Additionally, the practitioners in RDPA stated that they also conducted the Evaluation of Regional Development Performance for the regencies and cities in this province.

The evaluations for local budget implementation are an on-going evaluation and are considered as the focus of the provincial government due to the high correlations to the budget. This type of evaluation is usually known as the Performance Evaluation of Programs/Activities as Direct Expenditures of Provincial Government Budget (*Evaluasi Capaian Kinerja Pelaksanaan Program/Kegiatan Belanja Langsung Urusan APBD Provinsi*) or can be called *Evaluasi Triwulanan* or quarterly evaluation. It requires every working unit to evaluate their achievements of financial realisation and absorption and physical achievements each quarter and report to the RDPA to be compiled.

Finally, the state budget evaluation is mainly monitoring and evaluation for activities or programs financed by the State Budget or the deconcentration funds in the region. There are several reporting systems for this where the local government needs to report to several line ministries.

5.3 Conclusion

This chapter has provided an overview of the case study locations of this study: West Sumatra Province and DI Yogyakarta Province. These are two provinces of the 34 provinces in Indonesia. West Sumatra is one of the 10 provinces in Sumatra Island, while DI Yogyakarta is one of the six provinces in Java. West Sumatra has the typical provincial government governance, while DI Yogyakarta has a special status in the administration of government affairs.

Today, Yogyakarta is the only province in Indonesia where a traditional ruler continues to exercise not only informal traditional influence but also formal political power. Headed by the Sultan, this provincial government consists of the various working units of the provincial administration and the provincial parliament, all of which are meant to work together to develop the region. However, the hybrid form of provincial government leadership has created political and cultural conflicts of interest and presented difficulties of allegiance among government personnel in performance of duties responsibilities.

Compared to West Sumatra Province, DI Yogyakarta Province has better ranking in planning performance (based on APN award as explained in Chapter 3), accountability performance (based on the accountability ranking by the Ministry of Administrative and Bureaucratic Reform), and Government Index (based on Indonesia Government Index by Kemitraan). In comparison with West Sumatra, DI Yogyakarta Province has better quality of human resources in the RDPA, as the leading agent for evaluations in planning at the provincial level, when viewed from the level of education although both provinces have shortages in the number of employees.

Regarding the types of evaluation practices, both provinces have the same types of practices. They mainly have three groups of evaluation practices: evaluations of planning documents, evaluations of local budget implementation, and evaluations of state budget in the region. The current state and challenges of the evaluation practices in these two provinces will be elaborated in detail in Chapter 6 and the expectations of the evaluation practices from the stakeholders will be discussed in Chapter 7.

CHAPTER 6 RESEARCH FINDINGS

The aim of this chapter is to present the main findings from the fieldwork in order to answer the research questions of this study. This chapter will be divided into two main sections. Section 6.1 provides a working definition of an effective evaluation practice in the context of regional development planning. It aims to answer the first sub-question of this research: “What is an effective evaluation practice in the context of regional development planning?” It draws on the literature review in Chapter 2 which describes the concept of effectiveness, evaluation standards, and effective evaluation across disciplines with the adaption of effectiveness framework, and also the results of interviews and focus group discussions to develop the proposed definition of effective evaluation practice in regional development planning. Section 6.2 explains what generates effectiveness in evaluation practices, for whom, under what circumstances and why, by focusing on the elements of context, the mechanisms, interrelation between mechanisms and elements of context, and the outcomes. This section aims to answer the second and third sub questions: “How do evaluation practices support the regional development planning processes?” and “What are the circumstances that aid evaluation practices to be effective especially in supporting regional development planning process? What are the factors that create an enabling environment for effective evaluation?”

The chapter is based on the fieldwork data which includes interviews, focus groups, and documents, such as official reports and government regulations. However, it is important to note that most of the data comes from individual interviews. This study involves participants from the national level and also participants from case study locations, as explained in detail in Chapter 3. Given that the Regional Development Planning Agency is the key stakeholder of evaluation practices at this level, the participants from this institution are the main sources of data in this research. The evaluators in RDPA are the practitioners who conduct evaluations while the policy makers in RDPA are the intended users and also the data providers for evaluation practices. Therefore, interviews with the evaluators and policy makers in RDPAs are the main sources of information in this study.

6.1 Effective evaluation practice in regional development planning

In order to examine ‘what works and for whom’, it is necessary to clarify what is meant by “working” or what constitutes “effective” evaluation practice. As explained in Chapter 3, this study adapts the effectiveness framework from environmental assessment context into evaluation practice in regional development planning. This framework involves four dimensions of effectiveness and the relevant criteria for evaluation practice in regional development planning context. As explained in Chapter 3, this study mainly focuses on three dimensions of effectiveness, which are procedural, substantive, and transactive effectiveness considering more extensive data and evidence are needed to analyse the normative effectiveness. The following section presents the results of primary data collection, showing how the stakeholders in regional development planning defined effective evaluation practice based on the effectiveness framework that has been developed in this study.

6.1.1 Effectiveness Framework of Evaluation Practices in Case Study Locations

The stakeholders’ perspectives or perceptions of effectiveness of evaluation practices in regional development planning in each case study location are summarised in Figure 6.1 and Figure 6.2. Not every interview participant shared their perspectives on every criterion in this effectiveness framework; some were more concerned about specific criteria and somehow chose to disregard the others. For example, no participants from the NGOs category expressed their perspective about the “Integration of evaluation into the planning process” criterion. Moreover, it is important to note that not all participants in the same group of stakeholders have the same perspective about the effectiveness of evaluation practices based on the criteria in this effectiveness framework; some have slightly different views. Therefore, these figures give an overview of the majority perspectives of stakeholders interviewed. These are unpacked further in the following section.

In Figure 6.1 and Figure 6.2, the rows show different types of effective evaluation criteria, divided into three categories based on the effectiveness framework developed and explained in Chapter 3, which are procedural effectiveness, substantive effectiveness, and transactive effectiveness. The columns show the ratings from the four groups of stakeholders, which were Evaluators, Policymakers, Academics, and NGO officials.

Overall, in West Sumatra Province, none of the stakeholder groups evaluated any of the evaluation criteria to have been successfully met. This is shown by the absence of coloured green cells in Figure 6.1. The evaluators judged four of the criteria to have been partially met, as shown by cells coloured yellow. The policymakers and the NGOs judged just one of the criteria to have been partially met, while the academics thought that none of the evaluation effectiveness criteria had been met, as indicated by all cells coloured red. However, it can be noted also that there are clear differences between the state and the non-state stakeholders' views. The NGO group had different views to those of the other stakeholders. Their different perspective might be because they were not actively involved in the evaluation process. Therefore, they did not have enough information on which to assess the evaluation practices with all of the criteria in the effectiveness framework. They could only share their views on certain criteria. However, as outsiders, their views were more independent considering they were not part of the process (Hadiwinata, 2003). The evaluators and policy makers, the state stakeholders, seemed to view the state of evaluation practices more favourably. In particular, most of the evaluators reported that they were trying their best to conduct the evaluation practices within available resources. However, they also reported that resources were very limited. Consequently, the outcomes were not as good as they had hoped. These findings will be unpacked further in the next section.

Effective evaluation criteria	Groups of stakeholders			
	Evaluators	Policy makers	Academics	NGO Officials
Procedural Effectiveness Criteria				
a. Relevant procedures,	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion	Partly succeeded to meet the criterion
b. Integration of evaluation into the planning process,	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion	No evidence on this criterion
c. Involvement of stakeholders in the process,	Partly succeeded to meet the criterion	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion
d. Delivery of results to relevant stakeholders,	Partly succeeded to meet the criterion	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion
e. Objectivity, clarity and understandability of evaluation results.	Partly succeeded to meet the criterion	Failed to meet the criterion	Failed to meet the criterion	No evidence on this criterion
Substantive Effectiveness Criteria				
a. Support accountability,	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Failed to meet the criterion	Failed to meet the criterion
b. Support informed decision-making.	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion
Transactive Effectiveness Criteria				
a. Resource management,	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion
b. Specification of roles.	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion

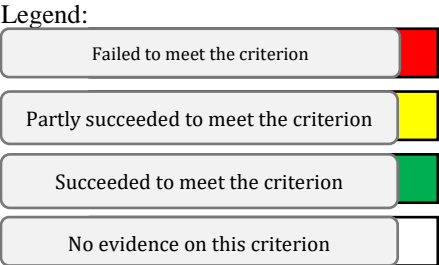


Figure 6.1 Effectiveness of Evaluation Practices in Regional Development Planning in West Sumatra Province

Figure 6.2 shows stakeholders’ perspectives of effectiveness of evaluation practices in DI Yogyakarta Province. Almost all groups of stakeholders in this province perceived that the evaluation practices in this province at least partly succeeded in meeting most of the criteria in the effectiveness framework. It can be noted also that the academics group’s views were very positive in this province. This is arguably because in DI Yogyakarta Province, most of the interviewed academics were actively involved in the regional development planning process, including the evaluation process. Therefore, they tended to view the evaluation practices favourably since they were part of the team supporting the practices and can be considered as insiders. Additionally, since academics were actively involved in the evaluation process, they had richer information

about the effectiveness of the evaluation practices in this province. This can be seen in more detail in the next section.

Figure 6.2 shows that, as in West Sumatra Province, the NGO group had the most diverse views compared to the other groups of stakeholders, but for different reasons. In DI Yogyakarta Province, NGO officials were actively involved in the regional development planning process as well, including in evaluations, although not being part of the main actors such as the academics. The main reason why their assessment of the effectiveness of evaluation practices is relatively worse than the views of other stakeholder groups is arguably because of the nature of NGOs. NGOs in Indonesia tend to be very independent and critical but in a constructive way (Hadiwinata, 2003). Additionally, DI Yogyakarta as one of the centres of education in Indonesia, has many highly credible NGOs (MacRae and Hodgkin, 2016). Therefore, their officials are relatively well-versed in conveying their critical views.

Figure 6.2 also shows that the policy makers seemed to perceive the “relevant procedures” and “resource management” criteria more favourably than the evaluators group. This is mainly because the evaluators had more detailed information and, therefore, they had a more critical view. For the “relevant procedures” criterion, most of the evaluators reported that there were conflicting regulations on the evaluation practices. Therefore, they stated that for this reason, the evaluation practices can hardly be effective. The same reason applied to the “resource management” criterion. All of the evaluation practitioners complained that the available resources were far from sufficient for supporting them in conducting evaluations. Consequently, it was very challenging for them to produce effective evaluation practices. All of those points will be elaborated in the next section.

Effective evaluation criteria	Groups of stakeholders			
	Evaluators	Policy makers	Academics	NGO Officials
Procedural Effectiveness Criteria				
a. Relevant procedures	Failed to meet the criterion	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Failed to meet the criterion
b. Integration of evaluation into the planning process,	Failed to meet the criterion	Failed to meet the criterion	Partly succeeded to meet the criterion	Failed to meet the criterion
c. Involvement of stakeholders in the process,	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion
d. Delivery of results to relevant stakeholders,	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Failed to meet the criterion
e. Objectivity, clarity and understandability of evaluation results.	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Failed to meet the criterion
Substantive Effectiveness Criteria				
a. Support accountability,	Succeeded to meet the criterion	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion
b. Support informed decision-making.	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Failed to meet the criterion
Transactive Effectiveness Criteria				
a. Resource management,	Failed to meet the criterion	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion	Partly succeeded to meet the criterion
b. Specification of roles.	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion	Failed to meet the criterion

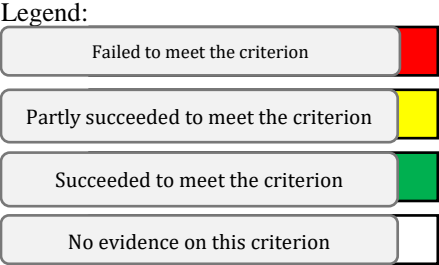


Figure 6.2 Effectiveness of Evaluation Practices in Regional Development Planning in DI Yogyakarta Province

The summary figures show that evaluation practices in DI Yogyakarta Province were perceived to be better than the evaluation practices in West Sumatra Province. However, it is important to note the different cultural backgrounds of the people in these two provinces and the effect that can have on practices. People in Sumatra, including the ethnic groups of West Sumatra Province, generally have a culturally-based “speak your mind” attitude, and tend to share their opinions and critiques. Javanese people, on the other hand, including those in DI Yogyakarta, have a “*nrimo*” cultural attitude where they tend to be accepting and grateful for what they have, while tending to minimise expression of complaints (Djakababa, 2010). This is supported by the statement from participants, especially one of the academics as follows:

After all, the people of Jogja are famous for their Javanese culture, which is quite obvious, some of them are obedient to their superiors or their elders; they tend to be *nrimo*. (Academic participant)

Therefore, assessment of evaluation practice in Sumatra is also influenced by the people's cultural background. However, in relation to mechanisms which actually contribute to the effectiveness of evaluation practices, matters are much more complex.

As can be seen from Figure 6.1 and Figure 6.2, none of the criteria in the effectiveness framework were fully met at either of the case study locations. However, the evaluation practices in DI Yogyakarta Province were perceived to be better in meeting the criteria. Additionally, the only criterion generally perceived to be almost fully met or partially met both in West Sumatra Province and DI Yogyakarta Province was the "support accountability" criterion. Most stakeholder groups, with the exception of academics and NGO groups in West Sumatra Province, agreed that evaluation practices in both provinces were able to support accountability in regional development planning. Moreover, the evaluation practitioners in DI Yogyakarta stated that they were quite confident to say that the evaluation practices fully support accountability. To back up this view, they point to an online reporting system called *Sengguh*, a verification process which they called "feedback desk", and also a routine quarterly meeting which is led by the DIY Governor where all the Heads of the Working Unit must be present.

Based on the effectiveness framework discussed above, it can be synthesised that an effective evaluation practice in regional development planning is one which fully meets the criteria under the procedural, substantive, transactive, and normative effectiveness dimensions. Therefore, this study defines an effective evaluation practice as one which is procedurally, substantively, transactively, and normatively effective. In other words, evaluation practice can be categorised as effective when it is implementing the principles governing evaluation process, achieving the set objectives, delivering the outcomes with minimum resources, and achieving the normative goals of evaluation practice.

The Effectiveness Framework developed in this study is useful in defining what is effective evaluation practice and also determining the state of evaluation practices in case study locations while the Context-Mechanism-Outcome Configuration, which will be elaborated in the following sections, is useful for unpacking the elements of context and mechanisms that influence the

effectiveness of evaluation practice. Additionally, the discussion in this section, complemented by Figure 6.1 and Figure 6.2, has shown that the effectiveness framework can be used as an assessment tool to determine current level of evaluation practice, and to monitor improvements.

6.2 What generates evaluation practice effectiveness, under which circumstances and why

This study argues that effective evaluation practice requires the identification of the key underlying mechanisms that influence change and understanding of how mechanisms interact with contextual factors. This study notes that it is also possible for mechanisms to reinforce context. This section provides evidence from fieldwork participants of this study about these elements of context and mechanisms and how they operate. Findings in this section show the links between the elements of context, mechanism, and outcome presented in a form of a table. This is arranged to better portray evaluation practice effectiveness by determining what makes evaluation practice work effectively or ineffectively in particular circumstances.

As described in the methodology chapter, an intervention theory for each of the case study locations is formulated in this study based on the fieldwork data. As for the elements of context, most can be drawn directly from the answers found in the interviews and focus group discussions. Additionally, some elements of context are drawn from the analysis of official documents. As for mechanisms, some of them are drawn directly from the answers of participants and some of them are drawn by inferring from their answers.

In this study, context refers to the characteristics of the stakeholders of an evaluation practice. It may also include the nature of the issue in question as well as the social, political, and geographical features that affect the implementation of evaluation practice. It is important to understand how a specific element of context acts to produce a specific outcome as it plays an essential role in the implementation of evaluation practice, while mechanism operates within a particular context in order to generate a particular outcome.

6.2.1 Effectiveness of Evaluation Practices in West Sumatra Province

In order to unpack the effectiveness of evaluation practices, this section explains the Context-Mechanism-Outcome configuration (CMOc), which is usually called the “intervention theory”, in

relation to case study locations of this research. In this section, findings from West Sumatra Province are described. The findings are presented under the headings of the three categories of effectiveness, which are procedural, transactive, and substantive. The fourth category, known as normative effectiveness, is not a part of the analysis in this study, as noted in the methodology chapter. This section begins with an overview of the key elements of context pertaining to the important mechanisms identified. This is followed by a description of how the mechanisms functioned, and how they generated the outcome. Most of the mechanisms are inferred from the interviewees' answers, considering that mechanism in a realist perspective is more in-depth and more about the underlying reasoning and resources.

6.2.1.1 Procedural Effectiveness

Procedural effectiveness refers to an assessment of how an evaluation practice complies with acceptable standards and principles. The findings show that a number of elements of context play a role in supporting the key mechanisms in the procedural effectiveness category. The comprehensive Context-Mechanism-Outcome configuration under the procedural effectiveness category for West Sumatra Province can be seen in Table 6.1. The first column shows the identified elements of Context which are believed to activate the specific Mechanisms, seen in the second column, to operate and produce the third column Outcome Level 1, which are direct outcomes that lead to the next level of outcomes seen in column four, Outcome Level 2 . All the configurations in this chapter aim to portray the most significant relationship between elements of context and mechanisms. However, in reality, the participants indicated that almost all of the identified elements of context are contributing somehow to trigger mechanisms, although some are more significant than others. Therefore, these configurations aim to highlight the most significant relationships.

Table 6.1 CMO Configuration in West Sumatra Province under the Procedural Effectiveness Category

Context	Mechanism	Outcome Level 1	Outcome Level 2
<ul style="list-style-type: none"> • Conflicting regulations • Conflicting timeline between planning and evaluation • Too many reporting systems • Limited capability of practitioners • Limited quantity of practitioners 	Limited understanding	Evaluation practitioners at RDPAs tend only to follow procedures that they understand and have more pressure from higher authorities	Evaluation practice is partly effective procedurally
	Low confidence		
	Overwhelm		
<ul style="list-style-type: none"> • Political mandate/decentralised governance • Limited evaluation culture • Low stability of position 	Low commitment		
	<ul style="list-style-type: none"> • Presence of legal framework • Political mandate/decentralised governance • Compliance culture 		
<ul style="list-style-type: none"> • Presence of legal framework • Political mandate/decentralised governance 		Top-down pressure	
<ul style="list-style-type: none"> • Conflicting timeline between planning and evaluation 	Confusion	Evaluation is not well integrated into the planning process	
<ul style="list-style-type: none"> • Limited networks of evaluation unit • Limited support from senior management • Limited evaluation culture 	Low commitment		
	Limited awareness		
<ul style="list-style-type: none"> • Conflicting regulations • Limited capability of practitioners • Conflicting timeline between planning and evaluation 	Limited understanding		
<ul style="list-style-type: none"> • Limited networks of evaluation unit • Limited support from senior management • Limited evaluation culture 	Limited inclusiveness	Not all relevant stakeholders are actively involved in the process	
	Low commitment		

Context	Mechanism	Outcome Level 1	Outcome Level 2
<ul style="list-style-type: none"> • Limited evaluation culture • Low stability of positions • Low stability of institutions 	Limited awareness		
<ul style="list-style-type: none"> • Conflicting timeline between planning and evaluation • Limited capability of practitioners • Limited references 	Confusion	Most evaluation results are considered unclear	
	Limited understanding		
<ul style="list-style-type: none"> • Limited evaluation culture • Limited support from senior management • Limited networks of evaluation unit 	Limited consensus		
<ul style="list-style-type: none"> • Limited networks of evaluation unit • Limited support from senior management • Limited evaluation culture 	Limited inclusiveness	Evaluation results only shared with limited stakeholders	
<ul style="list-style-type: none"> • Limited capability of practitioners • Poor quality of evaluation results 	Limited confidence		
<ul style="list-style-type: none"> • Low stability of positions • Low stability of institutions 	Low commitment		

The participants on the whole agreed that legal framework/regulations and political mandates regarding evaluation practices are in place which trigger the top-down pressure mechanism to occur. However, the participants also stated that some of the existing national regulations conflicted with each other and that regulations were applied to all provinces in a one-size-fits-all approach without taking into consideration the different characteristics and conditions. Additionally, some interviewees argued that the existing regulations are too complex and rigid. They said that those regulations are not in line with the spirit of decentralisation and they also suggested that those regulations should be reviewed and simplified to fit with what is actually needed by the local governments. Those elements of context may have contributed to the limited understanding and low confidence evident among the practitioners at RDPAs:

Then the evaluation reports that are currently mandated are too complicated so they must be simplified and there are too many of them. And the majority of evaluations are not suitable to regional needs, especially those mandated by the

central government. One of them was under Permendagri [Ministry of Home Affairs Regulation] Number 86 Year 2017. In trying to fulfil the mandates contained in that regulation alone, we have run out of time and resources, especially human resources. This can be seen from the physical thickness of this regulation, the amount of detail and the number of issues it regulates; we are also confused just trying to understand it. (Evaluation Practitioner in the Regional Development Planning Agency)

Additionally, the limited evaluation culture in West Sumatra and the low stability of positions for evaluation practitioners led to the limited commitment. On the other hand, commitment, including commitment to follow the rules, is considered to be essential. One of the evaluation practitioners spoke about the importance of commitment and also the training to improve their capabilities:

In my opinion, the most important mechanism is commitment to the rules. Even though there are many rules regarding evaluations, so it requires commitment to obey the existing regulations. If there is a commitment, so I will understand better oh this must be this. Later, we will also think about how we can improve our analytical skills as evaluators, this is what is still lacking. The quality of our labour, our quality to analyse is still lacking. It needs additional training as well as how to improve the analysis for the evaluators. We don't just collect data, right? We should analyse the data more. From the analysis, we will look at this, later we can provide input and follow-up in the next planning. (Evaluation Practitioner in the Regional Development Planning Agency)

Moreover, some of the interviewees stated that the strong compliance culture affects the compliance mechanism to operate and push the evaluation practitioners at RDPAs to follow the existing procedures that they could understand and follow within the limited resources. However, due to the limited capabilities and quantities of evaluation practitioners, the participants reported that they were overwhelmed with all of the evaluation mandates. Consequently, they had to prioritise those evaluations that higher authorities, such as the MOHA (seen as the “big boss” of the Local Governments), considered the most important. They hoped that there will be capacity building on evaluation to help the evaluation practitioners at the local level by explaining what they really need to evaluate in the future, whether the planned activities or programs have been achievable or well implemented:

The big boss of local governments is indeed the Ministry of Home Affairs. So we have to comply even though we think most of the forms are complicated and not that beneficial for us. (Evaluation Practitioner in the Regional Development Planning Agency)

From the document analysis (as explained in Chapter 4), it can be confirmed that the legal framework to integrate evaluation into the planning process does already exist. One of the most discussed regulations by the participants is the MOHA Regulations Number 86 Year 2017 which stated clearly that evaluation should be part of the regional development planning documents. However, the findings showed that the policy makers were confused by the evaluation results because they could not obtain the needed information. The existing evaluation results did not meet their expectations. The following statement from one policy maker highlights the limited consensus of the indicators among different stakeholders (limited consistency of used indicators), and their confusion regarding the evaluation results:

That's why I'm also confused. When we refer to the RPJMD [Regional Medium-Term Development Plan], that indicator is our guide. The indicator is like this, this is the direction of our policy. Right from policy down to program, program down to activities. But when, try to compare chapter nine in the RPJMD with the evaluation results. I'm just confused. Why is this indicator being assessed, why is this indicator being assessed. The basis (of the evaluation) is different again. Well, they think, especially now if you look at Permendagri 86 of 2017 the same as Permendagri 54 of 2010, that's the first part of the data and information. It's different from our chapter nine. Then, the PE Directorate measures other things. Yes, that's why it's confusing and so the results of the evaluation don't connect with what we need for planning. (Policy maker participant)

One of the academics mentioned that the limited understanding about the evaluation methods creates confusion among the evaluators:

Tools or methods for conducting evaluations must also be mastered, but now it is still unclear, which creates confusion, especially among the evaluators. (Academic participant)

Additionally, many participants reported that there is conflicting timeline between planning and evaluation and too many reporting systems from different ministries which also trigger the confusion among evaluation practitioners. One of the policy makers raised the concern that evaluation results only tended to be made available at a very late stage, when the plan document was due to be finalised. Therefore, the policy makers could not use the evaluation results as the basis of their planning because of the late availability. Additionally, most of the evaluation practitioners complained about the overlapping reporting system:

It's just that we have to report on a lot of evaluations [to the different ministries]. That's what actually takes a lot of time. There is the annual evaluation, LKPJ [Description Accountability Reports], and so on. Can't we just produce the one

report? Sorry, but it's just the same report adapted for different technical forms. In my opinion, if we still have to report to the top [national level] we should all reach an agreement [between different ministries] to integrate it into a single, comprehensive report. That's what we'd like to happen. (Evaluation Practitioner in the Regional Development Planning Agency)

Similarly, another evaluation practitioner also reported about too many reporting systems and also the lack of coordination between the regulating ministries:

We also hope for a synergy of coordination between the ministries. [We need to know] what the definition and expectation of these evaluations are from the perspective of the Ministry of Home Affairs, from the perspective of the National Development Planning Agency, and from Menpan [Ministry of Administrative and Bureaucratic Reform]. Why Menpan? Because Menpan also issued a regulation on evaluation whereby local governments must report. Because each ministry places different values on the results of the evaluation, consequently, local leaders will have different views. This makes us hypocrites. In practice, the regions want AA scores [the highest score] from Menpan, so we follow Menpan's formula, and we think the evaluation pattern is also different, the regions don't understand. Then for LKPJ or LPPD [Regional Government Administration Report], the Ministry of Home Affairs' formula is used. (Evaluation Practitioner in the Regional Development Planning Agency)

Interviewees revealed that the Evaluation Division in West Sumatra Province RDPA was only formed in 2017. Before, it was a sub-division under the Statistics and Development Control Division. This shows that perceived importance of evaluations was increased in 2017, but also shows that there was little institutional stability in this province. Moreover, because it was a new division, the networks of this division with other divisions in RDPA and other working units outside of the RDPA were still limited. Additionally, evaluation culture within the RDPA was also limited where there was limited acceptance on the importance of evaluation. The findings indicate that those elements of context might affect the limited *awareness* and commitment of senior management about this situation and about their role in integrating evaluation into planning process. As an outcome, evaluation was not well integrated into the planning process.

Another identified unintended outcome is that not all relevant stakeholders are actively involved in the evaluation process. Some interviewees stated that the limited networks of evaluation unit and the limited support from senior management in facilitating active communication or networks between evaluation unit with other working units limited inclusiveness and awareness. On the other hand, the participants including policy makers and NGO officials, highlighted the importance of involving all of the relevant stakeholders in order for evaluation to be effective:

Evaluation can be said effective from a process point of view, maybe when involving people who are involved and who are affected by the activity itself. Those who make plans must also be involved in the evaluation, the people who benefit from the activity and the program also participate, that's from the process side. (NGO participant)

Furthermore, interviews also revealed a lack of stability of positions and institutions in West Sumatra Province. One focus group participant mentioned that in one of the municipalities in West Sumatra Province there were cases where some officials were promoted one day and demoted the next. Focus group participants reported that this had happened due to political reasons:

There were cases where some officials are promoted that day and the next day they were demoted. Some also promoted that day in the morning, but in the afternoon they got dismissed. (FGD Participant from RDPA)

This statement demonstrates the low the stability of positions is in West Sumatra Province. All of these elements of context arguably influence the commitment of stakeholders in the evaluation process.

The other unintended outcome identified is that most evaluation results are considered unclear. Several policy makers reported that they perceived the evaluation results as compilation of data only, whereas there were no clear insights arising from them. This could be derived from the limited understanding of stakeholders (especially the evaluation practitioners), and also lack of consensus among them. The confusion and limited understanding seemed to occur due to the limited capability of practitioners. All evaluation practitioners mentioned that there were very few capacity building initiatives in evaluation:

The understanding of the evaluation method is still limited, so we only do what is possible, so the results of the evaluation are also plain. Of course, this also makes us not sure that the results are quality and will be used. (Evaluation practitioner participant)

They also reported that most of them were learning by doing, were not supported by proper knowledge and skills, and also there was no routine capacity building in evaluation. No one in the Evaluation Division had a formal background education in evaluation, and only one had the opportunity to participate in a workshop on evaluation. Therefore, they felt they needed to have a routine capacity building in evaluation:

So, we need training on evaluation. And the person who does the evaluation or participates in the training is expected to be continuous, meaning that this person

should not be a functional planner who can move from evaluation to planning but there should be an evaluator function as well. So, they are consistent and continuously evaluate, eventually they will show from that habit, oh yes, there are these additions that must be evaluated so that their knowledge and evaluation skills will increase from time to time. Oh yes, this is a recommendation that we have to give. So, it is not only evaluating but also can make a recommendation against what has been evaluated. So that's what we expect like that. (Evaluation practitioner participant)

Additionally, the evaluation practitioners reported that the conflicting timeline between planning and evaluation meant that usually the required data to conduct a proper evaluation was not available in time to finish the evaluation and submit it as part of the planning process. Furthermore, the participants also commented that there was limited consensus around evaluation results where different stakeholders had different expectations. This was associated with the lack of an evaluation culture and also with limited support from senior management in trying to reach consensus.

The limited networks of the evaluation unit, limited support from senior management, limited evaluation culture, and low stability of positions and institutions were considered as significant contextual factors which affected the mechanisms of limited inclusiveness and low commitment to occur. Additionally, the participants stated that the limited capability of practitioners and poor quality of evaluation results negatively impacted their own confidence in delivering the evaluation results. As an outcome of these contextual factors and mechanisms, the evaluation results were only shared with a limited set of stakeholders.

6.2.1.2 Transactive Effectiveness

As explained in Chapter 3, this study argues that transactive effectiveness refers to the proficiency in use of resources to achieve objectives. Regarding resources for evaluation practices, most of the participants regarded the quantity and quality/capability of practitioners as the most important resources. Unfortunately, all the participants reported that they were inadequately trained and experienced practitioners. Additionally, the number of practitioners available was very limited. Moreover, they stated that financial resources and references on evaluation, such as books, journals, and online resources on evaluation, were vital resources in addition to human resources. References which were addressed by the participants were literature on evaluations in Bahasa

Indonesia and other supporting documents other than the existing government regulations. However, the majority of participants stated that they did not have access to reference literature on evaluation in Bahasa Indonesia; they only had access to existing government regulations. Consequently, all of the evaluation practitioners interviewed reported that they had limited understanding and that they were overwhelmed. One of them highlighted that they have to conduct many evaluations but have limited time and resources which made them feel overwhelmed:

According to the regulation, there are a lot [of evaluations]; RPJMD, RPJP, Renstra, RKPD, Renja, City, District. All of them have to be done but not all have been carried out due to limited time and resources and because there are so many. We were overwhelmed so we had to prioritise. (Evaluation Practitioner in the Regional Development Planning Agency)

Similarly, another practitioner highlighted that they did not have the time and resources, especially human resources, to conduct all of the mandated evaluations. Additionally, some of the interviewees reported that the IT system support for evaluation was almost non-existent. Participants mentioned that the West Sumatra Provincial Government is developing an online planning system which is called the *Sakato* Plan. Some of the officials mentioned that they were planning to incorporate evaluations into the *Sakato* Plan but they did not know exactly when it would be realised.

Both evaluation practitioners and academics identified the importance of sufficient financial resources to conduct the expected evaluations. One of the academics stated that they needed sufficient funds to conduct the evaluation properly, including the data collection for evaluation. However, based on their experience, they claimed that the budget for evaluation was still insufficient:

One more that is needed, commitment to support financially. We sometimes need to find data, but it's not supported. Continuing to budget for data seems like the most distant priority. (Academic participant)

Although evaluation practitioners raised concerns about inadequate funding for evaluation and limited understanding on their part, one of them mentioned that if there is strong commitment then they will still do the best evaluation possible with the available resources:

If you have a commitment, anything can be done. Whatever, even though there is no funding, it can still be done. If we have committed, we actually commit and run everything we can do while learning step by step. (Evaluation practitioner participant)

Moreover, the evaluation practitioners stated that the commitment for evaluation is still low within RDPA. Some of the influencing contextual factors were the limited evaluation culture and also the lack of specification of roles for evaluators. Furthermore, stability of positions and institutions in West Sumatra Province was reported to be lacking. Therefore, the commitment was relatively low because the evaluation practitioners could be re-assigned anytime to other positions outside of the evaluation division. They suggested that the national government, specifically the Ministry of National Development Planning/BAPPENAS, should develop a dedicated career path for evaluation practitioners in public institutions, similar to the dedicated career path for planners which is called the functional planner:

Bappeda also has several functional planners. We hope that Bappenas will not only make it [dedicated career path] for planners, but for evaluators too. (Evaluation practitioner participant)

Participants also mentioned that the limited awareness of the importance of evaluation and the need to properly manage the resources for evaluation also played an important part. At the time of interview, the participants reported that awareness and appreciation for planning and evaluation were still far from proportional. Therefore, most of the resources were allocated for the planning process while the evaluation process only received the remaining resources.

All of the above descriptions regarding transactive effectiveness in West Sumatra Province are summarised in Table 6.2. From the findings presented, it can be concluded that the key elements of context and mechanisms were not supportive. Therefore, the resources for conducting evaluations in West Sumatra Province were found to be inadequate and the specification of roles was absent. Consequently, evaluation practice was transactively ineffective.

**Table 6.2 CMO Configuration in West Sumatra Province under the
Transactive Effectiveness Category**

Context	Mechanism	Outcome Level 1	Outcome Level 2
Limited evaluation culture	Limited awareness	Resources have not been managed properly	Evaluation practice is transactively ineffective
Limited support from senior management			
Limited capability of practitioners	Limited understanding		
Limited evaluation culture			
Limited quantity of practitioners	Overwhelm		
Limited capability of practitioners			
Limited IT system support for evaluations			
Limited support from senior management	Low commitment		
Limited evaluation culture			
Limited financial resources			
Low stability of positions			
Limited capability of practitioners	Limited understanding	Unclear specification of roles	
Limited evaluation culture			
Limited support from senior management	Low commitment		
Low stability of positions			
Low stability of institutions			

6.2.1.3 Substantive Effectiveness

Substantive effectiveness was described as the achievement of expected objectives. Criteria that were drawn from the interviews include “support accountability” and “support informed decision-making”. All participants agreed that monitoring and evaluation are means to support accountability in regional development planning which is supported by existing regulations both from the national government and also from the provincial government. There are legal frameworks and political mandates for evaluation practice to support accountability. These two elements of context and also the compliance culture appear to trigger the top-down pressure and compliance mechanisms. These top-down pressure and compliance mechanisms are triggered by the pressure from the existing regulations and mandates which dominate and push the evaluation practitioners to do their best in conducting evaluation practices to support public accountability.

The respondents reported that one of the most significant evaluation practices which support accountability is the Performance Evaluation of Programs/Activities as Direct Expenditures of Provincial Government Budget (*Evaluasi Capaian Kinerja Pelaksanaan Program/Kegiatan*

Belanja Langsung Urusan APBD PROVINSI). Usually known as “*Evaluasi Triwulanan*” or quarterly evaluation, which requires every working unit to evaluate their achievements of financial realisation/absorption each quarter. The participants stated this is arguably the most effective evaluation practice where almost all relevant stakeholders actively participate. This is because the evaluation results are monitored very closely by the Governor. These results are announced quarterly and also discussed in the quarterly meeting led by the Governor where all of the heads of working units are required to be present. Therefore, it can be a humiliation for the head of the working unit if the performance of their unit is below the expected target. All participants mentioned that the political will of the Governor to use humiliation worked as an effective disincentive for working units to perform to a high standard in order to avoid the humiliation. As a result, this led to compliance and commitment to complete designated tasks on time:

Because every Working Unit with the Governor has a performance agreement where we promise to complete activities as much as 100% in terms of performance, and above 95% in terms of finance. So, the absorption of funds is at least 95%, that is what we promise to the Leader. So as a result, in the mind of OPD [Regional Working Unit] is how to spend the budget, how to achieve such a target without really thinking about what has been implemented, what is the long effect, what is the impact on society. Because if they don't reach the target, they will be asked by the Governor so they tend to do their best to avoid being exposed to their shortcomings in the forum. (Evaluation practitioners in RDPA)

This factor of close monitoring by the Governor also indicates that the performance measurement's high relevance with the accountability role of evaluation is one of the significant elements of context under this effectiveness framework category. Furthermore, although the capability of practitioners was limited, this element of context promoted the commitment and self-determination of the evaluation practitioners.

One of the participants also mentioned that the senior management and the local leaders were already beginning to show concern about the importance of evaluation. However, this concern was limited to the accountability role only because it was strongly linked to the budget. The interviewee stated that the focus of senior management and the local leader was mainly on the absorption of the budget and assessing whether the budget realisation goal had been achieved, not about the real outcomes and impacts of development interventions.

Another important aspect to highlight is that under the new Law of Local Government, Law No. 23 Year 2014, local leaders must submit their accountability reports on time. If they do not comply

with this obligation, the national government will apply a penalty. There are stages of penalty starting with a reprimand letter, then special coaching and, finally, if necessary, the temporary dismissal of the local leader. This is considered as a strong disincentive to failure in reporting on time which motivates the local government to focus on complying with this accountability report obligation but not focusing on the quality of information.

Table 6.3 CMO Configuration in West Sumatra Province under the Substantive Effectiveness Category

Context	Mechanism	Outcome Level 1	Outcome Level 2
Presence of legal framework	Top-down pressure	Evaluation practice support accountability	Evaluation practice is not fully substantively effective
Political mandate/decentralised governance			
High association between accountability and performance measurement	Compliance		
Compliance culture			
Growing IT system support for evaluations			
Political mandate/decentralised governance	Growing political will		
Limited evaluation culture			
Supportive political system			
Limited evaluation culture	Commitment		
Supportive political system			
<ul style="list-style-type: none"> • Evaluation is not well integrated into the planning process • Problematic planning process • Most evaluation results are considered unclear • Limited evaluation culture 	Limited understanding	Evaluation practices could not support informed decision-making properly	
	Limited evaluation culture		Limited appreciation
Low stability of positions	Low commitment		
Evaluation results only delivered to limited stakeholders	Limited inclusiveness		
<ul style="list-style-type: none"> • Local leader as the “director” of the regional development planning • Political factors • Limited evaluation culture 	Growing political will		

Almost all of the interviewees and participants in the focus group mentioned how evaluation practices should support and inform decision-making. This is one of the supporting statements:

So, in my opinion, the evaluation results are effective when the regional head considers the evaluation results important and uses them for the next development planning. (Academic)

However, the participants reported that all of the key elements of context are not in a supportive condition, since evaluation was not well integrated into the planning process, the planning process was problematic, most evaluation results were considered unclear, the evaluation culture is limited, and sharing of evaluation results was only with a limited number of stakeholders. Therefore, these elements of context arguably affect the mechanisms of understanding, appreciation of evaluation, commitment, and inclusiveness.

This study notes that the Governor as the local leader showed recognition of the importance of evaluation to support policy making. One of the indications being the commitment to lead every three-monthly evaluation meeting, which can be seen as significant political will to support evaluation. However, the interviewee also stated that the current evaluations were not fully effective because most of the evaluations were output evaluations only and the results were unclear since not all indicators were measurable and supported by strong analysis. Therefore, the interviewee reported that it was difficult to properly use evaluation results to support decision-making. There were many times when the experts from the statistics bureau were asked for support with the latest data and then the inner circle team would make some kind of rapid analysis based on those data to support the decision-making process. Further, the interviewee stated that the main reason for ineffectiveness of evaluation was due to the limitations of evaluation practitioners, both in quantity and quality. However, it can be concluded from the interviewee's statements that there was significant political will to support evaluation. as can be surmised from the local leader's description of efforts to improve the evaluation practices in the region:

There are regular quarterly meetings where all SKPDs are required to attend without being represented and I always chair the meeting. The evaluation is quite effective but not completely because we are in the process of improving. Yes, so the hope is that there will be outcome evaluations, not only outputs. Then the quantitative and measurable indicators are being compiled, and we are also learning how to evaluate them properly. Ah, to be honest, us, the West Sumatra Regional Government, it can be said that quality evaluators are very insufficient. The quantity is lacking, the quality also shows that it is not high quality. Both are not as expected. (Local leader participant)

In line with the statement from the Governor, one of the policy makers reported that the low quality of evaluation results affected the quality of their planning:

Without analysis, the evaluation results are just numbers, and sometimes the numbers don't align with the performance indicators we use. So the planning misses the target. That's the problem. Underprepared and not on target so it doesn't solve the problem. Here's an example, I plan for 2019, for instance. I'll look at performance indicators but there is no data let alone analysis. I don't know what the conditions are like and, in the end, the planning is on track but doesn't answer the question. We must know, for example, that the target is five but only reaches three, then what's the problem? We don't want there to be a problem but not know what it is. (Policy maker in the Regional Development Planning Agency)

Additionally, academics also reported that the utilisation of evaluation was low because of it being mainly limited to output evaluation:

Evaluation cannot be used because most of the evaluations are output evaluations only. If I look at the document, it's only output, not at a higher level. (Academic participant)

Most of the policy makers pointed out that the current evaluation results did not significantly support them in policy-making. One of them even stated that the existing evaluation practices were useless:

In my opinion the current evaluation is of no use. Rather, I think that the [evaluation] division's role is unclear so it would be better to disband it. (Policy maker in the Regional Development Planning Agency)

Policy makers, members of local parliament, and the NGO official also mentioned the copy-paste practice in the planning process where the officials copy the same policies or programs from the previous period of planning into the next period of planning. So, it seems that they acknowledged it as a common practice and could not do anything significant about it although they all knew that it was not a good practice:

So, it's true when it's said that the SKPD [Regional Working Units] or the local government just copy paste from the previous year when preparing the RKPD [Regional Government Working Plan]. Because we don't know the results from the previous one, where the problem is. In the PE [Controlling and Evaluation] Division, and I've said this so many times, when the target was five but they only achieved four, they were asked what the problem was, not just the numbers, but why it was like this, where's the analysis? (Policy maker in the Regional Development Planning Agency)

For example, in 2016, we wanted to reach a HDI [Human Development Index] of a certain figure, a certain figure for APM [Net Enrolment Rates] achievement, economic growth of a certain rate. We in the DPRD [Local Parliament] ask, "what are the indicators that were considered?" Secondly, once it's determined, which

OPD [Regional Working Unit] supports the indicator? Based on what, how much funding is needed, which OPD is implementing it? They don't understand, and they don't understand evaluation either. So there is the term 'copy paste'. From previous years. It was shown that, at the Department of Agriculture last year, 10% was copied from the previous year with an additional 10% more. It shouldn't be like that. (Local parliament member participant)

Moreover, the policy makers stated that there were so many different factors and stakeholders that they needed to consider when developing regional development plans. These included the national strategic agenda, which is based on the new law of local government, Law No.23/2014, the aspirations of the citizens, the political direction of the Local Leader, and also the evaluation results. Considering there is a conflicting timeline between planning and evaluation, as briefly alluded to in the procedural effectiveness section, usually the required information from evaluation was not available when the policy makers needed to finish the planning process. Consequently, they said that it was almost impossible for the policy makers to utilise the evaluation results in informing their next period of planning. Therefore, sometimes some of them needed to improvise by using the copy-paste practice.

Academic participants also questioned the existing role of evaluation in supporting informed decision-making in the regional development planning processes. One of them stated:

In my opinion, it's unsatisfactory. First, because the data is not readily available, second, because the indicators are also not well established from the outset. (Academic participant)

However, both academic participant and local member of parliament noted that the real problem was the planning itself. Both groups of stakeholders highlighted that the regional development planning processes in West Sumatra Province were problematic:

...the evaluation can only be conducted properly if the plans are developed correctly; only then it can be evaluated. But the plans and the planning process themselves are the problem. (Academic participant)

[Regional development] planning [in West Sumatra Province] is currently not effective, so the evaluation is similarly ineffective. (Local parliament member participant)

Additionally, based on the decentralisation and regional autonomy policy, the local leader now has more authority to direct their own region's development plans. As a result, it is relatively difficult to avoid the local leader's personal interest and circumstances in influencing the decision-

making in regional development planning processes, including whether the local leader's political will is considered in the evaluation results. One of the participants reported how frustrating it was for evaluation practitioners when their recommendations arising from the evaluation results were overlooked and never followed up by the policy makers, and particularly by the local leader:

Sometimes we have prepared in a certain way. Whether this is used in the planning or not is unclear. At most it's used as baseline data for future planning. But sometimes the results of the evaluation are also not considered in the planning process. Those who formulated the plans had evaluation data, but they still proceeded without considering the results of the evaluation. (Evaluation Practitioner in the Regional Development Planning Agency)

One of the evaluation practitioners said that policy makers may choose not to act on recommendations because political interest is considered more important. Therefore, evaluation practices could not support the decision making properly:

There is a political element. This means that there are activities that we deem unnecessary, but these activities also involve heads of agencies, heads of divisions so they ask that the activities are retained. There is also a request from a member of the parliament. That's how it is, so we maintain the program or activity even though we have submitted the evaluation results [that the program/activity is unnecessary]. Frankly speaking, the EKPD [a type of evaluation practice] that we put together from a team of experts, they have already lost interest. How come our recommendations aren't followed up by the Governor? We submit them every year when we do an evaluation. After all, the same OPD [Regional Working Unit] was in the wrong. (Evaluation Practitioner in the Regional Development Planning Agency)

Additionally, one of the academics also mentioned that evaluations are very political:

The problem is also that the evaluations are a very political burden, it's just lip service. In actual fact, they just want to meet compliance measures, they don't care about the content and don't care about the quality so long as there's a report. (Academic participant)

Consequently, the evaluation practices could not properly support informed decision-making in the regional development planning process. Therefore, evaluation practices in West Sumatra Province were not fully or substantively effective.

From all the explanations and interviews discussed in this section, it can be deduced that there are different sets of elements of context and mechanisms which are perceived to work together in producing the effective evaluation practice under the three dimensions of effectiveness, these being procedural, transactive, and substantive effectiveness. Furthermore, it can be concluded that

in West Sumatra Province, most of the elements of context which impact the supportive mechanisms did not operate properly in producing the expected outcomes. Therefore, the majority of the evaluation practices in in this province were procedurally, transactively, and substantively ineffective.

6.2.2 Effectiveness of Evaluation Practices in DI Yogyakarta Province

Similar to the previous section, this section will unpack the hypothesised causal pathway between context, mechanism, and outcome which is mainly based on the perspectives of the participating stakeholders in the second case study location, DI Yogyakarta Province. The findings are presented under three headings: procedural, transactive, and substantive effectiveness.

6.2.2.1 Procedural Effectiveness

As with West Sumatra Province, the data from DI Yogyakarta shows broadly identical elements of context which significantly determine the key mechanisms under the procedural effectiveness to operate. The main difference is that DI Yogyakarta have more positive contextual factors compared to those in West Sumatra Province. Consequently, the positive mechanisms have better support to operate and produce better outcomes than in West Sumatra Province.

The comprehensive Context-Mechanism-Outcome configuration under the procedural effectiveness category for DI Yogyakarta Province can be seen in Table 6.4. As with Table 6.1, which portrays the CMOC under the procedural effectiveness category for West Sumatra, the first column in this table shows the identified elements of context which are considered to activate the specific mechanisms, shown in the second column, to operate and produce the Outcome Level 1 which leads to Outcome Level 2.

Table 6.4 CMO Configuration in DI Yogyakarta Province under the Procedural Effectiveness Category

Context	Mechanism	Outcome Level 1	Outcome Level 2
<ul style="list-style-type: none"> • Conflicting regulations • Conflicting timeline between planning and evaluation • Too many reporting systems • Limited capability of practitioners • Limited quantity of practitioners 	Overwhelm	Evaluation practitioners at RDPAs tend only to follow procedures that they understand and have more pressure from higher authorities	Evaluation practice is partly effective procedurally
<ul style="list-style-type: none"> • Continuous support from academics 	Growing understanding		
	Growing confidence		
<ul style="list-style-type: none"> • Political mandate/decentralised governance • Growing evaluation culture • Relatively stable position 	Adequate commitment		
<ul style="list-style-type: none"> • Presence of legal framework • Political mandate/decentralised governance • Compliance culture 	Compliance		
<ul style="list-style-type: none"> • Presence of legal framework • Political mandate/decentralised governance 	Top-down pressure		
<ul style="list-style-type: none"> • Conflicting timeline between planning and evaluation 	Confusion	Evaluation is not well integrated into the planning process	
<ul style="list-style-type: none"> • Adequate networks of evaluation unit • Adequate support from senior management • Growing evaluation culture 	Adequate commitment		
	Growing awareness		
<ul style="list-style-type: none"> • Conflicting regulations • Limited capability of practitioners • Conflicting timeline between planning and evaluation 	Limited understanding	Almost all relevant stakeholders are actively involved in the process	
<ul style="list-style-type: none"> • Growing evaluation culture • Supportive political system 	Growing political will		
<ul style="list-style-type: none"> • Adequate networks of evaluation unit • Adequate support from senior management • Growing evaluation culture 	Adequate inclusiveness	Some evaluation	
<ul style="list-style-type: none"> • Growing evaluation culture • Relatively stable positions • Relatively stable institutions 	Adequate commitment		
	Growing awareness		
<ul style="list-style-type: none"> • Conflicting timeline between planning and evaluation 	Confusion		

Context	Mechanism	Outcome Level 1	Outcome Level 2
<ul style="list-style-type: none"> Limited capability of practitioners Limited references Continuous support from academics 	Growing understanding	results are considered unclear	
<ul style="list-style-type: none"> Growing evaluation culture Adequate support from senior management Adequate networks of evaluation unit 	Growing consensus		
<ul style="list-style-type: none"> Limited networks of evaluation unit Limited support from senior management Limited evaluation culture 	Adequate inclusiveness	Evaluation results shared with almost all stakeholders	
<ul style="list-style-type: none"> Limited capability of practitioners Adequate quality of evaluation results 	Growing confidence		
<ul style="list-style-type: none"> Relatively stable positions Relatively stable of institutions 	Adequate commitment		

As reported by participants in West Sumatra Province, the majority of participants in DI Yogyakarta Province also agreed that legal framework/regulations and political mandates regarding evaluation practices were in place which triggered top-down pressure and compliance mechanisms to occur. Likewise, the participants in DI Yogyakarta also complained about the conflicting regulations regarding evaluation practice and the too heavy burden of many reporting systems. One of the evaluation practitioners pointed out these issues:

In terms of the timing of the governance itself, our situation is quite difficult, especially with the Permen [Ministerial Regulations], the management time is very tight, right? For example, [a deadline like] one month after the [end of the] fiscal year and so on, that is what we often find difficult to fulfil because it is related to the data. And in terms of the content, really, we are still somewhat confused. If we follow the Permendagri [Minister of Home Affairs Regulation] to the letter, it's rather difficult, especially the formats, the formats and also the substantial analysis, sometimes it's very difficult. Meanwhile, on top of the Permendagri, there are also other regulations. Different regulations require different formats and indicators or measures that must be evaluated. (Evaluation practitioner in the Regional Development Planning Agency)

Participants in DI Yogyakarta complained about too many reporting systems but acknowledged they were mainly for compliance purposes:

With LKPJ from the Ministry of Home Affairs, for example, the substance is the same as LKJIP from Kemenpan, there are just a few different sections, but I think it's only a superficial difference. If the substance requires the same data, why should there be two reports with the same content? So there are some things that

might not have been efficient so far in that regard. (Evaluation practitioner in the Provincial Working Unit)

After [discussions] in the focus group and the like, it turned out that the main problem with the evaluation was that it was done purely for reporting purposes. So we write a lengthy evaluation report and then they just file it, what I mean is that there's no follow-up. The year after, it was much the same; there were also evaluation reports that more or less had the same content because the evaluation results were never followed up on. (Academic participant)

The conflicting regulations and excessive reporting systems, combined with other elements of context, such as limited capability and quantity of practitioners, were indicated as key elements of context which made the evaluation practitioner feel overwhelmed. However, some of the evaluation practitioners reported that they were supported by the academics as their partners in conducting evaluation practices. So, whenever they were confused or stuck, including when there was a new regulation regarding evaluation practice which was confusing or seemed to be conflicting with other regulations, they always discussed the issues with the academics for a better understanding and help in finding solutions together. Therefore, they reported that their understanding and confidence grew continuously:

That's the macro policy of the Governor. So, in the Regional Government of DIY, one OPD [Provincial Working Unit] is accompanied by a doctoral expert from UGM [Gadjah Mada University]. Except for [the] culture [working unit] where they might have someone from ISI [Indonesian Institute of the Arts Yogyakarta]. So it is already the Governor's policy that each OPD is to be accompanied by experts, except in Bappeda. In Bappeda there are experts per sector, each sector can get one expert. We have an MoU with UGM. For evaluation, we have annual contracts with the experts, but not from the MoU package, we have experts from UGM MPKD [Magister Program in Regional and City Planning] and one from a private university. (Evaluation practitioner in the Regional Development Planning Agency)

The continuous support and commitment from the academics to be partners of the local government in conducting evaluations were also evident in this province. This is one of the statements from the academic which showed their commitment:

That's a commitment that I often pay attention to, that's why I talk to my friends too, now that commitment must be maintained, properly maintained. Personally, I will be committed to always being ready to help my friends in the local government, I will not take it into account, the important thing is that I have the ability, what I do can be useful for the region. That's my commitment. (Academic participant)

However, although they were supported by the academics, the practitioners said that they still felt that there were not enough resources to properly conduct all of the required evaluations. This, they said, was because they had too many responsibilities to conduct different types of evaluations, whereas the resources, especially the number of practitioners available, were limited. Consequently, they were feeling overwhelmed. Considering these descriptions, evaluation practitioners at RDPAs tended only to follow the prioritised procedures that they understood and were the cause of the greatest pressure from higher authorities.

Similar to the West Sumatra Province, the conflicting timeline between planning and evaluation was also considered by the majority of participants as an influential element of context in DI Yogyakarta Province. Almost all of the participants from evaluation practitioner and policy maker category mentioned this factor. One of them stated:

The problem is the time clash. So maybe the time clash was because when we were still preparing the RKPD [annual planning document] evaluation, our colleagues from the planning divisions had already requested the [evaluation] data. Sometimes, once we do the evaluations, the RKPD document is already done too. So, this is a time management issue because we can't see from the context of evaluation alone, but on the one hand, when we do evaluations, our planning colleagues always demand the evaluation results immediately. There is Renja [the working unit's annual plan], there is RKPD and all kinds of things, and sometimes, with limited human resources and so many documents, we are just overwhelmed. (Evaluation practitioner in the Regional Development Planning Agency)

During the interviews and focus group discussion, the participants stated that the conflicting timeline between planning and evaluation elements of context was a significant influence on their ability to conduct evaluations and planning effectively. Since there was little time to prepare planning after evaluation reports were received, the confusion mechanism had a stronger negative influence on them than the positive mechanisms for growing awareness and commitment. Additionally, conflicting regulations combined with limited capability of practitioners, led to limited understanding and poor outcomes. Therefore, although the networks of evaluation unit, support from senior management, and evaluation culture were considered to be adequate by most of the participants, the conflicting timeline issue still resulted in confusion and the ultimate outcome that the evaluation was not well integrated into the planning process. Additionally, from the statements of the participants, it can be concluded that the conflicting timeline added to the confusion of evaluation practitioners, whose main priority became the importance of submitting

“on time” as a basis of the planning. This in turn resulted in some evaluation results being considered unclear, which affected planning.

As Table 6.4 shows, there are several differences in terms of the state of the elements of context and mechanisms in West Sumatra Province and DI Yogyakarta Province. Compared to West Sumatra Province, DI Yogyakarta appears to have more supportive elements of context and mechanisms. For example, the networks of evaluation unit were perceived to be better by most of the participants, while the stability of positions was also perceived to be higher by the participants. Several mentioned that the special position of the Sultan as the long-term Governor of the province contributed highly to the stability of the unit. These elements of context were considered important in facilitating relevant stakeholders to preserve their commitment and to achieve better outcomes. Additionally, some participants also reported that the organisational culture in government institutions in DI Yogyakarta Province was relatively open to evaluation:

In our workplace, the culture of evaluation has improved. Previously, we only used the manual system, now the manual system has shifted to online, the online system that was developed that I mentioned earlier. The evaluation culture has developed along with that system. With regard to that, the Evaluation Division has been asked to give a presentation of the evaluation notes in the initial stages of planning. Because field A, for example, can't possibly criticise another field because the horizontal will later clash. So usually the Evaluation Division explained “previously this construction was like this, or like that”. So, for me, as a planner going forward, if the Provincial Working Unit submits it [proposal for an activity/program to be included in the next development plan], I have already heard from the evaluator first. I see it kind of like that. (Policy maker in the Regional Development Planning Agency)

In line with the policy maker's opinion, evaluation practitioners in this province also stated that recently the evaluation culture was growing, although it still needed improvement:

When it comes to openness, it's open. Even when we meet in the internal exposure meeting, we show who is the worst and who is the best, we present the conditions just as they are. We are open, these are our findings, there is an overspend, is that the case? We still say that, even though we didn't appoint the person concerned, they have already received a warning notice and that they understand that it's not to happen like that again. (Evaluation practitioner at the Provincial Working Unit)

Another evaluation practitioner highlighted the changes of evaluation culture in their region over time:

In the past, when people came to Bappeda, the Heads of OPDs were conscientious when they talked about planning because they wanted to get the money, so they themselves had to attend. But in terms of evaluation, let alone structural officials, only ordinary staff attend. But now the situation has changed over the past three years. By means of various strategies that we use, such as the assessment system, we can see the results in terms of compliance from the OPD side are pretty good, as well as the performance results. Control and evaluation now have a place. So, that is what we develop by creating systems that force them to change. So, our control and evaluation are now effective because we have developed systems. This evaluation culture changes in line with that system. (Evaluation practitioner at the Regional Development Planning Agency)

Those statements above indicate that there was a growing evaluation culture in this province.

As a result of having more supportive elements of context, some of the mechanisms in DIY were perceived to be more positive than those in West Sumatra Province, including inclusiveness, awareness, and level of commitment. Therefore, these mechanisms enabled almost all relevant stakeholders to actively be involved in the process, while evaluation results were shared with almost all stakeholders.

6.2.2.2 *Transactive Effectiveness*

For the transactive effectiveness category, the influencing elements of context are observed to be almost the same between the two case study locations. The main differences noted are that practitioners in DI Yogyakarta Province have better IT system support, better evaluation culture, and better support from senior management. The evaluation practitioners in DI Yogyakarta stated that they had an online evaluation system called *Sungguh* that had been in place since 2008:

We have a system of control and evaluation that was initially called “Monev APBD” but we changed the name to “Sungguh”, “Sungguh Jogja Prov.co.id”. This is where the OPD [Regional Working Units] must report on the entire implementation of the activities and program. Because aside from the implementation of the activities and program, we also have what’s called an ROPK application [one of the procurement applications]; it’s the Operational Plan for Implementation of Activities. Actually, it’s closely linked to the report on the implementation of the program of activities, so if the OPD claims this quarter the

achievement is, say, 10%, then it can be seen what that 10% consists of.
(Evaluation practitioner in the Regional Development Planning Agency)

However, the evaluation practitioners reported that this IT system created a problem, in that many of the Provincial Working Units were staffed by older workers whose computer and IT skills were limited, thus affecting the capability of the human resources at the unit. Additionally, due to the limited number of practitioners in the Provincial Working Units, the online reports sometimes could not be submitted on time because the person in charge was in training or on leave and no other officials were available and capable to prepare and submit the online report.

In DI Yogyakarta Province, the participants reported that they were aware that the number and capability of the evaluation practitioners, and the references about evaluation was also limited:

In terms of quantity, it's really lacking. We have three sub-divisions here, one of which is a sub-division of APBD [Regional Revenue and Expenditure Budget] Management, within which there is also a state budget which is also a strategic program. Actually, we are in the Special Region of Yogyakarta, Ma'am, we get special funding, so we also evaluate special funding, then sub-division performance with relatively limited human resources, so we have eighteen people in the whole division, if I'm not mistaken. That is for one division, with three sub-divisions. One sub-division only has one staff member. From a human resources level, staffing numbers are less than ideal. (Evaluation practitioner in the Regional Development Planning Agency)

To address the limitations of staff and capability of practitioners, one of the actions taken by the RDPA was to arrange cooperation with local experts from universities to support them. However, the evaluation practitioners also mentioned that although they had continuous support from academics, those academics were experts in development planning or in data management, not specifically in evaluation planning. They reported that usually both the practitioners from RDPA and the academics brainstormed and learned together. So, not all of their questions or issues could be directly answered or resolved with support from the academics. One of them also stated that an on-going issue they were facing was the need to develop proper instruments to evaluate the quality of government interventions considering there were limited references on evaluations in planning (such as articles and books) in Indonesia.

Table 6.5 CMO Configuration in DI Yogyakarta Province under the Transactive Effectiveness Category

Context	Mechanism	Outcome Level 1	Outcome Level 2
Growing evaluation culture	Growing awareness	Resources have not been managed properly	Evaluation practice is transactively ineffective
Growing support from senior management	Growing understanding		
Limited capability of practitioners			
Growing evaluation culture	Overwhelm		
Limited quantity of practitioners			
Limited capability of practitioners			
Growing IT system support for evaluations	Growing commitment		
Growing support from senior management			
Growing evaluation culture			
Limited financial resources			
Low stability of positions			
Limited capability of practitioners	Growing understanding	Unclear specification of roles	
Growing evaluation culture	Low commitment		
Adequate support from senior management			
Low stability of positions			

In terms of capability of practitioners, the practitioners mentioned that their knowledge and skills in evaluation were limited because most of them did not have formal education background in evaluations, and some of them were from the planning divisions. This is because government officials must accept placement in any division where they are assigned. However, the policies of the Sultan as Governor of the province has established greater stability for evaluation practitioners in comparison to West Sumatra Province where personnel may be re-assigned to other divisions without recourse. Due to the limited evaluation capability and other factors, the commitment among practitioners was not high enough but was considered to be improving with growing awareness. Additionally, capacity building for improving the knowledge and skill of evaluation practitioners was reported to be very limited:

Training or workshops on evaluation are perhaps quite uncommon here, maybe once a year, but we [staff members] have to take turns. This year, there's been almost nothing so far. (Evaluation practitioner in the Regional Development Planning Agency)

However, because some of them had a relatively high educational background, such as Masters degrees, and also long-term support from local experts, their level of confidence in conducting evaluation practice appeared to be higher compared to the level of confidence of practitioners in West Sumatra Province. One of the NGO officials mentioned that human resources in RDPA were among the best in this provincial government:

...usually Bappeda staff are the brightest ones. Those who receive the best [score in the civil servant exam] will usually be placed at Bappeda. If you're assigned there [in RDPA] it means your exam results were good, high scores are required to be placed there [in RDPA]. (NGO Official)

Despite these positive indicators and that they are highly qualified civil servants with support from experts, the evaluation practitioners said that they still felt that they did not have enough resources to properly conduct all of the required evaluations. This is because they claimed to have many responsibilities to conduct different kinds of evaluations, whereas the resources, especially the number of practitioners, were limited. Consequently, they were still overwhelmed by the task.

Participants also mentioned that awareness of the importance of evaluation was growing. However, they also commented that awareness of the need to properly manage the resources for evaluation was relatively lacking. Most of the resources in this province were allocated for the planning process, leaving only the remaining resources for the evaluation process. Therefore, it can be concluded that the resources for evaluations in this province were not managed properly in a balanced manner that would contribute to their effectiveness.

Since the specification of roles for evaluators was still lacking, an evaluation practitioner could be assigned a role in the Planning Division in Bappeda or even to other positions in the Regional Working Units. Consequently, similar to the suggestion from participants in West Sumatra Province, participants in DI Yogyakarta Province argued for the necessity to build a specific evaluation position and career in the public sector institutions, especially Bappeda:

...there needs to be a functional evaluator who knows what to do in the evaluation, what evaluation method is good to use, what recommendations will be given, what are the internal factors that make this evaluation into this, what are the external factors, whether the SWOT method should be used, for example, right? (Evaluation Practitioner Participant)

Considering these findings from interviews with the evaluation practitioners, it can be concluded that the key elements of context and mechanisms under the transactive effectiveness in DIY

Province, although an improvement on those in West Sumatra Province, were not supportive enough to achieve the desired outcomes. The resources for conducting evaluations in DI Yogyakarta Province were not managed properly and the specification of roles for evaluation practitioners was still unclear. Therefore, as an overall outcome, the evaluation practice in the DIY Province was still not considered to be transactively effective.

6.2.2.3 *Substantive Effectiveness*

Similar to the opinions of the participants from West Sumatra Province, all participants from DI Yogyakarta Province agreed that monitoring and evaluation are means to support accountability in regional development planning which is supported by existing regulations both from the national government and from provincial government. It can be concluded that the legal framework and political mandate for evaluation practice to support accountability in the DIY Province do exist. These two elements of context, together with the compliance culture, appeared to trigger the top-down pressure and compliance mechanisms. These top-down pressure and compliance mechanisms were triggered by the pressure from the existing regulations and mandates which dominated and pushed the evaluation practitioners to do their best in conducting evaluation practices to support accountability.

Likewise, the participants in DI Yogyakarta also reported that one of the most significant evaluation practices which supported accountability was the quarterly evaluation or “*Evaluasi Triwulanan*”. In DI Yogyakarta it is usually called “*Rapotan*” or quarterly reports. In this type of evaluation practice, every working unit has to evaluate their achievements of financial realisation/absorption and physical achievements per quarter. The participants stated this was arguably the most effective evaluation practice where almost all relevant stakeholders actively participated. This is because the evaluation results were monitored very closely by the Governor. These results were announced quarterly and also discussed in the quarterly meeting led by the Governor where all of the heads of working units had to present. Therefore, it would be humiliating for the head of the working unit if their unit had a lower performance than the expected target. These factors illustrate the strong political will of the Governor as the local leader in ensuring the system of evaluations functions to support accountability:

So this is what we do, based on the quarterly results, all OPD [Provincial Working Units] will be given a report card, like on the 18th [of this month]. The Governor will issue the report card for the entire OPD. But in DIY [Province] those meeting with the Governor are usually the best and the worst. This is done deliberately to shame people as a psychological strategy. From the perspective of evaluation, it turns out that this is more effective for managing performance compared to remuneration. Because [if] the head of the OPD, as the manager, is humiliated [for underperformance], especially in DIY [Province], it is published in the newspaper. Not only it is announced in the quarterly meeting, but it gets published in the print media. It seems that the [impact of] psychological stress is quite extraordinary compared to if it [the resulting sanction] were only linked to remuneration. They might then think if they were unable to reach the [desired] TPP [remuneration or employee bonus], that is fine. This month they will just let the TPP fall short, but next month they will try to do better. That's what would happen if it were only about money. But if it affects their good name, their reputation, well, they will usually care about that and then call upon their executive [to work harder]. So that's why we [also] issue the worst report card, that is actually a form of punishment. (Representative of the Governor)

This statement by a representative of the Governor has also been confirmed by an official in one of the Provincial Working Units who highlighted that there are incentives and disincentives in DIY Province as part of the political will to support evaluations. This participant reported that humiliation acted as the disincentive to failure to meet the requirements of reporting on time, while retaining a performance financial bonus acted as an incentive to complete the task well:

In the past, financial targets were achieved by a certain percent, ok tick, the target was achieved and that's it. Now, that's not enough, there must be evidence and that will later influence the OPD's [Provincial Working Unit's] ranking in the report card. The performance outlined in the report card also influences personal income [of officials]. So, when performance isn't good, it decreases the ranking of the OPD, the institution's ranking, and this influences the acquisition of TPP. TPP is employee bonuses. Yes, it's usually called remuneration too. [The payment of] TPP is highly dependent on the achievement of agencies and, secondly, also depends on individual performance which is called SKP. Jogja [DIY Province] manages TPP well. The indicators Jogja uses for assessing TPP is considered to be good. So the reward and punishment is felt. It directly impacts the income earned by individuals. So that kind of control mechanism already exists. (Official in the Provincial Working Unit)

These statements show that performance measurement's high relevance with the accountability role of evaluation is one of the significant elements of context under the substantive effectiveness

framework category. Furthermore, the growing evaluation culture and supportive political system reinforces the commitment of relevant stakeholders in DIY province.

The participants in DI Yogyakarta mentioned that the senior management and the Local Leader were aware and concerned about the importance of evaluation. One of them stated:

The commitment level is pretty good, especially starting from the leader. As I said, the Governor, yes, he's quite concerned about it. As I said, he always leads each event relating to the report card, he always leads them every 3 months.
(Evaluation practitioner in the Regional Development Planning Agency)

However, it is important to note that this concern was mostly limited to the accountability role because of its high relevance to the budget. The interviewee stated that the focus of senior management and the local leader was still mainly on the absorption of the budget and whether the target had been achieved. They did not put much effort into determining the real outcomes and benefits of development interventions for the citizens. However, they stated that they had started to determine some alternatives to properly evaluate the factual outcomes and impacts but were still struggling considering the lack of capability and references.

Table 6.6 CMO Configuration in DI Yogyakarta Province under the Substantive Effectiveness Category

Context	Mechanism	Outcome Level 1	Outcome Level 2
Presence of legal framework	Top-down pressure	Evaluation practice support accountability	Evaluation practice is not fully substantively effective
Political mandate/decentralised governance			
High association between accountability and performance measurement	Compliance		
Compliance culture			
Growing IT system support for evaluations			
Political mandate/decentralised governance	Growing political will		
Growing evaluation culture			
Supportive political system			
Limited evaluation culture	Commitment		
Supportive political system			
Continuous support from academics	Synergy with experts		
Adequate networks of evaluation unit			
<ul style="list-style-type: none"> • Evaluation is not well integrated into the planning process • Some evaluation results are considered unclear • Growing evaluation culture • Limited capability of practitioners 	Growing understanding	Evaluation practices are partly support informed decision-making	
	Limited appreciation		
Growing evaluation culture	Growing commitment		
Low stability of positions			
Evaluation results shared with almost all stakeholders	Adequate inclusiveness		
<ul style="list-style-type: none"> • Local leader as the “director” of the regional development planning • Political factors • Growing evaluation culture 	Political will		

Almost all of those interviewed and those who took part in focus group discussions, particularly the policy maker participants, mentioned how evaluation practices should support informed decision-making. One of them said:

We plan based on data. Based on the reality of the previous year which we draw from evaluation results. That becomes our resource for planning, the things identified as obstacles experienced in the previous year becomes input [for planning]. Then we look at the conditions, the data. For example, retention figures

in the education sector, then APK [participation rate] and so on. We look at performance from the previous year. The data for 2017 becomes input for planning so that the same challenge doesn't present again or we create a new program to build upon previous achievements. (Policy maker in the Regional Development Planning Agency)

However, the participants reported that some of the key elements of context were not supportive in DIY Province since evaluation was not that well integrated into the planning process and most evaluation results lacked detail, depth, and were still considered unclear. These results were mainly an effect of the conflicting timeline between planning and evaluation, which caused the process to be rushed and inadequate. Therefore, these elements of context arguably affected the understanding and appreciation of evaluation. The policy makers and NGO participants mentioned that the evaluation results might not properly inform the planning process because it lacked the required information:

In my opinion, it's only the level of detail that still needs to be improved. So far, the evaluation is still limited to the realisation of programs [budget] and macro indicators, but for the quality of activities and programs, the conditions have not yet been reflected [in the evaluation data]. (Policy maker in the Regional Development Planning Agency)

It's not yet in-depth, not yet in accordance with actual needs. If you look at the rules, that's what the rules are like. There definitely must be creativity; in my opinion that creativity must come from the apparatus. (NGO Official)

In agreement with this opinion from the NGO Official, one of the evaluation practitioners stated that the evaluation results form might not properly support the planning process because the existing regulations mainly demanded that they focus on the accountability role of the evaluation practice:

Particularly those [evaluation questions] based on the form from Permen 54 or 86 aren't yet able to answered, it's still just what the problem is and how much was achieved. It can't capture the strategic technical obstacles if a problem like that arises. The analysis still lacks depth. (Evaluation practitioner in the Regional Development Planning Agency)

Therefore, the evaluation practitioners reported that together with the support from academics, they were continuously trying to develop more appropriate evaluations for supporting regional development planning processes within the available resources.

Another participant shared their perspective on how the political factor becomes influential and may take precedence over other considerations in the regional development planning processes:

At first, the draft of the plan was determined but then, due to certain factors, in the midst of the process, someone said that “this” [additional activity] must be included or the target recipient should be changed. Those changes were due to political factors but I can’t refuse directions from the higher-ups. It still happens, but it’s not as bad as in other regions, in certain cases it still exists. It's not too central like it was in the past; in 2013 it was still brutal. It was usually [resulting from] a request from member of parliament. (Policy maker in one of the Provincial Working Units).

Additionally, academics also had similar opinions about the influence of politics in planning:

If I can be honest, our country has tended to overdo the political aspects. So any planning has been tampered by political aspects so that everything is based on the short term. Why is it all happening, because of the cost of politics in Indonesia. Because we are still a political country, right, the cost of politics is very high. Every 5 years you have to replace [the regional head], the maximum 10 years have been completed, this is “aji mumpung” in Javanese. So those who are elected try to get a minimum return on investment or get a profit within those 5 or 10 years. (Academic participant)

One of the policy makers in RDPA expressed their opinion that the follow-up of evaluation results most of the time is out of the RDPA control. It is more up to the Heads of the Working Unit as the Person in Charge (PIC) of development programs and activities and the direction of the Governor:

In my opinion, the conditions in Jogja [DIY] now all depend on the PICs [Heads of Provincial Working Unit as the PIC of development programs and activities]. Now the larger the sector they are responsible for, the greater the likelihood that the results of monitoring or evaluation might not be followed up on promptly. Only strategic issues are prioritised for follow up. So the key lies with each PIC and also the direction or support of the Governor. (Policy maker in the Regional Development Planning Agency)

Additionally, the evaluation practitioner respondents reported that most of the recommendations from evaluation results tend to be ignored; the Governor is the one who decides whether to follow up the recommendations or not. One of them stated:

Actually, there is already a regulation in place that our evaluation results must include recommendations to the Governor to be conveyed to the agencies. But it depends on what the Governor's response is like. There are several recommendations that become problematic in planning documents. The same problems are identified and reported over and over again. The problems and recommendations are reported year after year but there has never been any real

intervention. (Evaluation practitioner in the Regional Development Planning Agency)

In line with the evaluation practitioner's statement, several academic respondents reported that the Governor had a strong influence in the planning process. This is understandable and common across regions in Indonesia because provincial governors hold significant power, since decentralisation policies have created greater regional autonomy. However, these participants reported that sometimes the personal circumstances and interests of the Sultan, as Governor, and the royal family also significantly influenced the planning process:

Oh there are still many political factors at play here, Ma'am. Even though the leader tends to remain the same, unlike other regions that change all the time, but the Governor has a family. There are a lot of stories, like when the eldest daughter became the chair of the KNPI for youth [National Youth Committee], the Disdikpora [Provincial Working Unit for Education, Youth, and Sports] was the funding source. If KNPI wanted to undertake any activities, there had to be funding from them [Disdikpora]. I don't know how that will manifest in Renja [annual working plan of the Provincial Working Unit], and there are lots of other stories [like that]. (Academic participant)

This participant also added that the RDPA could not do anything to refuse the demands from the Governor although sometimes the request was not in accordance with the regional development plan:

The Head of Bappeda always acquiesces to whatever the Governor wants. He is the Governor's right-hand man. Whatever the Governor wants will be ensured by the Head of Bappeda. Although sometimes not in accordance with RKP or RPJMD, for example. However, the Governor can adapt it to fit. It's a bit complicated, actually. When I was asked why Yogya is good [in terms of governance]? My answer is the bureaucrats. The bureaucrats are highly skilled and able to interpret directions [from the Governor] which are often very abstract, and they are always able to improvise [to meet demands]. (Academic participant)

Another participant mentioned one of the cases which was personally initiated by the Governor but because it was not based on the development plan then it developed into a controversy. However, no one dared to officially report or criticise the Governor and the matter was quickly settled without further embarrassment to the Sultan or the Governor's office. From this finding, it is important to note that the special position of the Governor as the Sultan has a significant, perhaps ultimate influence in the decision-making process of DI Yogyakarta Province regional development planning:

NGOs are good here [DIY Province], the bureaucrats are good, it's only political parties at the provincial level that wouldn't dare face the Sultan [Governor]. In 2007, there was a big case about the CDMA project or the cable. Actually, it was clearly the direction of the Governor but, yes, in the end it was the Regional Secretary who was held responsible. The regional Ombudsman Institution was established by the Governor and funded from the Regional Budget. There's only one rule there and that is that they cannot criticise the Governor. (Academic participant)

This participant also stated that for evaluation results to be followed up by the Governor, they needed to be formulated and presented in such a way that was in line with the personal circumstances and interests of the Governor:

[The Governor is] open-minded but cannot be criticised. Because he is [also] the king, so if we meet him, we have to be deliberate in choosing our words, our tone of voice and intonation to ensure what we say is respectful and received openly. Once he accepts our idea, with the results of the evaluation that we submit, then it will be immediately followed up. That's what the pattern is like. So, there is a dimension of [the Governor] wanting to listen so long as it doesn't bother him. (Academic participant)

The descriptions by this academic participant suggest that the political will of the Governor as the local leader is a significant mechanism which determines whether evaluation results are considered usable or utilised in informing decision-making.

Despite these indications of improper process, one of the academic participants stated that evaluation practices in DI Yogyakarta had improved in terms of effectiveness. Moreover, the opinion was that DI Yogyakarta is the leader in this area when compared to other provinces:

In my opinion some evaluations are relatively effective, but I really take my hat off to my colleagues [evaluation practitioners in the RDPA]. We will continue to learn, yes, even as academics who are sometimes requested to collaborate, we learn from the process. For example, the evaluation starts with focusing only on the input, so the budget spending. Then it is included as part of the performance measurement, the assessment of the report cards that is conducted every three months. After that, we are now thinking not only of input but also output; outputs of activities or programs are discussed. Then, yes, it was included as an innovation improvement process from the procedure scheme that was included in both the hard system and the human system. For example, when it comes to performance incentives, it's a system, but also a software system. With e-planning, Jogja has been learning. We evaluate development but we also evaluate how we evaluate it ourselves. I think Jogja is, in some respects, somewhat leading compared to other regions. (Academic participant)

Based on the explanations and interviews with respondents, it can be concluded that the evaluation practices in DI Yogyakarta were partly effective in informing decision-making in the regional development planning process, but that there were significant anomalies in processes of approval and decision-making in regional development. Furthermore, it can be concluded that DI Yogyakarta Province had more favourable elements of context compared with West Sumatra Province. Consequently, the supportive mechanisms, such as understanding, confidence, commitment, and inclusiveness, could operate better in producing the expected outcomes. Therefore, evaluation practices in DI Yogyakarta were partly effective procedurally and substantively. However, although some elements of context under the transactive effectiveness performed better compared with West Sumatra Province, most of the key elements of context were still not in a favourable condition. As a result, the evaluation practices in DI Yogyakarta were still transactively ineffective. It is also interesting to note that the special power of the Governor as the Sultan and traditional leader of the Sultanate had a significant influence on the decision-making process in the regional development planning of DI Yogyakarta Province. Consequently, the effectiveness of evaluation practice was also highly dependent on the political will, awareness, and also appreciation by the Governor of the importance of evaluation. Additionally, it is also interesting to highlight the role of humiliation as a means of enforcing political will in both case study locations which led to compliance of government officials. This point will be discussed further in Chapter 7.

6.2.3 Identified Mechanisms and Contexts

This section summarises key mechanisms and elements of context identified in the fieldwork. The three columns in Table 6.7 show whether the mechanism is considered to generate effective evaluation practice or to generate ineffective evaluation practice or both. Based on the literature and fieldwork findings, it is important to note that *political will* mechanism is believed to contribute to both effective and ineffective evaluation practice, depending on whether the political will supports evaluation practice or tends to ignore it. Each of the mechanisms in this table will be discussed further in the next chapter. However, considering the next chapter focuses on the ‘what works’, two of the identified mechanisms (*overwhelm* and *confusion*) are the opposite/negative

forms of the other mechanisms. Therefore, the next chapter only focuses on the eleven positive mechanisms which support the effectiveness of evaluation practice.

Table 6.7 Summary of Identified Mechanisms and Elements of Context

MECHANISMS	Fieldwork	
	Generates effective evaluation practice	Generates ineffective evaluation practice
(1)	(2)	(3)
Understanding	√	
Commitment	√	
Inclusiveness	√	
Top-down pressure	√	
Compliance	√	
Confidence	√	
Overwhelm		√
Confusion		√
Awareness	√	
Consensus	√	
Appreciation	√	
Political will	√	√
Synergy with experts	√	

Additionally, based on the empirical data discussed in Section 6.2, Table 6.8 portrays the working positive mechanisms in each of the case study locations. It can be noted that compared to West Sumatra Province, DI Yogyakarta Province has slightly better working mechanisms which support effective evaluation practice, including adequate inclusiveness, growing consensus, adequate synergy with experts, growing understanding, adequate commitment, growing confidence, and growing awareness. Consequently, the mechanisms are more influential in producing better outcomes than in West Sumatra Province.

Table 6.8 The Identified Positive Mechanisms in Each Case Study Location


No.	Mechanisms	Case Study Location	
		WS	DIY
1	Inclusiveness	Limited	Adequate
2	Consensus	Limited	Growing
3	Synergy with experts	Limited	Adequate
4	Understanding	Limited	Growing
5	Commitment	Low	Adequate
6	Confidence	Low	Growing
7	Awareness	Limited	Growing
8	Appreciation	Limited	
9	Top-down pressure	High	
10	Compliance	Adequate	
11	Supportive political will	Growing	


Furthermore, based on the empirical findings, this study has mapped the relationships between each of the identified elements of context with the identified mechanisms, as can be seen in Table 6.9. The first column of this table shows that there are nineteen elements of context which are considered to influence the effectiveness of evaluation practice in the regional development planning context. The next columns show the key mechanisms that have been identified based on the findings from the fieldwork. The green box illustrates a positive relationship while the red box illustrates a negative relationship.

Table 6.9 Relationships between Elements of Context with the Identified Mechanisms

Elements of Context	Mechanisms												
	Inclusiveness	Consensus	Synergy with experts	Understanding	Commitment	Confidence	Overwhelm	Confusion	Awareness	Appreciation	Top-down pressure	Compliance	Political will
Harmonious legal framework				Positive	Positive			Negative			Positive	Positive	Positive
Political mandate/decentralised governance					Positive						Positive	Positive	Positive
Compliance culture					Positive							Positive	
Integrated reporting systems				Positive			Negative	Negative					
Sufficient capability of practitioners				Positive		Positive	Negative	Negative	Positive	Positive			
Sufficient quantity of practitioners							Negative						
Continuous support from academics	Positive	Positive	Positive					Negative					
Harmonious timeline between planning and evaluation				Positive		Positive		Negative					
Adequate networks of evaluation unit	Positive	Positive	Positive										
Adequate support from senior management	Positive	Positive			Positive								
Strong evaluation culture	Positive	Positive		Positive	Positive			Negative	Positive	Positive		Positive	Positive
Stable positions					Positive								
Stable institutions	Positive				Positive								
Sufficient financial resources					Positive		Negative					Positive	
Sufficient references				Positive				Negative					
Adequate IT system support for evaluations												Positive	
Sufficient political system	Positive		Positive		Positive				Positive			Positive	Positive
Good evaluability of plans				Positive				Negative					
Highly articulated roles and responsibilities for evaluation					Positive	Positive	Negative						

Notes:

Positive relationship 

Negative relationship 

As the table shows, each of the elements of context was perceived to trigger different mechanisms. Additionally, some elements of context appear to have more influence than the others. Table 6.9 reveals “strong evaluation culture” as the most significant element of context which was considered to have a significant influence in triggering most of the identified mechanisms to operate in generating effective evaluation practice. Participants from a majority of stakeholder groups, including evaluation practitioners, policy makers, academics, NGO officials, and local leaders, referred to the importance of evaluation culture although they did not mention the exact term due perhaps to their limited knowledge of evaluation culture. However, they stated their strong concern related to culture where evaluation becomes an integral part of the organisation culture. Similarly, NGO official, academic, and policymaker participants also expressed the view that evaluation culture is definitely needed to support the effectiveness of evaluation:

What supports the effectiveness of evaluation is a critical culture, a very critical organisational culture, a culture that is very comfortable with criticism, with feedback, comfortable with news of failures and then a culture that motivates or provides incentives for those who are critical or can analyse. Culture which is very open, that is, a culture that is very open to evidence. The point is open to data-based information. A culture that is less sensitive to criticism or failure. (NGO participant)

Moreover, an academic participant expressed a similar view:

The culture of the Regional Government, especially the Provincial Government, must support evaluation, such as they are open to criticism of evaluation results, actively carry out and utilise evaluation results, and so on. (Academic participant)

Additionally, a national policy maker stated that evaluation culture is important but not yet existing in regional governments in Indonesia. The participant reported disappointment with many RDPAs which only instructed third parties to conduct the evaluations without any substantial supervision from the RDPA. That is one of the reasons, it was stated, why the evaluation results were not high quality and meaningful for informing policy making because the assigned third party did not exactly know what was expected from the evaluations:

There may not yet be a culture for evaluating. The majority of Bappeda only carry out routine work, I mean. Just business as usual. And that was because they didn't know how to do it so to make it easier, they only ordered a third party to do it. This kind of culture is what gets in the way. The third party is allowed. If I see it, we can cooperate. Only the concept must be from the local government, my concept is like that, please work on it. (National policy maker)

In line with the view of the national policy maker, an evaluation practitioner also stated that the problem at the moment is the lack of evaluation culture:

Maybe the problem is the culture. Not accustomed to seeing evaluation as important so that the commitment is also limited. If it is not the Governor who leads directly, then representatives of the Head of Service will come, only the staff or the three echelons. (Evaluation practitioner)

Additionally, one of the NGO officials pointed out that the current bureaucratic culture is problematic:

In this area, yes, it is not yet effective, so maybe this factor is also a factor in the regional head. Regional head factor then cultural factor, bureaucratic culture. That is, bureaucratic culture. Not saying that everything is like that, but the bureaucrats say that the regional government they do the things they are obliged to do, but that's it. (NGO participant)

All those individual perceptions from the interviews about evaluation culture were also confirmed in the focus groups in both case study locations. In the focus group in West Sumatra Province, all participants, including the evaluation practitioners, policy makers, and academic, agreed on the importance of evaluation culture which was expected to lead to prioritisation and institutionalisation of evaluations in their organisations:

....there needs to be an organisational culture that is normally accustomed to regarding evaluation as something that must be done and is indeed the main criterion in determining future planning. So if it has become a culture we usually don't become afraid, Ma'am, to say that oh this is actually enough, no need to continue based on the evaluation results. Now we don't have the courage, ma'am, there are those who do not depend on the individual. (Policy maker participant in FGD)

Similarly, the discussion from the focus group in DI Yogyakarta also agreed that evaluation culture is the most important element of context:

In my opinion the first is the culture, right? If that culture has been built, the system will follow itself, evaluation will run on its own. (Evaluation practitioner participant in FGD)

Yes, it's the same, because if the culture works, evaluation will run on its own, so if people change, the quality can still be maintained because there are already values. (Evaluation practitioner participant in FGD)

From this section, it can be noted that there are different sets of key elements of context which activate the influential mechanisms to function in producing effective evaluation practice under the procedural, transactive, and substantive categories of effectiveness. From the case study locations of this research, it was revealed that not all the existing elements of context were favourable. As a result, both positive and negative mechanisms were activated. This study has focused on positive mechanisms that generate effective evaluation practice in order to provide substantial recommendations to improve the effectiveness of evaluation. Therefore, the discussion in the next chapter focuses on the positive mechanisms, especially the most influential mechanisms.

6.3 Conclusion

This study adapts the effectiveness framework from an environmental assessment context for evaluation practice in regional development planning. The Effectiveness Framework developed in this research is useful in framing and determining the state of evaluation practices in case study locations while the Context-Mechanism-Outcome Configuration is useful in unpacking the elements of context and mechanisms which influence the effectiveness of evaluation practice. The CMOC is a proposition that summarises the hypothesised causal pathway between context, mechanism, and outcome which are mainly based on the perspectives of the participating stakeholders. Therefore, the effectiveness framework and the CMOC complement each other.

From the findings, it can be determined how the group of stakeholders in regional development planning assessed the evaluation practice effectiveness based on the effectiveness framework slightly differently. For example, in West Sumatra Province, the evaluators and policy makers (the state stakeholders) seemed to view the state of evaluation practices more favourably than the non-state stakeholders. It is interesting to note that the significant elements of context in both case study locations are broadly identical and almost all are interrelated. However, the findings show that DI Yogyakarta Province have slightly more favourable elements of context compared to those in West Sumatra Province. Consequently, the mechanisms are more positive in DI Yogyakarta Province than in West Sumatra. Additionally, it can also be determined that some elements of context seem to be more significant compared to the others, especially evaluation culture. Similarly, some mechanisms appear to be more significant compared to the others, including inclusiveness, commitment, and political will. These will be discussed further in the next chapter, as will other positive mechanisms that support the effectiveness of evaluation practice.

Another matter that can be drawn from the findings of the fieldwork is that, although evaluation practices in both case study locations were not fully effective procedurally, transactively, or substantively, they do serve as effective means for accountability. This is further discussed in Chapter 7.

CHAPTER 7 DISCUSSION

The aim of this chapter is to further elaborate the findings of this research by drawing on the literature review in Chapter 2 and empirical evidence presented in Chapter 6. This will enable the researcher to address the main research question: How and to what extent are evaluation practices effective in supporting regional development planning processes? The previous chapters explored literature relating to the study, outlined relevant theoretical and methodological frameworks, and presented findings from the case studies. Chapter 6 has partly addressed the questions of this study by examining the effectiveness of evaluation practices in case study locations using the effectiveness framework, and the elements of context and mechanisms that are considered to be significant. In summary, this chapter discusses the findings from the two case study locations, compares them, and shows the relationships between empirical findings and literature to find answers to the questions this study has posed.

The chapter is divided into three sections. Section 7.1 addresses the first sub-question of this study which is about the definition of an effective evaluation practice in relation to regional development planning. It links the literature in Chapter 2 and the empirical findings presented in Chapter 6 to develop a working definition of effective evaluation practice in the domain of regional development planning.

Section 7.2 discusses the role of evaluation practice in regional development planning in order to answer the second sub-question of this study by linking the literature and the findings from empirical data. It can be concluded that the findings in this study are broadly consistent with and complement the existing literature that the main rationale of conducting evaluation in regional development planning is to further support the planning process considering evaluation has two primary roles: the accountability role and the learning or improvement role.

Section 7.3 addresses the last sub-question of this study, which is: “What are the circumstances that aid evaluation practices to be effective especially in supporting regional development planning process? What are the factors that create an enabling environment for effective evaluation?” This section discusses further the identified key elements of context and mechanism based on the empirical data from Chapter 6 and links them with the relevant literature. The last section concludes the discussions in this chapter.

7.1 Working Definition of an Effective Evaluation Practice

The results of the systematic scoping review at the beginning of this study showed that there is no explicit definition in the literature on effective evaluation practice in supporting regional development planning. The purpose of this section is to develop a working definition. According to the framework developed in Chapter 3 and the findings from the fieldwork in Chapter 6, an effective evaluation practice in regional development planning is one that meets all of the criteria in the procedural, substantive, transactive, and normative effectiveness dimensions. In other words, evaluation practice can be categorised as effective when it implements the principles governing evaluation process, achieving the set objectives, delivering the outcomes at minimum resources, and achieving the normative goals of evaluation practice. This is the first effectiveness framework developed for evaluation practice in the field of regional development planning. Therefore, it is important to note that the examples of literature used in this section only provided partial descriptions of the criteria necessary for the proposed effectiveness framework.

Based on the effectiveness framework developed in this study, as presented in Chapter 3, this study proposes that an evaluation practice can be perceived as effective *procedurally* when it implements all the principles governing evaluation process, including following relevant procedures, integrating evaluation into the planning process, involving stakeholders in the process, delivering results to relevant stakeholders, and the evaluation results having objectivity, clarity

and understandability. The findings of this study are broadly aligned with the literature reviewed in Chapters 2 and 3 including those by Baker and McLelland (2003), Bond *et al.* (2013), Cashmore *et al.* (2004), Chanchitpricha and Bond (2013), and Theophilou *et al.* (2010). An effective evaluation, according to the participants of the study, include the following elements: regulatory relevance, uses appropriate methods, it is an integral part of the planning process, it has power in the planning process. Furthermore, an effective evaluation should be congruent with plan documents, should involve relevant stakeholders and should reflect a consensus amongst them, and should have relevant benefits for targeted groups. Additionally, the participants also define an effective evaluation as having clear results, is easy to understand, is credible and objective, uses relevant indicators, is of a high quality, has measurable results, is supported by strong analysis and reliable data, and is well communicated.

Moreover, this research argues that an evaluation practice can be considered effective *substantively* when it achieves set objectives, including supporting accountability and informed decision-making. This is aligned with the literature which has been discussed in Chapter 2 where Kusak and Rist (2004) and Seasons (2002b) discuss how evaluation practice supports accountability, while Davies (2012), Head (2016), Sanderson (2002), and Segone (2008) discuss evaluation as one of the supporting evidences to inform key decision makers in policymaking. Participants in this study also echoed similar criteria although not using exactly the same terms, where they mentioned these criteria when defining an effective evaluation: informing the public, including routine reporting, utilised, and supporting decision-making.

Additionally, an evaluation practice can be deemed effective *transactively* when it delivers the outcomes at minimum resource, including managing resources properly and having specification of roles. Many articles in the existing literature have discussed these criteria, for example Seasons (2002b: 45), Olejniczak (2013: 1662), and Waldner (2004). Waldner (2004) highlighted that resources need to be managed properly considering planners are often involved in the “front-loading” of resources where most of the resources are allocated towards the plan development while the evaluation of plans only receives minimal resources. Participants in this study also

reported similar points, where they highlighted these criteria as part of an effective evaluation definition: resource efficient (including human resources, time and space, and finance), sufficiency of IT system support and practitioner capability, and specified roles for evaluators.

Finally, this study proposes that an evaluation practice can be categorised effective *normatively* when it helps achieve the normative goals of evaluation practice, including improving the plan or policy and developing or making changes in relevant institutions. Although no study in the evaluation practice literature is using the term “normatively effective evaluation practice”, many studies discuss the related criteria under this normative effectiveness category including Cousins and Bourgeois (2014), Head (2016), Oliveira and Pinho (2011), Roberts (2006), and Torres and Preskill (2001). Head (2016) stated that rigorous evaluation practices can significantly improve policy development. The present study also provides empirical evidence where participants reported similar criteria. They tend to view an effective evaluation as the one which contributes to improvement of policy or development plan, and also improvement of evaluation culture in an organisation.

One of the interesting findings is the difference between respondents in different stakeholder groups in terms of their emphasis on the categorisation of evaluation effectiveness. Most of the evaluation practitioners indicated that they were more concerned about the procedural effectiveness of evaluation practices, where they were expected to comply with existing mandates. However, they were also very aware of criteria in other categories of effectiveness, especially the “support accountability” criterion. One of the them stated:

Too many types of evaluations are mandated. Most of them were through Permendagri 54/2010 which has now been replaced by Permendagri 86/2017. Once we have conducted the evaluation and submitted the report to the central government, there's never been any feedback provided to us, it was unclear what the evaluation was for. Our impression is that we have to submit the evaluation reports just to fulfil the obligations, we just have to comply. (Evaluation practitioner in the Regional Development Planning Agency)

The policy makers, on the other hand, showed more inclination towards substantive effectiveness, but this differed according to their position. For example, policy makers in the Regional Development Planning Agencies, the leading agents involved in the formulation of the regional development plans, expected the evaluations to support their decision-making process in formulating the regional development plans:

Evaluation is very important because it informs future planning. When I was preparing a plan, the direction of the policy was available but the information to determine the target should come from the evaluation results. So, in my opinion, an evaluation is effective if it can inform us, as policy makers, in determining appropriate targets in planning [documents] (Policy maker in the Regional Development Planning Agency)

In line with the previous statement, another policy maker also expressed this expectation. The policy maker reported that the lack of information from the current evaluation results resulted in continuation of the same programs for consecutive years without a clear basis. Therefore, there were hopes that future evaluation results could inform the planning team about which programs and activities should be included in future development plans:

It is hoped that, based on the results of the evaluation, we can decide that a program isn't required because the target has already been achieved, for example. Or that it's not the kind of program that the community needs based on the results of the evaluation. But the reality is that there are programs/activities that are run each year without a clear basis. So, an effective evaluation can determine whether a program/activity needs to be continued or not, or needs to be changed or enhanced or reduced. I think that is the ideal. (Policy maker in the RDPA)

Additionally, the policy maker in the Ministry of National Development Planning Agency highlighted his ministry's expectations of evaluations at the regional level:

Evaluation is very important because the results of an evaluation become the foundation for planning. So based on evaluation, the regional government can set more rational targets, then it's more focused. Now, for Bappenas, the evaluation-based planning in the regions means that Bappenas knows exactly which regions have the biggest problems and need intervention so that, later on, the gap analysis of the achievements or standards that have been set can be quantified. If we have service standards, the results of the evaluation each year show which regions are still below

the standards and we can fill that gap. So, the results of evaluations in the regions also become input for planning at the national level. (Policy maker in the Ministry of National Development Planning)

However, not all policy makers at the national level, such as the officials from the Ministry of Home Affairs, shared the same perspective. Some of them saw evaluation mainly as a tool for accountability and somehow disregarded evaluation as important input for planning:

Indeed, every year and every quarter, Bappeda must report the results of their evaluation to us so that we can see the extent to which the plans they have made have been carried out. That's all. (Policy maker in the Ministry of Home Affairs)

Respondents representing NGOs and academics, on the other hand, were more inclined to emphasise substantive effectiveness, with a particular focus on the effectiveness criterion of “support informed decision-making”. One of the NGO respondents argued:

Evaluation should also be able to be the basis for planning, especially in the determination of development targets. But in reality, that setting of targets does not use the evaluation results. And this is what I think might be the big challenge. (NGO official)

Almost in the same vein, one of the academics stated that:

Evaluation is expected to make the planning make more sense; making sense in terms of whether it really addresses the existing needs and problems. Because planning is supposed to be a means of problem solving anyway. That's what it's about. Continuing to make sense of the planning means that we should learn from the results of the evaluation, right? For example, if a particular target is set, and it turns out that the evaluation results show that it's not possible, then we have to be more realistic in setting targets. (Academic participant)

Respondents in this study revealed a range of criteria in defining effective evaluation practice. A close reading of their responses showed that what the respondents revealed is not far different from theoretical positions on evaluations. However, it is important to note that this is the first study in the evaluation practice field to categorise the criteria into the four dimensions of effectiveness discussed above. Additionally, findings of this study provide more detail to the expected criteria of effective evaluation practice.

In summary, this study has proposed a comprehensive and systematic working definition of an effective evaluation in the regional development planning context where effectiveness is categorised into procedural, substantive, transactive, and normative effectiveness. It is the first definition that has been proposed for an effective evaluation considering there is no other definition that can be found in the existing literature in the planning, policy, and evaluation arena. Furthermore, the fieldwork data contributes to understanding which criteria belong to each of the four categories of effectiveness. Additionally, the fieldwork findings also add to knowledge by showing divergent viewpoints between respondents in different stakeholder groups in terms of their emphasis on the categorisation of evaluation effectiveness. Now that the answers for the first sub-question of this study have been presented, the next section will address the second sub-question of this study.

7.2 Role of Evaluation Practices in Supporting the Regional Development Planning

The findings from empirical data and analysis of this study align with the conclusions from literature (explained in Chapter 2) that evaluation practices support regional development planning processes in two ways. First, evaluation practice supports accountability (Seasons, 2002b; Kusak and Rist, 2004). Second, evaluation practice supports learning and improvement (Faludi, 2000; Torres and Preskill, 2001; Patton, 2002; Sanderson, 2002; Seasons, 2003; Roberts, 2006; Segone, 2008; Davies, 2012; Cousins and Bourgeois, 2014; Head, 2016). It is important to note that the perspectives of the participants presented in this section relate to the ideal roles of evaluation practice, not the current condition of evaluation practice in their regions.

Evaluation supports accountability by assessing the extent to which the objectives of an intervention are being or have been achieved. Evaluation can be used as a means to promote good governance, better accountability, innovation and reforms, and modern management practices

(Kusak and Rist, 2004: 60). Evaluation is considered to have a significant role to good governance, especially for enhanced transparency, although it is context-specific (Dahler-Larsen and Boodhoo, 2019). Seasons (2002b: 52) states that evaluations are conducted to inform program stakeholders and to ensure accountability by an organisation. In the evaluation context, accountability serves as a means to secure legitimacy. Evaluation is now considered a legitimising function and good governance practice (Chouinard, 2013). One of the greatest benefits of evaluation is that it holds planners and those involved in plan preparation accountable; this can be a way to legitimise the field of planning (Brody and Highfield, 2005; Laurian *et al.*, 2010; Oliveira and Pinho, 2010a).

Many participants from all of the stakeholder groups in this study reported this accountability role of evaluation. All of them support the existing literature about the importance of this accountability role of evaluation in supporting regional development planning:

Its role is very important, meaning that we evaluate the past activity program whether it is useful or not. Its role is to measure achievement, find out whether there is an excess or insufficient allocation of funds, and so on. And also to find out the existing problems and the recommendations to solve them.” (Evaluation practitioner participant)

“That is very important, because number one is the evaluation result. Because from the evaluation, the results are visible. What has been achieved what has not. There are three things: suitable, achieved, or not yet achieved. So, these are three things that can be seen from the results of the evaluation. (Policy maker participant)

These perspectives are consistent with the literature which stated that evaluation can provide information about the progress, the achievement, and the value of government interventions (Leeuw and Furubo, 2008; Weiss *et al.*, 2008). Additionally, one of the academics highlighted the accountability role of evaluation to show and determine whether the current local government is performing or not. The academic also mentioned that being accountable to the community is very important. Citizens are more likely to live happily when they feel their circumstances change for the better or there is hope to have a better quality of life. Evaluation is expected to show the status of the people's lives and be able to give hope to the community to progress further. This is an interesting finding which has not been emphasised in the existing literature:

In my opinion, evaluation is to assess from the past how much has been achieved. For me, there are those for the administration section and the politics section, to be able to show a fair assessment of whether this executive is performing or not. Second, being accountable to the community is also very important. People sometimes live happily when we feel that we have changed for the better and so that's an interesting thing. Or if there is hope, there is progress. Our evaluation can show the status of our community life and so on, how much we are going up or down or so so, it might also give people hope and happiness and so on. (Academic participant)

A participant from the NGO group also perceived that evaluation can help the government to assess whether they are on the right path:

Ideally, of course, to see how far the regional vision and mission have been achieved, are they going in the right corridor or not. If that is important, in my opinion, monitoring evaluation includes a survey on the satisfaction of implementing program activities to achieve the vision and mission, so that the government can measure itself whether they are on the right path or not. If not, of course there can be improvements, so that they can effectively carry out their program of activities to achieve the vision and mission. (NGO participant)

Similarly, another NGO participant also argued that evaluation is needed to understand the relationship between planning and implementation and also for the government to assess whether they can achieve their vision and mission which were stated in the planning documents:

Yes, for me, that means evaluation needs to be improved, to see, the first is to see suitability. The suitability of the planning results with their implementation, the second is to prove the hypothesis. Proving their own hypothesis that in order to achieve their vision for the next 5 years that which is stated in the planning document can be proven by evaluation, it proves that. (NGO participant)

On the other hand, the learning role focuses on deriving lessons from the past experiences with the intention of informing related stakeholders in order to improve future practice. This role therefore relies on the existence of feedback mechanisms and the operational flexibility needed to function as a learning organisation (Molas-Gallart, 2012). Olejniczak *et al.* (2016) state that one of the main functions of evaluation in policymaking is to build a stronger knowledge base. Cousins and Bourgeois (2014) and Torres and Preskill (2001) also discuss how evaluation relates to organisational learning. Evaluation helps in supporting capacity building to better design,

implement and value planning interventions; and in determining lessons which can be utilised to guide future planning (Faludi, 2000; Patton, 2002; Roberts, 2006). In summary, evaluation fosters continuous learning in planning, from promoting assessment of plans, and aiding the achievement of intended outcomes (Seasons, 2003; Head, 2016), to providing constant support for improvement in the planning profession (Oliveira and Pinho, 2010b, 2010a, 2011).

Similar to the previous role, many participants from various stakeholder groups in this study also reported about the learning or improvement role of evaluation. All of them supported the existing literature highlighting the necessity of this role in supporting regional development planning. They also stated their high expectations regarding this role to improve evidence-informed policymaking in the regional development planning processes. All of the policy maker participants discussed their expectations of evaluation to support the policymaking in planning process. One of the policy makers at the national level reported concerns about the mis-interpretation of the role of evaluation in regional development planning. The participant reported that many regional apparatuses seem to forget about the real value of development planning and evaluation. The interviewee then went on to emphasise that the real importance of evaluation is not about the formality of carrying out evaluation practice but how it can significantly be a valuable support for the next period of planning process:

They do not know, many of the regional apparatus do not understand the value of development planning and its evaluation. It is thought that development is only a process of carrying out activities. Planning is structured as a tool to achieve that development value, and evaluation as a development function is needed to ensure that planning plays an optimal role in this regard. So it's not just a formal evaluation activity and fulfilling the agency's duties and functions. But it serves as input to complement planning. (Policy maker at the national level)

Many policy maker participants reported the importance of evaluation as supporting evidence to make the plans appropriate to existing needs and conditions, and also to check whether implementation is going according to plan. If it is in accordance with the plan, the evaluation is needed to identify the obstacles and what needs to be done to overcome them:

In fact, the role of evaluation is very important in development planning. For example, we plan early and then evaluate whether this is in accordance with what was implemented or not. If it is not suitable, where are the constraints and obstacles? Is this necessary the plan to be continued. If this is lacking, where is it lacking, what is needed? So as evidence for making a next plan. If there is no that can't be. This means that we just make things up again, up 10% from last year and so on. If you already have an evaluation, oh yesterday we did this oh no, this should be replaced, oh this might be the focus here, oh this one doesn't need a time because the area doesn't match, it means we need this. Well, there we can plan well according to their needs and conditions. (Policy maker participant)

Another policy maker participant also mentioned that evaluation plays a very important role in development planning in ensuring the plan implementation does not deviate from what the government is aiming for based on the agreed plans. If there are unexpected external conditions which affect the implementation of the plan, evaluation can help in identifying them and suggest recommendations on how to move forward with the current conditions. However, he also stated that evaluation can be effective when the development plans are of good quality. Additionally, another policy maker participant highlighted the need for planning to be based on evaluation results to improve its accuracy and precision. The participant mentioned that without evaluations, regional development plan is only like a dream without a clear foundation:

Yes, that was for planning the future. Without evaluation, what the success of the previous years was will not be known. The success or failure of seeing the extent to which the program has been implemented. Now to plan ahead which things are still empty, which are still not being implemented. That is what we will patch in the future from the previous year's target as a guideline for future planning. We saw the success of the achievements and problems that existed in the previous year, right like that. Without evaluation, of course we cannot make plans that are more concrete, precise and accurate. (Policy maker participant)

Additionally, a policy maker from the national level provided a similar perspective:

Evaluation is very important. Because based on the results of the evaluation, we understand the problems in the regions. So that the evidence in the form of the results of the evaluation will be used as input for planning. So that evidence based, it must be based on evaluation. Without evaluation, there is not sufficient evidence. (National policy maker participant)

The above points raised by participants are strongly related to existing literature stating that evaluation helps evidences to inform key decision makers in evidence-based policymaking (Sanderson, 2002; Segone, 2008; Krizek *et al.*, 2009; Davies, 2012; Head, 2016). Segone (2008) suggested that evaluation (together with monitoring) should aid evidence-based policy making, especially in facilitating the selection of policies, and facilitating public argumentation among policy makers and relevant stakeholders. To do so, evaluation should be both technically sound and politically relevant. Therefore, he proposed the dialogue between the suppliers and users of evidence should be strengthened to bridge the gap between the information needed by policy makers and the information provided by evaluators. Additionally, Head (2016) stated that evaluation provides evidence that contributes to balanced policies and legitimate governance. However, he preferred the term evidence-informed policy making considering that within public policy discussion it is clearly demonstrated that the objective evidence of scientific knowledge, including evaluations, cannot drive policy in a democratic political system. In this politicised context, some kinds of evidence are inevitably seen as more relevant than others for underpinning policy positions (Head, 2016).

In this study, academics and evaluation practitioners also highlighted the learning role of evaluation. Evaluation must be done to ensure that the policy objectives are achieved. If the objectives are not achieved, evaluation is expected to identify the issues to ensure that the same mistakes will not be repeated in the future. As stated by several policy makers and practitioners in this study, ideally evaluation can show lessons whether the specific interventions in the plan documents need to be continued, replicated in other areas, or stopped because the evaluation shows they are not useful:

Ideally, the objective of the evaluation is clear to ensure that the policy objectives are achieved, if not achieved, what are the gaps, whether the people or the equipment or what or the regulations? So actually it must be done because otherwise it will repeat the same mistakes. Yesterday we thought it was good, for example, but we didn't know what was good, so in my opinion it is important and someone needs to ensure that the evaluation results lead to the next policy. If we talk about evidence-based policy, for example, one of the forms of evidence is the results of the evaluation, there is no need

for field research anymore, we can identify what is actually missing there. We have this best practice, let's replicate it in this place for further policies like that. (Academic participant)

Supposedly, for example, in policy making, the main basis should be, apart from the strategic issue, the first is evaluation, because this is based on records, historically what was planned, what was the implementation. Whether it's during the implementation or after the implementation, what is it like? So the evaluation is a lesson for us whether this needs to be continued or if it wants to be bigger or the policy needs to be stopped because we think it is useless. (Evaluation practitioner participant)

Additionally, some participants also raised the role of evaluation to ensure the continuity of planning between periods. One of the NGO participants highlighted this point arguing that local leaders would change based on the elections, but regional development should be sustainable:

The role of evaluation is first, it can improve the continuity of planning between periods. If the evaluation is good, yes, if the evaluation is effective as I mentioned earlier then what happens is a continuation. Then the second evaluation makes the planning make more sense. Make sense in one sense, it really answers existing needs, answers existing problems. Because planning should be a means of problem solving anyway. So that. Continuing to make planning makes more sense, it means that the evaluation results become a lesson, yes, for example, if the target is set this way, the evaluation results show that it is impossible so evaluation makes us be more rational in setting targets. I guess I think the two are the main thing. (National NGO participant)

Based on these interviews, it can be concluded that participants' perspectives in this study converge in supporting and complementing the statements from the existing literature that the main rationale for conducting evaluation in regional development planning is to further support the planning process. Expectations of participants with respect to evaluation are broadly similar at the national level and in the two study locations.

Regarding the learning and improvement role, as discussed in earlier paragraphs of this section, this research aligned with existing literature in identifying the importance of evaluation in acting as supporting evidence to make the plans appropriate to existing needs and conditions, checking whether implementation is according to the plans, identifying the issues to ensure that the same mistakes will not be repeated in the future, and providing evidence on whether specific interventions should be continued, replicated or stopped. Moreover, in contrast to the existing

literature, this study also emphasises the role of evaluation in ensuring the continuity of planning between periods. This continuity was considered important by participants in this study since the local leaders often change with elections, and the newly elected local leader tends to have their own vision and mission. Therefore, lessons from evaluation are needed to support sustainable regional development plans. Furthermore, extending the existing literature, especially by Head (2016), this study supports the perspective that positions evaluation as one of the examples of scientific evidence that supports evidence-informed policy making considering that the regional development policy-making in Indonesia occurs in a democratic political system which is characterised by many trade-offs and compromises between stakeholders.

As stated earlier in the introduction of this section, the interviews with the participants presented in this section relate to their opinions and expectations of the ideal roles of evaluation practice. However, the empirical findings presented in Chapter 6 show that the evaluation practices in the case study locations are not very effective. Therefore, the next section will discuss what generates effective (and ineffective) evaluation practice and under what circumstances.

7.3 What Generates Effective Evaluation and Under Which Circumstances: Refined Explanation

This section discusses what generates effective evaluation practice and under which circumstances. The refined explanation was achieved by bringing together the findings from the fieldwork, especially the identified key elements of context and mechanisms which generate effective evaluation practice, as described in Chapter 6, and the literature review. This section consists of three subsections. Section 7.3.1 discusses the refined intervention theory, Section 7.3.2 unpacks the identified mechanisms in more detail, while Section 7.3.3 discusses the importance of context. To make it easier to identify mechanism and element of context in the discussions,

every mechanism will be written in *italic* while element of context will be written inside quotation marks (for example: “evaluation culture”) in the following sections.

7.3.1 The Refined Context-Mechanism-Outcome Configuration (CMOc)

This section discusses the refining of the initial intervention theory (or usually called CMOc) based on crossing analysis findings from empirical data with existing literature discussed in previous sections, especially in Chapter 2. After building CMOs specific to each case study location, cross-case comparisons and synthesis were conducted. The CMO has been refined by identifying the elements of context that were common across both case studies, re-examining the associated mechanisms and outcomes, and cross-analysing with the relevant literature. In summary, the refined CMO was developed according to the results of this study about ‘what works’.

Before embarking on details about the refined CMOs, it is important to summarise the mechanisms that have been identified in this study. Table 7.1 shows the summary of eleven key mechanisms identified based on the findings from the fieldwork (described in Chapter 6) and the developed initial intervention theory based on the systematic scoping review (in Chapter 1). The first column of this table shows that there are eleven mechanisms which are considered important to produce effective evaluation practice in the context of regional development planning. The next columns (columns 2-3) show the source from which the mechanism was derived, whether it was from the systematic scoping review or from fieldwork or from both. As the table shows, three mechanisms are consistent with those which were found in the initial systematic scoping review results and fieldwork findings; while eight mechanisms are derived from the fieldwork data only. Therefore, the researcher conducted a broader review to find related literature for each of the eight mechanisms to unpack and discuss all of the mechanisms further in this chapter.

It is important to note that the two negative mechanisms identified in Chapter 6 were not considered for this Chapter, which focuses on ‘what works’. Additionally, based on the literature and fieldwork findings, it is interesting to note that political will mechanism generates both effective and ineffective evaluation practice, depending on whether the political will is in favour to support evaluation practice or tends to ignore it. Further discussions will focus on the positive mechanism, which is *supportive political will*. Elaborations for each of the mechanisms in this table are discussed in the next subsection.

Table 7.1 Summary of Identified Positive Mechanisms

MECHANISMS	IDENTIFICATION SOURCES	
	Systematic scoping review	Fieldwork
(1)	(2)	(3)
<i>Understanding</i>	positive	positive
<i>Commitment</i>	positive	positive
<i>Inclusiveness</i>	positive	positive
<i>Top-down pressure</i>		positive
<i>Compliance</i>		positive
<i>Confidence</i>		positive
<i>Awareness</i>		positive
<i>Consensus</i>		positive
<i>Appreciation</i>		positive
<i>Supportive political will</i>		positive
<i>Synergy with experts</i>		positive

This study proposes a comprehensive CMO which explains the workings of all the identified relevant elements of context and necessary mechanisms in producing an effective evaluation practice (summarised in Figure 7.1). This study also proposes specific CMOs which explain the workings of the significant elements of context and significant mechanisms in producing a procedurally effective evaluation practice (summarised in Figure 7.2), a transactively effective evaluation practice (summarised in Table 7.2), and a substantively effective evaluation practice (as can be seen in Table 7.3).

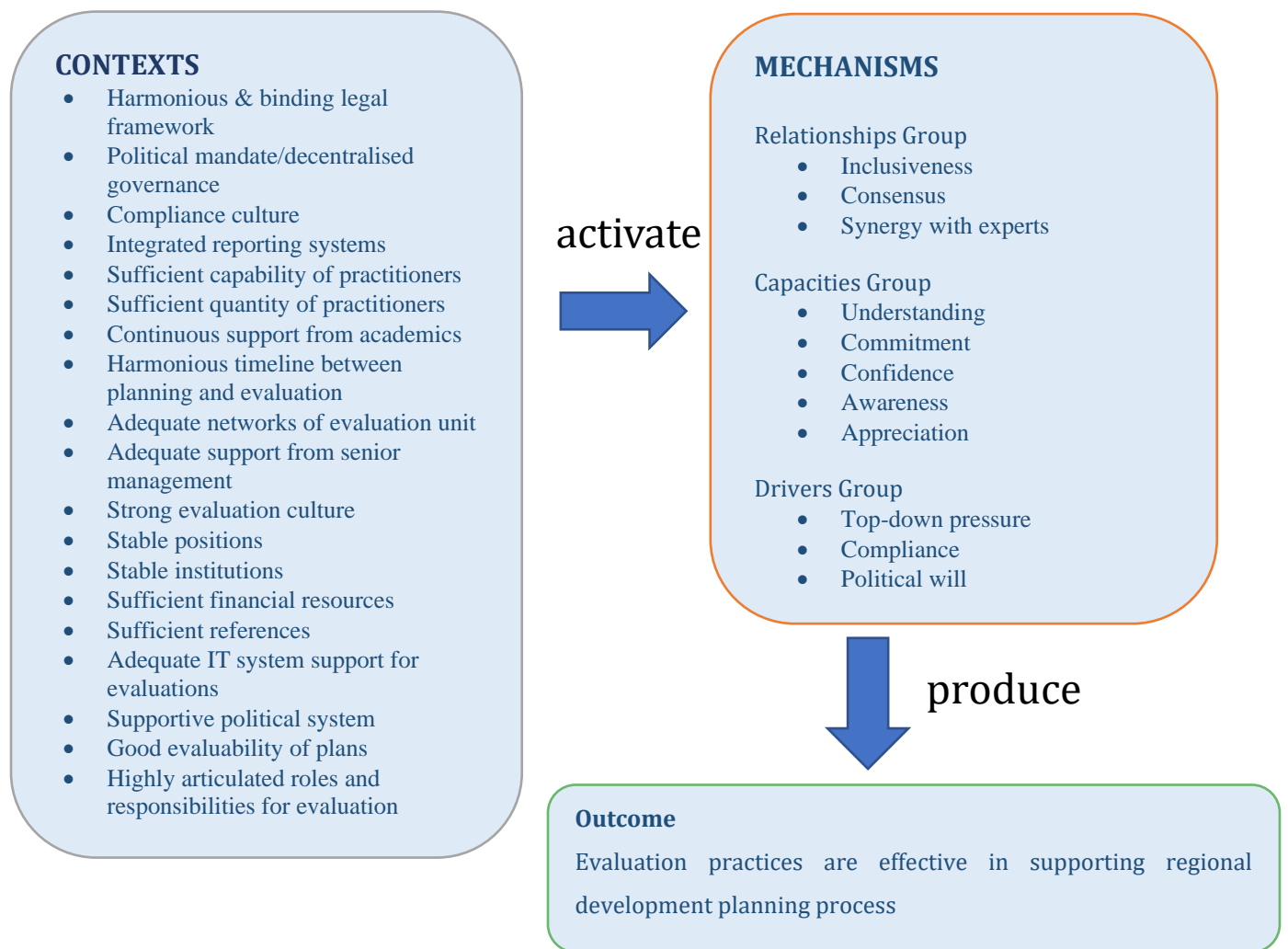


Figure 7.1 Proposed Comprehensive CMOc

Based on the similarities of their characteristics and to make discussions of the mechanisms more structured, the key mechanisms identified in this study can be classified into three groups as follows: relationships group, capacities group, and drivers group. In the relationships group, the *inclusiveness* mechanism is indicated to be the most significant mechanism. In the capacities group, *commitment* is the one mechanism which appears to have the most influence on the effectiveness of evaluation practice. In the drivers group, *political will* is the one which appears to

have more influence on the effectiveness of evaluation practice compared to *top-down pressure* and *compliance*. All of the identified mechanisms are unpacked further in Section 7.3.2.

This study sought to find causal explanation for three dimensions of evaluation effectiveness: procedural, transactive, and substantive. Empirical data on participants' perceptions provides initial insight into the key patterns that need explanation. The interview and focus group data were very useful for identifying these patterns, recognising participants' experiences and understanding how different stakeholders interpret those experiences. However, in this realist study, it is necessary to look beyond the empirical level of reality and comprehension through a number of possible explanations. Theory and prior literature are useful in the analytical process to uncover the causal mechanisms that may not be directly observable in the empirical data and may be arrived at through analysis by the researcher. The proposed CMO configuration (CMOc) is presented under the headings of the three categories of evaluation effectiveness: procedural, transactive, and substantive.

Figure 7.2 shows the proposed CMO configuration for the procedural effectiveness category. The first column shows the key elements of context which are considered as the most significant ones in facilitating the key mechanisms (in the second column) to occur. The mechanisms in the same block are perceived to work together in producing the outcome level 1 in the third column. Then, all the outcomes of level 1 are expected to work together in producing the outcome level 2 (in the fourth block of Figure 7.2).

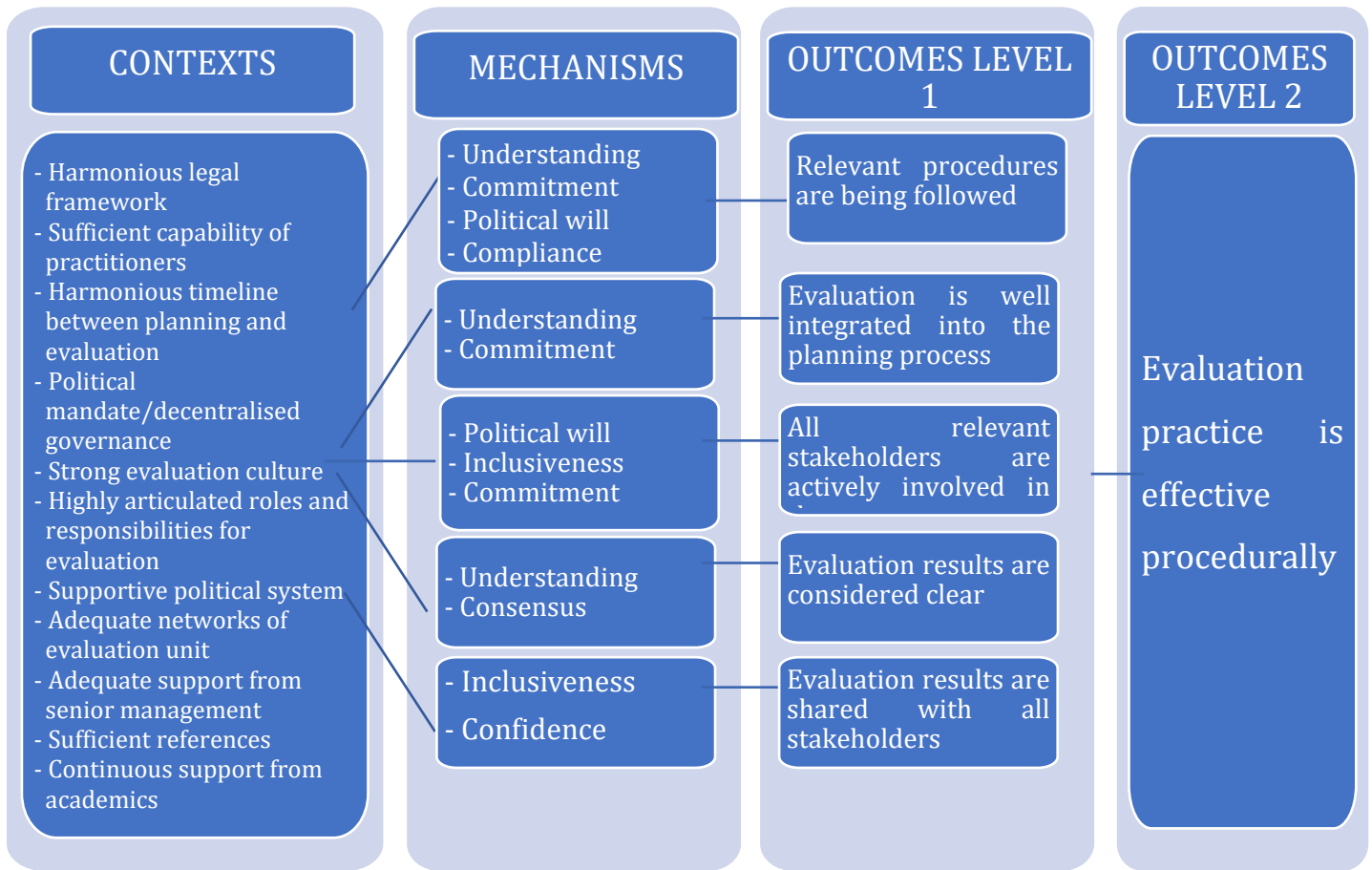


Figure 7.2 Proposed CMOc for the Procedural Effectiveness Category

As can be seen in Figure 7.2, this study has identified the most necessary elements of context for promoting procedurally effective evaluation practice. For example, this study proposes that “harmonious legal framework”, “sufficient capability of practitioners”, and “harmonious timeline between planning and evaluation” are the necessary elements of context to facilitate *understanding* together with *commitment* and *political will* to ensure all relevant procedures of evaluation practice are being followed. Furthermore, when all of the level one outcomes (relevant procedures are being followed, evaluation is well integrated into the planning process, all relevant stakeholders are actively involved in the process, evaluation results are considered clear, and evaluation results

are shared with all stakeholders) are achieved, then evaluation practice can be considered as effective procedurally.

This study argues that in order for the evaluation to be procedurally effective, seven significant mechanisms are needed: *understanding*, *commitment*, *political will*, *compliance*, *inclusiveness*, *consensus*, and *confidence*. The combination of *understanding*, *commitment*, *political will*, and *compliance* is perceived to work together in ensuring that all the relevant procedures are followed. When the stakeholders, especially the evaluation practitioners, understand the procedures (including what are the procedures, what are the benefits of following the procedures, and how to follow the procedures), then it is expected that all of the relevant procedures will be more easily followed. One of the national policy makers highlighted the importance of understanding in ensuring that the existing procedures are followed:

So, in my opinion, the most important thing is HR, because when we have tools but the human resources do not want to learn how to evaluate then it's useless, right? We already have the tools, the procedures we already have, it's just how we encourage these human resources so that they want to learn. Yes, hopefully, InsyaAllah (God willing), I think there will be no problem in the future, they can understand and use the tools.
(National policy maker)

However, it is valuable to note that it also depends on who the benefits of those procedures are for. Procedures can be clear and easy to understand but that does not mean that they will be followed if the provincial government does not perceive that there will be any significant benefits for them from following all those procedures obligated by the central government.

Moreover, empirical findings reveal that *commitment* is central and appears to be underpinning other mechanisms. All of the group of participants highlighted the importance of commitment (as will be elaborated in detail in subsection 7.4.2. Unpacking the Mechanisms):

There should be trust and commitment that we have to build from all parties including SKPD [Regional Working Unit], now including from Bappeda itself. So, those who evaluate and those who are evaluated have to build their trust properly, there is responsibility, commitment to what is done. We both understand exactly what the role of evaluation is. That's what will be good. Because all this time I have often

encountered in several areas that they did an evaluation but there was no framework for what they wanted to achieve, no understanding. (Academic participant)

According to the interviews and focus groups results, the local leader's commitment was considered as the most influential aspect supporting other necessary mechanisms to operate. In this research, commitment and willingness of political leaders to support evaluation is referred to as *political will* (will be elaborated in detail in subsection 7.4.2.). So, when there is commitment from relevant stakeholders and political will from the local leaders to push the regional government officials to comply with all the mandatory procedures, then it is expected that all of the relevant procedures will be followed:

First, the leadership of the Governor, he commits to monitor. Then how the leadership is then brought down to the bottom. From Mr. Governor to Head of Bappeda, to Regional Secretary. Commitment from everyone, from above to the bottom, so everyone is aware of their respective roles. Second, there is an agreement led by the Governor regarding the evaluation procedures, instruments, and what indicators must be used in all regions (districts / cities) to be uniform and comparable. (Evaluation practitioner)

This study also identifies that *supportive political will* as the most vital mechanism since evaluation practice takes place in a political context. Literature has discussed the strong correlation between evaluation and politics. For example, Guba and Lincoln (1989) highlighted that evaluation is a political act as well as an investigatory process, and Taylor and Balloch (2005) stated that evaluation is socially constructed and politically articulated. Additionally, Bamberger *et al.* (2011) identified that politics influence evaluation design, implementation, and also reporting. However, these authors only discuss general correlation between evaluation and politics, whereas this study specifically argues that *supportive political will* plays an important part in producing effective evaluation practice, including the procedurally effective evaluation. This is especially applicable in an Indonesian context which has a hierarchical and patriarchal working culture. Moreover, as a result of decentralisation policy in Indonesia, local leaders hold an important position in influencing the decision-making at the regional level (Aspinall and Fealy, 2003), including in deciding which evaluation procedures to follow and which to ignore. As

explained in Chapter 5, it is also important to note that in DI Yogyakarta Province, the Governor as the local leader is also a Sultan (the traditional cultural leader and monarch of the region) who has been appointed to be the governor of the province for life. Consequently, the Governor in DI Yogyakarta Province has a very special power or authority as the head of the royal family of the region, which strongly influences the decision-making in the region (Harsono, 2018).

Figure 7.2 also shows that the combination of *commitment* and *understanding* is expected to work together in ensuring that evaluation is well integrated into the planning process. This is based on the findings that when the stakeholders are committed and understand the importance of evaluation in planning, and understand how to integrate evaluation in the planning process, then it is argued by this research that evaluation will be well integrated into the planning process. Consequently, the combination of *political will*, *inclusiveness*, and *commitment* is considered to work together in making sure that all relevant stakeholders are actively involved in the process. This combination is developed based on the findings, especially the empirical findings presented in Chapter 6, that *inclusiveness* is important to include all the relevant stakeholders to participate. However, *inclusiveness* needs to be supported by *political will* that ensures *commitment* of all relevant stakeholders to be actively involved.

Additionally, the combination of *understanding* and *consensus* is perceived to work together in producing clear evaluation results. This is based on the findings that when the main actors, especially the evaluation practitioners and policy makers as the main intended users, understand and have an agreement on the expected evaluation results, then the evaluation results will be considered clear:

So, it is not enough just to have the documents, but all related parties also need to understand the value of the evaluation and have an agreement on the expected results so that the evaluation can be useful. (Academic participant)

Lastly, the combination of *inclusiveness* and *confidence* is considered to work together in ensuring that evaluation results are shared with all stakeholders. From the empirical findings, it was apparent that confidence of the evaluation practitioners of their capability in producing a high

quality of evaluation results plays an important part in motivating them to communicate the evaluation results to all stakeholders. One of the participants' perspectives regarding this matter is worth noting:

Evaluators must also be confident in their evaluation results so that they can be communicated properly to planners. (National evaluation practitioner)

Therefore, this study proposes that, besides *inclusiveness*, *confidence* among the practitioners is also needed to ensure that evaluation results are delivered to all stakeholders.

As can be seen in Table 7.2, this study has identified the most influential elements of context to trigger combinations of mechanism to operate in producing transactively effective evaluation practice. For example, this study proposes that “strong evaluation culture” and “adequate support from senior management” as the most influential elements of context to create *awareness* together with *understanding* and *commitment* in ensuring resources are managed properly. Furthermore, when all of the level one outcomes (resources are managed properly and clear specification of roles) are achieved, then transactively effective evaluation practice can be produced. Although one or more of these identified influencing factors have been discussed in the existing literature, none have identified them as a working combination or revealed the relationships between all of them. Therefore, this study provides new findings in the form of working combinations of influential elements of context and key mechanisms in producing transactively effective evaluation practice.

Table 7.2 Proposed CMOc for the Transactive Effectiveness Category

Context	Mechanism	Outcome Level 1	Outcome Level 2
Strong evaluation culture	Awareness	Resources are managed properly	Evaluation practice is transactively effective
Adequate support from senior management			
Sufficient capability of practitioners	Understanding		
Strong evaluation culture			
Adequate support from senior management	Commitment		
Strong evaluation culture			
Sufficient financial resources			
Highly articulated roles and responsibilities for evaluation			
Sufficient capability of practitioners	Understanding	Clear specification of roles	
Strong evaluation culture			
Adequate support from senior management	Commitment		
Stable positions			
Highly articulated roles and responsibilities for evaluation			

Table 7.2 shows that a combination of *awareness*, *understanding* and *commitment* is arguably necessary to work together in ensuring that resources for evaluation are managed properly. The findings from this study show that if the stakeholders of evaluation are aware of the importance of evaluation, they will be committed and then manage the resources properly:

Well, the awareness and appreciation for monev (monitoring and evaluation) needs to be improved, so if it is considered important, there will be dedicated resources for evaluation and many parties are interested in getting involved. (NGO participant)

This will also be supported by the understanding of how to manage the resources appropriately. Additionally, this study proposes that the combination of *understanding* and *commitment* helps in making sure that roles for evaluation are specified clearly. The findings of this study reveal that *understanding* and *commitment* of the relevant stakeholders, including the senior management, are needed in order for roles and responsibilities for evaluation to be specified. If the resources are

managed properly, and there is clear specification of roles, it is expected that evaluation practice will be transactively effective.

Table 7.3 Proposed CMOC for the Substantive Effectiveness Category

Context	Mechanism	Outcome Level 1	Outcome Level 2
Binding legal framework	Top-down pressure	Evaluation practice support accountability	Evaluation practice is substantively effective
Political mandate/decentralised governance			
Binding legal framework	Compliance		
Compliance culture			
Adequate IT system support for evaluations			
Political mandate/decentralised governance	Political will		
Strong evaluation culture			
Supportive political system			
Strong evaluation culture	Commitment		
Supportive political system			
Continuous support from academics	Synergy with experts		
Adequate networks of evaluation unit			
Political mandate/decentralised governance	Political will	Evaluation practice support informed decision-making	
Strong evaluation culture			
Supportive political system			
Sufficient capability of practitioners	Understanding		
Strong evaluation culture			
Good evaluability of plans			
Sufficient capability of practitioners	Appreciation		
Strong evaluation culture			
Strong evaluation culture	Commitment		
Highly articulated roles and responsibilities for evaluation			

As can be seen in Table 7.3, this study has identified the most influential elements of context to trigger the significant mechanisms to operate in producing substantively effective evaluation practice. For example, this study proposes that “harmonious legal framework” and “political

mandate/decentralised governance” are the most influential elements of context to generate *top-down pressure* to operate together with *compliance, political will, commitment, and synergy with expert* in ensuring evaluation practice support accountability. Furthermore, when all of the level one outcomes (evaluation practice support accountability and evaluation practice support informed decision-making) are achieved, then substantively effective evaluation practice can be produced.

Additionally, Table 7.3 shows that this study proposes that the combination of *political will, understanding, appreciation, and commitment* is expected to work together in ensuring that evaluation practice supports and informs decision-making. Some of the supporting statements regarding this matter can be noted from these quotes:

... the first is the leadership of the Regional Head. Why is that, because it all depends on the leadership's commitment. The good use of evaluation results depends on the leader. Including evaluation depending on the leader. If the leader really is committed to the use of this evaluation, it will be used. (Focus Group Discussion)

... the results of the evaluation should be utilised optimally. There must be a commitment from the leaders both at the central and regional levels that the evaluation must be part of the planning. Because the evaluation that has been produced must be appreciated and used as input for planning future activities. So learning from the mistakes, improvements, it must be reflected in the new planning. (Evaluation practitioner)

This study found that if there is *supportive political will* where the political leaders support processes that are necessary to utilise evaluation in informing decision-making, and combined with *understanding, appreciation, and commitment* by relevant stakeholders, then evaluation practice can optimally support informed decision-making. If evaluation practice is optimally utilised to support accountability and informed decision-making, then the evaluation practice will be substantively effective. This is the first study which identifies that the combination of *supportive political will* with *understanding, appreciation, and commitment* by relevant stakeholders is required in producing a substantially effective evaluation practice. Therefore, it can be concluded that this finding adds knowledge to the existing literature.

To understand better about the identified mechanisms and contexts, and also the relationships between them, the next subsections will elaborate further on these points.

7.3.2 Unpacking the Mechanisms

This subsection will discuss the implications and significance of the identified mechanisms in relation to the broader literature and the Indonesian context. This is an investigation within the realist explanation of the causes of evaluation practice effectiveness, and of how the workings of these mechanisms are contingent and conditional.

Based on the similarities of their characteristics and to make discussions of the mechanisms more structured, this study classifies the identified key mechanisms into three groups as follows:

- a. Relationships group: *inclusiveness, consensus, synergy with experts*
- b. Capacities group: *understanding, commitment, confidence, awareness, appreciation*
- c. Drivers group: *top-down pressure, compliance, supportive political will*

All of these three groups complement each other in producing an effective evaluation practice as can be seen in Figure 7.3. However, some mechanisms in each group appear to have more significance in influencing the effectiveness of evaluation practice than the other mechanisms. This was indicated by how frequently the mechanism was discussed in the relevant literature and by the participants in this study. This high frequency of these mechanisms is also portrayed in the proposed intervention theories which were elaborated in the previous section.

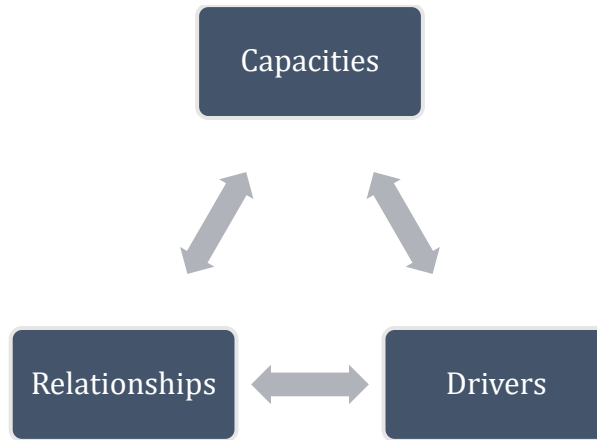


Figure 7.3 Groups of Mechanisms

7.3.2.1 Relationships Mechanisms Group

In the relationships group, *inclusiveness* is indicated to be the most significant mechanism. Moreover, based on support from relevant literature, majority of participants involved in this study have highlighted this. Generally, *synergy with experts* can also be considered as part of *inclusiveness*. However, this study separates this as another mechanism considering empirical data in DI Yogyakarta Province shows that *synergy with experts* is a specific relationship which has a significant influence on the effectiveness of evaluation practice in the regional development planning context. Therefore, this study identifies this mechanism as an individual mechanism.

Inclusiveness

In this study, *inclusiveness* refers to the consideration or accommodation or involvement of a range of stakeholders who are relevant to be involved in the evaluation process. This study has described the existing literature related to *inclusiveness* in Chapter 2, especially Patton (2008), Olejniczak (2013), and Sanderson (2000), although those existing authors do not use the exact term specifically. Most of them used the term stakeholder involvement or stakeholder participation and they discussed evaluations in general, not specifically in the regional development planning

context. One of the most authoritative researchers is Patton (2008) who was highly concerned with defined, finite potential user groups. Patton supports a breadth and depth of stakeholder involvement, with the goal being increased use of evaluations. The most recent research was a study by Deane *et al.* (2020) which discusses inclusiveness by facilitating evaluative thinking and decision-making to build stakeholders' evaluation capacity. This is known as the integrative approach to evaluation which aims to increase process use for intended users through shared decision-making, organisational learning, and capacity building while simultaneously producing a robust and relevant evaluation design suited to stakeholder needs and the evaluation context.

As found in the systematic literature review, Olejniczak (2013) and Sanderson (2000) discussed briefly inclusiveness although they did not use the exact term. Olejniczak (2013) showed concern regarding *inclusiveness* under the term of network of evaluation units. In this study, “network of evaluation unit” is considered as one of the elements of context which has a high influence in triggering *inclusiveness* as a mechanism. Olejniczak (2013) stated that the loose connection and relatively isolated position of evaluation units in relation to strategic and operational functions resulted in a limited use of evaluation results. Thus, the author suggested that evaluators must maintain *inclusiveness* with their counterparts.

Sanderson (2000) discussed inclusiveness briefly in policy evaluation as stakeholder participation where it was distinguished into two instances. The first one involved a more or less explicit value stance in favour of using evaluation processes to empower participants, to encourage the development of democratic values and to build capacity for participant self-evaluation. The second stance argued that evaluations must pose the question: ‘Whose values are to count?’ In other words, evaluations should seek to identify and elaborate the values of all stakeholders in a policy or program and identify the implications and effects for each. This stance has important implications in the context of complex social policy systems due to the potentially very wide range of stakeholders involved (especially in cross-sectoral initiatives) and the considerable scope for value conflict in relation to the underlying issues.

The *inclusiveness* mechanism in this study is strongly aligned with the second stance which was discussed by Sanderson (2000) where many participants in this study (especially the academics, policy makers, local parliament members, and NGO group of stakeholders) reported the importance of this mechanism. As explained in Chapter 6, most of the stakeholders mentioned this mechanism where they referred to including them in evaluation process. Specifically, all of the academic participants mentioned how important it is to include other stakeholders besides the ones from the local governments to improve the objectivity and the quality of the evaluation results. One of them mentioned that local parliament members, academics, and mass media are also relevant stakeholders:

Get everyone involved. The Provincial Bappeda should recap these results and then study them in a separate forum and involve the DPRD. DPRD and academics should also be involved. It's impossible like it is now that the "homework" is being done by the same one who evaluates it. Then the grade will be a hundred (rated as perfect). Yes, the evaluation must be carried out by various parties involved. And the most important thing for me is mass media (for communicating with the evaluation to the public). (Academic participant)

Similarly, NGO officials highlighted the importance to include non-government stakeholders to make evaluation practice more effective. Local parliament member also highlighted the importance for different stakeholders to be involved and complement each other with their different roles in evaluation. Most of the policy maker participants mentioned that it is important to properly involve them in the evaluation design and the evaluation process considering they are the intended users of the evaluation. Additionally, a policy maker participant also raised the need to involve all of the relevant stakeholders in evaluation process, including the research institutes and universities.

Despite the empirical findings revealing that the majority of the stakeholders acknowledge the importance of *inclusiveness*, the data from West Sumatra shows that *inclusiveness* in evaluation practice is still limited there. As discussed in Chapter 6, not all relevant stakeholders are actively involved in the evaluation process in this province. This is specifically highlighted by the group of stakeholders from academics, NGO officials, and even the policy makers. Some interviewees

in this research stated that the limited “networks of evaluation unit” and the “limited support from senior management” in facilitating active communication or networks between evaluation unit with other working units made the *inclusiveness* and *awareness* limited. In contrast, *inclusiveness* in DI Yogyakarta Province is perceived to be adequate where almost all relevant participants are considered to be actively involved in the evaluation process. Therefore, *inclusiveness* as a mechanism can operate more optimally in generating more effective evaluation practice in this province compared to West Sumatra Province.

Besides the importance of *inclusiveness*, it is also important to know about the issues related to it. Alkin, Hofstetter and Ai (1998) cited in Christie (2003) detailed three common concerns about stakeholder involvement: determination of which stakeholders’ interests should guide the evaluation, the extent to which the evaluator can adequately represent the interests of stakeholders, and advocacy. Taut and Alkin (2010) highlighted again about the importance to know whom to involve and how. In addition, Widmer and Neuenschwander (2004) briefly mentioned *inclusiveness*, stating that evaluators need to establish a good relationship with the people involved, and that this is one of the cornerstones in establishing a learning-oriented evaluation culture. They also stated that on the other hand, a close relationship between the evaluation and those in charge of the evaluandum is a clear disadvantage when aiming for trust in evaluation results among external users in an accountability-oriented evaluation. All of these literatures emphasise that *inclusiveness* is important; however, it is also important to know who, how, and when to involve people, and to what degree they should be involved.

Based on these discussions, both empirical findings and existing literature, especially Patton (2008), Olejniczak (2013), and Sanderson (2000)), support the statement in this study that *inclusiveness* mechanism influences the effectiveness of evaluation practices although this study is the first one which identifies *inclusiveness* as a specific mechanism because the other literature use the other terms, which are stakeholder involvement or networks or stakeholder participation. Additionally, this study argues that stakeholder involvement or participation is not enough.

Therefore, this study proposes the term *inclusiveness* as a mechanism to indicate that a high quality of involvement or participation is the significant mechanism.

Furthermore, this study contributes more in providing empirical evidence, especially in the domain of regional development planning, that inclusiveness is generating effective evaluation practice. Additionally, this study reveals that this *inclusiveness* has a high significance compared to other mechanisms in generating effective evaluation in regional development planning. Moreover, this study identifies the relationships or generative causations between the identified elements of context and mechanisms where the empirical findings of this research noted that this *inclusiveness* is considered to be activated by “networks of evaluation unit”, “evaluation culture”, “support from senior management”, “continuous support from academics”, “stable institutions”, and “supportive political system” elements of context where high sufficiency of all those elements of context would promote *inclusiveness*. Furthermore, as noted in the discussed literature above, this study acknowledges that there are challenges with inclusiveness that need to be taken into account, including the importance of knowing who, how, and when to involve people, and to what degree they should be involved. Additionally, the findings from case studies reveal that *inclusiveness* is also highly linked with *consensus*, since a high level of inclusiveness has been found to facilitate consensus between different stakeholders regarding evaluation practice, including the process and the expectation of the evaluation results.

Consensus

In this research, consensus refers to agreement between stakeholders, especially between the evaluators and policy makers as the main intended users, regarding evaluation practices which can include but are not limited to evaluation methods, evaluation results, and ways of communication in evaluation process. Hayton (2015) states that in order to improve evaluation practice, it is important to get acceptance by both evaluators and policy makers that evaluation is very imprecise. Hayton reported that the reason for the inexact nature of evaluation findings is that any public

sector intervention is characterised by a large number of variables, which policy makers tend to ignore, that is to say they ignore the wider context which conditions causality. Therefore, Hayton (2015) suggested that the policy makers should avoid trying to pressurise evaluators to provide definitive answers and evaluators need to acknowledge that regardless of the adopted approach, the evaluation results will always be subject to questioning if only as they are at best, a judgement. Additionally, the author also suggested that evaluators need to ensure that any evaluation is totally transparent as to the assumptions made, the methodology used, and the potential drawbacks of this approach. Rather than the spurious precision that characterises many evaluations, it is more realistic to provide a range of impact metrics and, if possible, any evaluation should attempt triangulation (using different approaches) (Hayton, 2015).

In line with Hayton (2015) position, empirical findings from this study show that consensus between stakeholders is needed. As shown in Chapter 6, most of the evaluation practitioners reported that they were hoping that all of the related stakeholders could have an agreement about which evaluation practice that they have to do and will be useful in supporting regional development planning processes, and what methods are appropriate to produce the expected and agreed evaluation results. Policy makers also reported the importance of consensus in evaluation practice. They would like to be involved in discussing the expected evaluation results so the results can be useful for them in the decision-making process. Unfortunately, this has not been happening in either case study locations. Findings from West Sumatra Province even showed that the policy makers were confused by the evaluation results because they could not obtain the needed information. The existing evaluation results did not match their expectations. One of the policy makers reported the different perspectives between evaluator and policy maker in perceiving evaluation practice where this difference leads to limited or no utilisation of evaluation results. Additionally, as described in Chapter 6, policy makers (especially the ones at the national level) shared the opinion about the importance of intensive study and discussions between related stakeholders to build consensus regarding evaluation practice considering that evaluation can be regarded as a new development function in Indonesia. It has only just begun to be legitimised and

carried out regularly under the new Planning Regulation in 2004, as described by the following interviewee:

Their methodology is also very weak especially because monitoring and evaluation are a fairly new development function area, only prominently introduced to the government apparatus since its enactment, since the issuance of Law 25 of 2004. Previously there were (evaluations) but sporadic and tended not to be implemented. Well, this is new so it's still evolving and I guess the progress is still slow. Therefore, there needs to be intensive studies and discussions between the related parties to build (evaluation practice) together. (National policy maker)

Academics from both case study locations also indicated their preference for *consensus*. One of the academics expressed his opinion that there should be an agreement about how to analyse evaluation data, how to verify, and where should the results be accessible:

There should be clarity and agreement, especially on how to process and verify it, but also where the data must be provided. (Academic participant)

Based on both empirical findings and existing literature, especially Hayton (2015), it is clear that *consensus* mechanism influences the effectiveness of evaluation practices. However, this study contributes new knowledge in the form of empirical evidence, especially in regional development planning context, that consensus contributes to generating effective evaluation practice. Furthermore, this study reveals that this mechanism has a strong relationship with the *inclusiveness* mechanism. The findings also show that this mechanism operates within circumstances where there is sufficient “evaluation culture”, “support from senior management”, “support from academics”, and “networks of evaluation unit” to build the consensus between relevant stakeholders. Vice versa, this mechanism could not operate optimally where these elements of context are not in a favourable condition. Furthermore, it can be noted as well that consensus might also play a part in building these contextual factors.

Synergy with Experts

In this study, *synergy with experts* refers to continuous partnership and collaboration with experts, including academics and private consultants, in the evaluation design and process. This mechanism can be understood through the partnership synergy theory (Lasker *et al.*, 2001), which suggests that leveraging of resources and skills of relevant stakeholders enhances intervention design, processes and realisation of outcomes. Findings from DI Yogyakarta Province show that the *synergy with experts* is significant in improving the effectiveness of evaluation practices. As stated in Chapter 6, all of the evaluation practitioner participants from DI Yogyakarta Province reported that this collaboration has helped them in the way of continuous knowledge sharing and discussions with the experts to design better evaluation practices and also to improve the evaluation process:

Furthermore, the last one is expert support, it is very helpful. Experts from universities, outside of government, they provided inputs from the outside. We have been in the system all this time. If there is an expert, there are other views. Then what is the term pentahalix? From universities, from the private sector, from the press. (Evaluation practitioner)

The academics who have partnered with the DIY Province local governments (specifically the RDPA) also show a similar perspective where they value the partnerships with the local governments and view that their collaboration has positive results, especially in improving the effectiveness of evaluation practices in DIY Province. Unfortunately, this mechanism could not be identified in West Sumatra Province because the collaborations with academics were not continuous. The interviews and focus group results in West Sumatra revealed that academics were only occasionally involved in evaluations in the province.

Empirical findings show that the *synergy with experts* mechanism has a significant influence on the effectiveness of evaluation practices. However, this study could not identify any literature regarding evaluation in planning which specifically discuss this mechanism. Other studies only briefly discuss about support from external stakeholders, such as research institutions and

individual consultants which tends to be occasional collaborations, not a continuous and synergise collaboration which is what is discussed in this study. Therefore, this study contributes in providing insight from empirical evidence that this mechanism is contributing in generating effective evaluation practice. Furthermore, this study reveals that the *synergy with expert* mechanism has a strong relationship with *inclusiveness* mechanism and can be considered as part of *inclusiveness* mechanism. However, since it only includes specific relationship with experts outside of the government, this study identifies it separately as a specific mechanism. Additionally, the present study shows that this mechanism operates within circumstances where there are sufficient “supportive political system”, “support from academics”, and “networks of evaluation unit” to build the continuous partnership with experts. Furthermore, vice versa, synergy with experts can also build these contextual factors.

Based on these discussions, it can be concluded that mechanisms under this relationships group are crucial in influencing the effectiveness of evaluation practice in supporting regional development planning processes. This is also supported by Carden (2018) who highlights the importance of relationship building for interventions in the policy arena (in this case, evaluation practice) to be effective. Following this conclusion, this study shows that techniques used to build relationship mechanisms include constant communication between relevant stakeholders, and regular partnership based on MoU (Memorandums of Understanding) between local governments and local experts (as have been practised in DI Yogyakarta Province for years).

7.3.2.2 *Capacity-related Mechanisms*

In the capacities group, *commitment* is the one mechanism which appears to have the most influence on the effectiveness of evaluation practice. This is because the findings reveal that *commitment* is underpinning the other mechanisms. This finding will be unpacked further in the following section.

Commitment

In this study, *commitment* mainly refers to the willingness of stakeholders, especially evaluation practitioners and policy makers (including local leaders), to support and take action to achieve the goals of evaluation practice. Their commitment to support processes that are necessary to the effectiveness of evaluation practice is considered very substantial. Some of the necessary support includes capacity building for evaluation practitioners, existence of an evaluation culture, and the provision of essential resources including human resources. Although this mechanism also operates among other stakeholders, the empirical findings and literature focusing on evaluation practitioners and policy makers indicated the decisive role of this group in the design and implementation of evaluation practice. However, this study also discusses *commitment* among other stakeholders, including academics, although the discussion is not as detailed as for evaluators and policy makers. It is interesting to note that according to empirical data, the *commitment* from the local leaders as the political leaders was perceived more as *supportive political will*. Therefore, the commitment from this specific stakeholder group will be explained in more detail in the *supportive political will* mechanism section.

Several studies in the literature, including the literature from the systematic scoping review results, provided evidence that *commitment* contributes to effectiveness of evaluation. For example, Seasons (2003) briefly discussed *commitment* in relation to demonstrable support from senior management and politicians. While this study noted that commitment for politicians also led to political will. Furthermore, Widmer and Neuenschwander (2004) emphasised that decision makers should be committed in acknowledging that defining the purpose as well as the intended utilisation of an evaluation is essential. They argued that only once these considerations have been addressed should the decision makers decide upon the most suitable institutional setting for the evaluation.

The findings in this study show that *commitment* is related to most of the other identified mechanisms in this study, not only with the ones in the capacities group but also the other groups, and there were strong indications that it was underpinning other mechanisms. Many of the

participants from interviews and focus groups pointed out that the commitment of the main stakeholders to evaluation practice, including the governor as the local leader, as one of the most important factors which influence the effectiveness of evaluation practice.

Figure 7.4 illustrates the relationships between commitment and the majority of other identified mechanisms in this study. As can be seen in the figure, empirical findings reveal that commitment drives other mechanisms. Across all groups of main stakeholders, *commitment* plays out relatively the same in underpinning *awareness*, *appreciation*, and *consensus*. However, some differences can be seen as well among the groups of stakeholders which relate specifically to their different roles in evaluation practice. These differences are also portrayed in Figure 7.4. For evaluators group, *commitment* appears to be underpinning *awareness*, *appreciation*, *consensus*, *compliance*, and *understanding* then *confidence*. For policy makers group, *commitment* seems to be underpinning *awareness*, *appreciation*, *consensus*, *inclusiveness* (especially for the senior management to involve all the related stakeholders), *political will*, and *top-down pressure*. For local leaders group, *commitment* is considered to be underpinning *awareness*, *appreciation*, *consensus*, *inclusiveness*, *top-down pressure*, *compliance*, and *political will* which led to facilitating *synergy with experts*. For academics group, *commitment* is perceived to be underpinning *awareness*, *appreciation*, *consensus*, and *synergy with experts*.

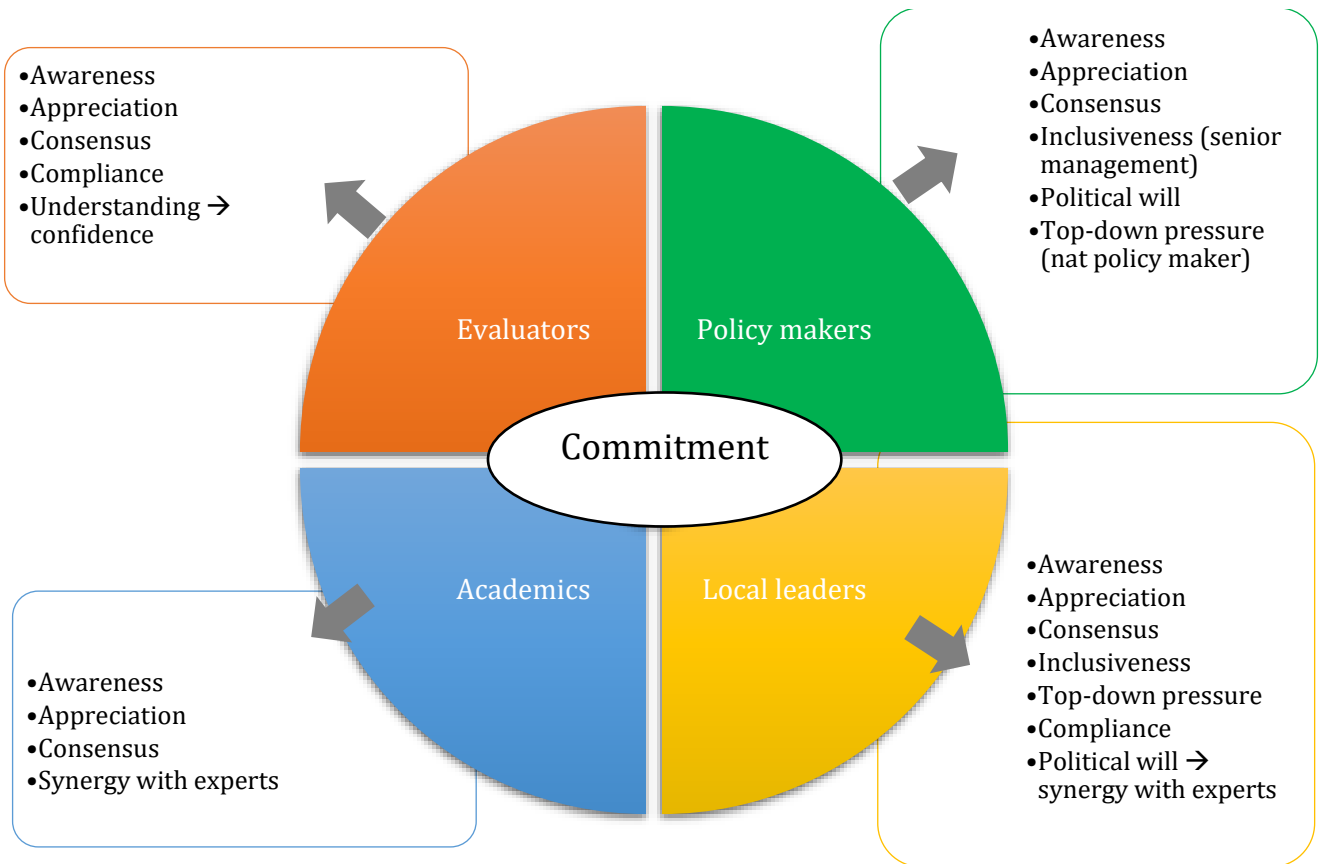


Figure 7.4 Commitment Relationships with the Majority of Other Identified Mechanisms

According to the interviews and focus groups results, the local leader’s commitment was considered as the most influential in supporting other necessary mechanisms to operate. In both case study locations, the governors as the local leaders showed their commitment and concern regarding evaluation. Figure 7.4 shows that for local leaders, *commitment* underpins other mechanisms including *awareness*, *appreciation*, *consensus*, *inclusiveness*, *top-down pressure*, *compliance*, and *political will* which appears to underpin *synergy with experts*. As explained in Chapter 6, evaluation practitioners and policy makers stated how the commitment from the local leader drives *top-down pressure*, *awareness*, and *compliance*. Additionally, evaluation

practitioners revealed the relationships between the Governor's commitment with top-down pressure, awareness, inclusiveness, and consensus.

As described in Chapter 6, evaluation practitioners also highlighted the relationships between local leader's commitment with top-down pressure, where the governor is considered to have top-down pressure to the regencies / cities in the province. Furthermore, the evaluation practitioners also mentioned the relationship between the local leader's commitment with *political will* which relate to the *synergy with experts*. In addition to *awareness*, *appreciation*, and *consensus*; *commitment* for evaluation practitioners is considered to underpin other mechanisms including *compliance*, *understanding* which underpins *confidence* to operate optimally. The evaluation practitioners explained the relationships between evaluator's *commitment* with *compliance*, *awareness*, *appreciation*, and *self-determination* which leads to *understanding*.

For academics, *commitment* is believed to lead to *consensus* and *synergy with experts* as well. For policy makers, *commitment* underpins other mechanisms including *inclusiveness* (especially for senior management), *political will*, and *top-down pressure* (especially for national policy maker), awareness, appreciation, and consensus. National policy makers' *commitment* is perceived to underpin the *top-down pressure*. Additionally, the quote from evaluation practitioner which is provided to explain Table 7.3 also highlighted how *commitment* from policy makers (including the leaders both at the central and regional levels, and also senior management level) underpins *inclusiveness*, *political will*, *awareness*, and *appreciation*. The evaluation practitioner stated that there should be commitment from the policy makers to make evaluation as an integral part of planning. This would create awareness and appreciation for evaluation which lead to *political will*, and therefore the evaluation results will be used as input for planning future programs and activities. Additionally, discussions from the focus groups in case study locations also support the statement about the importance of commitment, especially commitment from the local leader:

The first is the leadership of the Regional Head. Why is that, because it all depends on the leadership's commitment. The use of good results depends on the leader. Including

evaluation depending on the leader. If the leader really is committed to the use of this evaluation, it will be used. (Policy maker participant in West Sumatra Province FGD)

One of the academics gave input on how to ensure evaluators' commitment, one of which is the specification of roles, there should be functional evaluators who specifically work in the evaluation division. Similar with the functional planners which is already existing in Indonesia but with different field of work, specialised in evaluations:

Oh yeah, the people who do this evaluation should be in a special field doing the evaluation. If necessary, these people should be functional personnel, planners too, but their job is to do evaluations, provide input to planning. Now what is lacking is that there are functional people who work in the evaluation. At Bappenas, there are already some, although there are just a few. So there are people who are specifically dedicated and committed to think and work in the evaluation section, right, now in this area there is no one so that those who carry out this evaluation are administrative or structural people or staff who do not have good knowledge regarding the evaluation so that in the end the evaluation documents are just a handful. (Academic participant)

Similar to the suggestions from the academic, many of the evaluation practitioners shared their opinions about the necessity to build a specific evaluation position and career in the public sector institutions, especially Bappeda. This was explained in detail in Chapter 6.

Based on the empirical findings and extending the view of Seasons (2003) and Widmer and Neuenschwander (2004), this study concludes that *commitment* does influence the effectiveness of evaluation practices although this study is the first one to identify *commitment* as a specific mechanism. This study makes a contribution to literature that *commitment* acts as the underpinning mechanism and degrees of commitment in evaluation matters. Furthermore, this study contributes in providing empirical evidence, especially in the domain of regional development planning. Moreover, this study identifies the relationships or generative causations between the identified elements of context and commitment as a mechanism. As explained in Chapter 6, commitment as a mechanism triggered in circumstances where the stakeholders are well supported by government regulations such as the “harmonious legal framework” and “political mandate”; and other supporting elements of context such as “compliance culture”, “adequate support from senior

management”, “stable positions”, “stable institutions”, “sufficient financial resources”, “supportive political system”, and “highly articulated roles and responsibilities for evaluation”. Additionally, the findings also show the significant role of “evaluation culture” in enabling *commitment* to operate. Some of the participants mentioned that such a conducive working environment is more likely to get stakeholders to pay attention to evaluations and be committed to take part.

Understanding

In this research, *understanding* refers to the knowledge and comprehension to conduct evaluation practice effectively, including knowledge about the evaluation methods, communication of evaluation processes and results, and utilisation of evaluation results. Seasons (2002a) argued that it is important to understand the integration of quantitative and qualitative research, as well as triangulation of research methods as one of the facilitating factors for effective monitoring and evaluation in municipal and urban planning. Seasons also stated that a sufficient understanding of how to articulate or communicate evaluation purpose and approaches properly in publications is needed for evaluations to be effective. Similarly, Waldner (2004) suggested that planners should know how to conduct evaluation and how to place it in an appropriate context. Additionally, Olejniczak (2013) briefly suggested the importance of evaluation units to have understanding of information needed by the intended users.

Extending what has been discussed by Seasons (2002b), Waldner (2004), and Olejniczak (2013), the empirical findings from this research show that *understanding* is a significant mechanism to produce effective evaluation practice. As discussed in Chapter 6, the findings also reveal that this *understanding* has significant relationships with other mechanisms such as *confidence*, *appreciation*, *confusion* and *overwhelm*. The relationship with *confidence* and *appreciation* is a positive relationship, while level of *confusion* and *overwhelm* appear to have an inverse

relationship. As a result, when the *understanding* is high then the level of *overwhelm* and *confusion* would be low, and the vice versa, as confirmed in interviews:

The understanding and credibility of the evaluators are very important, very important. Because it affects what method will be used. It is not certain that the evaluator is willing to use a method according to the objective, but they use a method that they are comfortable using which is not necessarily suitable. That could become a problem too. (Academic participant)

Similarly, a national level policy maker participant also noted that it is important for the evaluation practitioners to understand the way to evaluate and have the motivation to keep on learning.

So, in my opinion, the most important thing is HR, because when we have tools but the human resources do not want to learn how to evaluate then it's useless, right? We already have the tools, the procedures we already have, it's just how we encourage these human resources so that they want to learn. Yes, hopefully, Insyallah (God willing), I think there will be no problem in the future, they can understand and use the tools. (National policy maker)

From the findings of this research, it can also be noted that the *understanding* mechanism is activated by “capability of practitioners” element of context but also is affected by other elements of context, such as “harmonious legal framework”, “integrated reporting systems”, “harmonious timeline between planning and evaluation”, “evaluation culture”, “sufficient references”, and “good evaluability of plans”. In both case study locations of this research, this *understanding* mechanism was found to be in low existence due to the insufficient condition of the elements of context including the conflicting legal framework, overlapping reporting systems, overlapping timeline between planning and evaluation, limited evaluation culture, and insufficient references. Therefore, the *confidence* and *appreciation* mechanisms were also inadequate. These conditions confirm that the *understanding* mechanism has a positive relationship with *confidence* and *appreciation* mechanisms.

Confidence

According to psychology literature, confidence is a belief about the validity of our own thoughts, knowledge or performance and relies on a subjective feeling (Luttrell *et al.*, 2013; Grimaldi *et al.*, 2015). In this research, *confidence* refers to the feeling of self-assurance mainly from the evaluation practitioners arising from their own abilities and or from the quality of the evaluations that they have commissioned:

Evaluators must also be confident in their evaluation results so that they can be communicated properly to planners.” (National evaluation practitioner)

“The expected competency of the evaluator will have three terms: confident, competent and credibility. Actually, they are interrelated. They can be confident when they are competent and credible. (NGO participant)

There are limited studies which discuss evaluators’ confidence. One of them is a study from Becker *et al.* (2015) which highlighted the importance of understanding the four indicators of buy-in (timeliness to evaluation team requests; quality and quantity of feedback received; interaction with decision-makers; and investment of in-kind contributions) to enable evaluators to proceed with confidence, knowing stakeholders are satisfied with the evaluation plan and progress. They also stated that it is necessary for novice evaluators who may lack the confidence to rationalise their evaluation approach and methods to their clients. Furthermore, they also mentioned that awareness of these indicators of buy-in can help to boost confidence, alleviate confusion and lay the foundation for an effective evaluation.

Based on empirical findings and extending the discoveries by Becker *et al.* (2015), this study proposes that confidence mechanism influences the effectiveness of evaluation practices although this study is the first one which identifies *confidence* as a specific mechanism. Furthermore, this study contributes more in providing empirical evidence, especially in the domain of regional development planning, that *confidence* is generating effective evaluation practice. As explained in Chapter 6, evaluation practitioners were the ones who highlighted this mechanism. Moreover, this study identifies the relationships or generative causations between the identified elements of

context and mechanisms where the findings of this research noted that this *confidence* is considered to be activated by “sufficient capability of practitioners”, “harmonious timeline between planning and evaluation”, “strong evaluation culture”, and “highly articulated roles and responsibilities for evaluation” elements of context where high sufficiency of all those elements of context would boost *confidence*. Additionally, the findings from case studies (as described in Chapter 6) have revealed that *confidence* is also highly linked with *understanding* and *awareness*, which suggests that high *understanding* and *awareness* would improve evaluators’ *confidence*. In short, this study adding to the literature by discussing in detail about the importance of *confidence* in supporting effective evaluation, and the relationships or generative causations between *confidence* as a mechanism with the relevant elements of context.

Awareness

In this study, *awareness* refers to stakeholders’ consciousness and concern about evaluation practice and its roles in supporting regional development planning processes. It was emphasised throughout the interviews that there is the need to create awareness about evaluation practice and its results as well as educating related stakeholders about the roles of evaluation. This way of thinking suggests that, if only people knew, they would be able to better utilise and be assertive in the evaluation process. Awareness raising is a critical component of the intervention theory as it can strengthen information flow, improve stakeholders’ access to evaluation results and also their appreciation to it. Therefore, *awareness* has a strong relationship with *appreciation*. One NGO participant raised the point about the need of public education to improve awareness about the importance of evaluation:

Then there also needs to be some kind of public education, awareness of all parties, especially implementers of policies or programs regarding the importance of evaluation. Because of our experience, not everyone understands that it's important. Often people feel that the evaluator is similar to the auditor, the same as the examiner,

so they might be a bit difficult to distinguish what I am checking for. (NGO participant)

One of the national policy makers argued that the awareness of the local officials about evaluation is still limited. Therefore, they are always trying to motivate and improve the awareness by constructing relevant regulations about evaluation.

If I think like this, we still have to motivate them. Then we still have to encourage them to do that. Although not only because of this, we have done it before. But indeed that condition makes them this, but yes we can't stop, we still have to. Because their awareness is also lacking like that. Sometimes they think maybe how to do it is less important, and they also think it's like a burden. (National policy maker participant)

Academics, NGO officials, and evaluator participants also had the same perspective about the limited awareness of evaluation's importance. They urged a more systematic efforts to improve the awareness and also the capacity of the relevant stakeholders, especially the evaluation practitioners.

....we hope that the awareness of this evaluation will be more widespread to all stakeholders and then there are also more systematic efforts to strengthen the evaluation activity itself, including strengthening the capacity of the evaluators. (Academic participant)

Literature also shows the importance of awareness about evaluation. Stockmann *et al.* (2020) stated that one of the factors that might hinder evaluation to be utilised regularly in the decision-making process probably is its weak position in general and the resulting low awareness of evaluation and its potential benefits. Additionally, Saunders (2012) and Hayton (2015) also briefly mentioned that awareness or concern of evaluation and implications of the evaluation are needed.

Based on these discussions, both empirical findings and existing literature support the statement in this study on the role of awareness mechanism in influencing the effectiveness of evaluation practices although this study is the first one which identify *awareness* as a specific mechanism. Furthermore, this study contributes more in providing empirical evidence, especially in the domain of regional development planning, that *awareness* is generating effective evaluation practice.

Moreover, this study identifies the relationships or generative causations between the identified elements of context and mechanisms where the findings of this research noted that this *awareness* is considered to be activated by “sufficient capability of practitioners”, “strong evaluation culture”, “adequate support from senior management”, and ”supportive political system” elements of context where high sufficiency of all those elements of context would raise *awareness*. These elements of context are identified as the key ones considering the main actors in this study are the government officials. However, the empirical findings also show that the awareness of external stakeholders such as academics and NGO official is needed as well.

Appreciation

In this study, *appreciation* refers to the value and priority given to evaluation by relevant stakeholders especially policy makers, where evaluation is considered important and beneficial for supporting regional development planning both for accountability and informed decision-making. This is consistent with Saunders (2012) who perceived appreciation as value and priority given to evaluation. He mentioned this as part of evaluative culture and organisational context which affect the use and usability of evaluation. However, he did not discuss appreciation in further details as his focus was on the different interpretations of use and usability of evaluation.

As described in Chapter 6, the NGO participant from this study mentioned the appreciation for evaluation needs to be improved because it is still limited, which is indicated by the limited resources provided for conducting evaluations. Additionally, policy maker participants stated that the appreciation for planning and evaluation should be more or less the same. Therefore, evaluation results will be considered as important inputs for planning. Moreover, an academic participant highlighted the importance of appreciation for evaluation by the local leader to ensure the utilisation of evaluation for the next period of development planning:

So in my opinion, the evaluation results are effective when the regional head considers the evaluation results important and uses them for the next development planning. (Academic)

As Chelimsky (1994) noted, the first requisite to useful evaluations is an appreciation that the evaluation is worth doing and that the findings will be useful for the intended users. The author also added that in the real world of busy service providers who are not predisposed to appreciate this work, evaluators may find it difficult to convince the users that the benefits outweigh the costs. Additionally, there is a resounding theme in the literature about the importance of practical training and hands-on experiences to develop an understanding of and appreciation for evaluation (Milstein *et al.*, 2002; Huffman *et al.*, 2008).

Based on these discussions, both empirical findings and existing literature support the statement in this study that appreciation mechanism is influencing the effectiveness of evaluation practices. Furthermore, this study contributes more in providing empirical evidence, especially in the domain of regional development planning, that *appreciation* by policy makers including the Local Leader is generating effective evaluation practice. Moreover, this study identifies the relationships or generative causations between the identified elements of context and mechanisms where the findings of this research noted that this *appreciation* is considered to be activated by “sufficient capability of practitioners” and “strong evaluation culture” elements of context where high sufficiency of all these elements of context would improve *appreciation*. Additionally, the findings from case studies reveal that *awareness* is also highly linked with *appreciation* where high *awareness* would improve *appreciation*.

7.3.2.3 Drivers Mechanisms Group

In the drivers group, *political will* is the one which appears to have more influence on the effectiveness of evaluation practice compared to *top-down pressure* and *compliance*. This is because political will is considered to lead to *top-down pressure* and *compliance*.

Supportive political will

In this research, *supportive political will* refers to the willingness and commitment of political leaders to support processes that are necessary to the effectiveness of evaluation practice, including the utilisation process. The importance of *supportive political will* was clearly mentioned by the participants in this study:

Political commitment or political will from the leadership level also influences. Sometimes we also cannot execute it without full support from the Leader. Incidentally, our Governor is quite concerned. He even brought in special people from the ministry to discuss indicators. It means that indicators are very important as a measuring tool in evaluation. (Policy maker)

Now, that means political will. Political will from the authorities, from policy makers to accept and follow up on the results of the evaluation. (Academic participant)

In addition to the views of participants in this study, *political will* has strong support in the relevant literature. *Political will* is important because the political influence can have major effects on the evaluation design, approach and methods; and politics also has the potential to influence the decisions made from the evaluation findings. Therefore, it is important to be aware of the positive and the negative aspects of *political will*. *Political will* is considered to be positive when the political leaders support evaluation, including the utilisation process in decision-making and support in allocating sufficient resources for evaluation. Political factors also have a strong effect on the stability of positions and public institutions. As noted in Chapter 6, in several cases officials were suddenly demoted due to political reasons. This was cited by respondents during FGD in West Sumatera Province, although specific details were not made available. However, the fact that this issue was supported by several respondents, indicates that such practices are common in this case study location. This portrays the significant political power of local leaders under the decentralised system in Indonesia. Stability is needed for evaluations to be effective considering it “allows the accumulation of experience on the demand side (that is institutions that contract out

studies), it increases the methodological quality of studies and, over time, the institutional capacity for strategic thinking” (Olejniczak, 2013: 1662).

On the other hand, it is also important be aware of the negative aspects of political will which lead to the ineffectiveness of evaluation practice. Guyadeen and Seasons (2018) also identified two main institutional challenges that inhibit the use of evaluation in planning agencies: organisational culture and political constraints. “The creation of plans is inherently a political process because politicians use plans to garner public support and, more importantly, elected officials are usually the decision-makers in planning” (Guyadeen and Seasons, 2018: 10). Politicians and also planners may be afraid that evaluation can reveal shortcomings or failures that reflect political judgements (Waldner, 2004; Laurian *et al.*, 2010).

Empirical findings in this study reveal that *supportive political will* mechanism influences the effectiveness of evaluation practices, whereas other studies discussed political factors in general which can include political system, political environment, political processes, and also political will. For example, Seasons (2003) discusses this point under the term of political realities which include political environment, political will, and electoral politics. He only briefly mentioned political will in one of the quotes from participant where the participant stated that “we work in a political environment. As such, the end result is not so much whether or not the policies are effective, it's whether or not the political will exists to implement those policies.” This is aligned with the perspectives of most participants in this study as explained in Chapter 6. However, Seasons (2003) did not discuss this matter further while this study discusses it further and argues that the impact of political will as contingent on these other political factors.

Furthermore, this study contributes more in providing empirical evidence, especially in the domain of regional development planning, that *political will* can be both generating effective and also ineffective evaluation practice. It can positively support effective evaluation practice when the political will is supporting evaluations, and it can hinder effective evaluation when the political will is not supporting evaluations. This is a new finding which clearly states the dual side of political will in influencing the effectiveness of evaluation practice.

This study also reveals that this *supportive political will* is a great deal more significant in comparison to other mechanisms in the drivers group in generating effective evaluation in regional development planning. Additionally, this study notes that humiliation plays a key role as a means of enforcing political will. However, this study does not recommend humiliation as an approach, as this would mean that fear of humiliation should drive relevant government officials, especially the ones that need to be accountable for their performance. Thus, humiliation may lead to a procedurally effective evaluation practice, but it may not result in a substantively effective evaluation practice. Moreover, this study identifies the relationships or generative causations between the identified elements of context and mechanisms where the findings of this research noted that this *political will* is considered to be activated by the “harmonious legal framework”, “political mandate”, “evaluation culture”, and “supportive political system”. If all these are in place, then humiliation will be no longer needed. Additionally, the presence of the laws or government regulations and other policies that obligate the regional local governments to commission evaluation practice are considered as necessary common basis.

Top-down pressure

In this study, *top-down pressure* interpreted as the exercise of control from higher levels of government to the lower levels of government. Several respondents reported that the pressure from national government to local governments is strongly felt and they considered it as one of the most influencing factors which motivated the evaluation practitioners to conduct evaluations. One of the evaluation practitioners clearly mentioned how they conduct some of the evaluation practices because of the top-down pressure from the ministries, as explained in Chapter 6. Another evaluation practitioner participant revealed the relationships between the Governor’s commitment with top-down pressure, awareness, inclusiveness, and consensus. The participant viewed top-down pressure from the Governor can enhance the commitment from all stakeholders in related institutions in the provincial government:

First, the leadership of the Governor, he commits to monitor. Then how the leadership is then brought down to the bottom. From Mr. Governor to Head of Bappeda, to Regional Secretary. Commitment from everyone, from above to the bottom, so everyone is aware of their respective roles. Second, there is an agreement led by the Governor regarding the evaluation instruments, and what indicators must be used in all regions (districts / cities) to be uniform and comparable. (Evaluation practitioner)

As explained in Chapter 6, evaluation practitioners also highlighted the top-down pressure from the Governor to the regencies / cities in the province considering Governor's role is as an extension of the central government and serves to facilitate the cities/municipalities. As described in Chapter 6, *top-down pressure* from the central government, especially from the MOHA for different kinds of reporting, also has a significant influence. However, it is important to note that *top-down pressure* tends to lead the regional governments to conduct the evaluation practices as mainly an empty exercise in procedural compliance. Therefore, it can be concluded that *top-down pressure* is not always positive in terms of producing effective evaluation practice.

It is argued in the literature that *top-down pressure* is a prevalent aspect of the paternalistic attitude of political leaders as well as the tradition of centralised governance, in spite of legal mandates of decentralisation. Additionally, this top-down pressure is regarded to be significant in case study locations of this study considering the existence of paternalistic leadership in Indonesia, especially in Java. The existence of paternalistic leadership was also pointed out by participants in this study, both in West Sumatra Province and more apparently in DI Yogyakarta Province. This is aligned with literature where Irawanto *et al.* (2012) found that traditional Javanese values encourage similar paternalistic leadership which is used in Chinese-based societies. This paternalistic leadership has been claimed to be one dominant leadership style in Asia (Cheng *et al.*, 2013). Moreover, Huang and Kim (2020) found that when the policy entails substantial class or distributive conflicts and bureaucratic friction, top-down pressure for compliance is a dominant driver for local policy adoption. Therefore, both empirical findings and literature showed that this *top-down pressure* mechanism is highly related to the *compliance* mechanism. Furthermore, this study provides specific insights from the Asian region, where *top-down pressure* and *compliance* are significant when there is existence of paternalistic leadership in the public sector. Furthermore,

this study identifies that this *top-down pressure* operates within a context where there is “harmonious legal framework” and “political mandate/decentralised governance”. However, presumably *top-down pressure* can work without these elements of context, but its impact may be less optimal.

Compliance

In this study, *compliance* is interpreted as the obedience of public sector officials, especially evaluation practitioners, to follow the existing laws or government regulations and other policies that obligate them to conduct evaluation practice. Therefore, this mechanism has a strong relationship with *top-down pressure*. Many participants in this study reported views about this mechanism:

The problem is also that the evaluations are a very political burden, it's just lip service. In actual fact, they just want to meet compliance measures, they don't care about the content and don't care about the quality so long as there's a report. (Academic participant)

After [discussions] in the focus group and the like, it turned out that the main problem with the evaluation was that it was done purely for reporting purposes. So we write a lengthy evaluation report and then they just file it, what I mean is that there's no follow-up. The year after, it was much the same; there were also evaluation reports that more or less had the same content because the evaluation results were never followed up on. (Academic participant)

According to the empirical findings and literature, it can be concluded that compliance is influential but it should not be merely a procedural exercise, to complete the obligation, or to complete the decision-making process only:

From what I said earlier, I think the most important thing is policy. The policy emphasizes that this evaluation is not just an obligation but is a necessity. The policy that makes it possible at the regional government level to comply is to carry out the evaluation. At the same time, the policies that make the evaluation results are followed up. It also needs that policy. So the policy is not only a policy that requires evaluation, but a policy that requires that evaluation results be

followed up. That's the most priority. The Ministry of Home Affairs has said that all regions will participate. (Academic participant)

Based on the above discussions, empirical findings in this study reveal that *compliance* mechanism influences the effectiveness of evaluation practices, whereas other studies only discussed it briefly considering the topics they focused on were different from those in this study. For example, McDonald *et al.* (2003) focused on the evaluation capability building in the public sector where they stated that making evaluation mandatory could promote a culture of token compliance, but voluntary adoption is much slower to take effect. Therefore, whether evaluation should be mandatory is one of the important decisions to be made in managing evaluation demand. Furthermore, Davies (1999) stated that evaluation cannot be forced on people because attempting to impose it will likely lead to goal displacement, unreliable information and an increase in the risk that program relevance will be diminished rather than augmented.

Extending this idea, McDonald *et al.* (2003) and Davies (1999) added that, to avoid a culture of token compliance, this mechanism should work together with other mechanisms, such as *understanding* and *inclusiveness* in order for relevant stakeholders to cooperate actively with evaluation practitioners so that they can do the best evaluation possible for both accountability and informed decision-making purposes. Furthermore, this study contributes more in providing empirical evidence, especially in the domain of regional development planning, that “harmonious legal framework”, “political mandate”, “compliance culture”, “evaluation culture”, “sufficient financial resources”, “adequate IT system”, and “supportive political system” as essential elements of context that enable *compliance*. In other words, compliance is dependent on these elements of context.

This section has discussed which key mechanisms generate effective evaluation practice and under which circumstances (elements of context) based on the findings from the case studies and the existing literature. This study also found that some mechanisms have not been properly discussed in the evaluation and public policy literature, such as *confidence* of evaluators. Therefore, this study proposes descriptions for each of the identified key mechanisms specifically in the field of

regional development planning to contribute to the literature. The evidence from West Sumatra Province and DI Yogyakarta Province also revealed a new finding that some mechanisms are more significant than the other mechanisms. Furthermore, this study classifies the positive mechanisms into groups to help in identifying the most significant mechanisms. In the relationships group, *inclusiveness* mechanism is indicated to be the most significant mechanism. In the capacities group, *commitment* is the one which appears to have more influence on the effectiveness of evaluation practice. This is because the findings reveal that *commitment* is underpinning the other mechanisms. In the drivers group, *political will* is the one which appears to have more influence on the effectiveness of evaluation practice compared to *top-down pressure* and *compliance*. After getting a detailed explanation of each of the identified mechanisms in this subsection, the next subsection will discuss the importance of context.

7.3.3 The Importance of Context

The findings of this research show that evaluation is a context-specific practice. This study identifies factors of evaluation practice effectiveness using a realist approach which detected the links of context to effectiveness by exposing the underlying mechanisms. Therefore, this study contributes to the better understanding of the role of context, which triggers a particular set of mechanisms that generate particular evaluation effectiveness. Additionally, this study identifies the necessary elements of context to allow the mechanisms to operate in generating effective evaluation practice.

As described in Chapter 6, Table 6.8 maps the relationships between each of the identified elements of context with the identified mechanisms. This table reveals “strong evaluation culture” as the most significant element of context which was considered to have a significant influence in triggering most of the identified mechanisms to operate in generating effective evaluation practice. This is consistent with Seasons (2003) who asserts that organisational culture is one of the most important influencing factors for effective evaluation. As explained in Chapter 6, participants from

the majority of stakeholder groups including evaluation practitioners, policy makers, academics, NGO officials, and local leaders referred to the importance of evaluation culture as part of the context:

And all these mechanisms can work if they are supported by the organisational culture [which support evaluation] and the support from senior management. (Academic participant)

Additionally, the individual perceptions from the interviews about evaluation culture were also confirmed in the focus group discussions in both case study locations. This is one of the confirmations:

Agree, evaluation culture first. If the quality of human resources is good but the organisational culture is not accustomed to evaluation then it still won't work, right? (Academic participant)

According to the literature, evaluative or evaluation culture is defined as the commitment to conduct and use information from evaluation in the decision-making process in an organisation, therefore evaluation becomes an integral part of the organisation culture (Owen, 2003; Cousins and Bourgeois, 2014). Dahler-Larsen and Boodhoo (2019) described it in two ways, first, evaluative mindsets and how people make evaluation fit into local beliefs and traditions; and second, the institutionalisation of evaluation. Mayne (2008) pointed out that developing evaluative culture is vital in order to make result-based management effective. A strong existence of evaluation culture is expected to foster the utilisation of evaluation results in the policymaking process. This is mostly because evaluation culture has the role to ensure there will be a continual process in using evaluation results in policymaking, not only at a particular period of time.

There are already a number of studies which describe how evaluation culture can be developed in public institutions. For example, Bustelo (2006) described the potential role of standards and guidelines in developing national evaluation culture in Spain, while McNamara *et al.* (2009) explained the emergence of evaluation culture in the public education sector in Ireland, and Ferguson (2003) examined the factors that have supported the development of evaluation culture

in the Australian War Memorial. Furthermore, Owen (2003) identified the key factors in developing evaluation culture at the Emergency Management Australia, while Mihalache (2010) described evaluation culture as part of their national evaluation system in Romania, and Mora and Antonie (2012) explained the role of leadership in supporting evaluation culture in Romanian public administration. From these studies, it can be noted that development of evaluation culture involves a range of activities and requires a relatively long time. Bustelo (2006) recommended the following: institutionalising and consolidating evaluation practice; creating a common language and terminology; defining the key criteria, issues, and elements of evaluation; creating a specific identity for evaluation, evaluators, and the evaluation units; mapping out evaluation references and options, and supporting training of evaluators. Moreover, Owen (2003) identified key factors that contributed to the development of an evaluation culture, including knowledge of where to get expert external evaluation assistance; a joint commitment from operational managers for the use of internal evaluation to help decision-making; high level support from the executive manager; and employment of external evaluators who use participatory/interactive forms of evaluation. The author also advocated for communication to all staff about how evaluation would or would not be used and how it might affect them; the use of teams and committees to develop evaluation protocols and formally receive findings; and a requirement that staff actually use findings to improve programs. Moreover, Owen (2003) recommended the development of in-house capacity to undertake aspects of evaluation as a routine part of work; the identification of obstacles to collecting systematic data; an initial focus on processes rather than outcomes, so that staff could change and improve their own programs; and over time, changing the focus of evaluation efforts to outcomes, not just processes.

The information and strategies from those studies can be adapted to improve evaluation culture in Indonesia. However, it is important to be aware of the existing challenges in Indonesia, including the local beliefs and traditions, to support the institutionalisation of evaluation. A more detailed recommendation on how to develop evaluation culture in Indonesia is provided in Chapter 8.

7.4 Conclusion

In Chapter 7, this study has presented a working definition of effective evaluation practice. It has also described the proposed Context-Mechanism-Outcome configurations which explain the workings of all the identified key elements of context and influential mechanisms in producing an effective evaluation practice under the headings of the three categories of evaluation effectiveness: procedural, transactive, and substantive. Additionally, this chapter has highlighted how this study identifies the most significant contexts and mechanisms, and the relationships between them.

This study extends the existing literature by defining all the positive mechanisms that have significant influence in producing effective evaluation practice and portraying the functioning of these mechanisms. Furthermore, the classification of mechanisms (relationships, capacities, and drivers) is useful to identify the most significant mechanisms. In the relationships group, *inclusiveness* mechanism is indicated to be the most significant mechanism. In the capacities group, *commitment* is the one which appears to have more influence on the effectiveness of evaluation practice. This study also makes another contribution to literature that *commitment* acts as the underpinning mechanism and that degrees of commitment in evaluation matter. In the drivers group, *supportive political will* is the one mechanism which appears to have the most influence on the effectiveness of evaluation practice. These findings contribute to the better understanding of key mechanisms which influence the effectiveness of evaluation practice. This is essential in informing related stakeholders on what they should do to improve the effectiveness of evaluation practice. Additionally, this study provides specific lessons for the Asian region, where *top-down pressure* and *compliance* are significant when there are paternalistic leadership values in the public sector.

This study also contributes to the better understanding of the role of context, which triggers a particular set of mechanisms that generate particular evaluation effectiveness. Additionally, this study reveals “strong evaluation culture” as the most significant element of context which was considered to have a significant influence in triggering most of the identified mechanisms to

operate in generating effective evaluation practice. In summary, this study argues that a strong evaluation culture provides conditions necessary for generating effective evaluation culture. It is important to note that this is the first study which defines effective evaluation practice, and also identifies the most significant contexts and mechanisms in the relationships among them.

CHAPTER 8 CONCLUSION

This thesis commenced against the backdrop of very limited knowledge which suggests that evaluation remains largely underutilised yet an essential part of regional development planning worldwide, particularly in developing countries such as Indonesia. An examination of theories concerned with the effectiveness of evaluation led to the identification of the gap in the literature on the topic, and revealed how this research can potentially add to the knowledge by filling this gap. Furthermore, it was explained how this research adapted and customised the effectiveness framework and the realist methodology to address the research questions.

Chapter 6 presented the findings, drawing on the field level data and conceptual frameworks of contexts and mechanisms that influence effective evaluation practice. The conceptual frameworks were discussed further in Chapter 7, where a refined framework was developed to present a number of contextual variables and mechanisms that support effective evaluation practice. The aim of this final chapter is to distil the key messages from the research and present them as a synthesis of the findings. The chapter consists of four major sections. Section 8.1 sums up the key findings of this research. Section 8.2 describes the implications for policy and practice. Section 8.3 identifies the limitations of the present study, while Section 8.4 outlines the directions for further research in this field.

8.1 Key Findings of the Research

This research has investigated and analysed the current evaluation practices in two of Indonesia's provinces. The provincial governments in the case study locations have been relatively successful in conducting evaluations for accountability purposes; however, evaluations remain largely aspirational since the value of these evaluations has not yet been optimised in the context of

supporting evidence-informed policymaking. Components of evaluation practice and culture have been introduced, but the provincial government's capacity to institutionalise evaluations has been limited, due to the overwhelming requirements for compliance in a complex national and sub-national bureaucratic environment.

The findings of this research were analysed by adapting the effectiveness framework to gain a better understanding of an effective evaluation practice, and using the realist framework to identify and uncover the underlying mechanisms and contextual variables. The definition, effectiveness framework, and intervention theory formulated in this study are the first to be developed in the literature. Therefore, this study offers a starting point to stimulate further debate on effectiveness of evaluation practice and invite others to test and develop the practices further.

According to the effectiveness framework developed in this research, an effective evaluation practice in regional development planning is one that meets all the criteria in procedural, substantive, transactive, and normative effectiveness dimensions. In other words, evaluation practice can be defined and categorised as effective when it has implemented the principles governing the evaluation process, achieved the set objectives, delivered the outcomes at minimum resources, and achieved the normative goals of evaluation practice. These working definitions are proposed to answer the first research question of this study.

An evaluation practice can be defined as effective *procedurally* when it has implemented all the principles governing the evaluation process, including that it has followed relevant procedures, integrated evaluation into the planning process, involved stakeholders in the process, delivered results to relevant stakeholders, and produced evaluation results that have been objective, clear, and understandable. An evaluation practice should be deemed effective *substantively* if it achieves the set objectives, including support of accountability and support of informed decision-making.

An evaluation practice can be considered effective *transactively* when it delivers the outcomes at minimum resource, including proper management of resources, and clear specification of roles.

Lastly, an evaluation practice is deemed effective *normatively* when it improves the plan or policy and develops or implements beneficial changes in relevant institutions.

The findings of this study suggest that evaluation practices support the regional development planning processes through their two primary roles: the accountability role and the learning or improvement role. This study is significant for it provides empirical evidence from an Asian country on the role of evaluation. Considering that virtually all existing literature found on this topic were from other regions, especially from North America and Europe, this study provides an overview on how the role of evaluation is perceived by key stakeholders in regional development planning field in an Asian country. This study also adds to knowledge by providing a more detailed explanation of the roles of evaluation in the field of regional development planning which are reporting public fund allocations, measuring and reporting progress or achievements, and assessing performance of local government as part of the accountability role.

Although evaluation practices in both case study locations were not effective procedurally, transactively, and substantively, the practices did manage to be effective as a means for accountability. This finding is consistent with the primary objective of evaluation in regional development planning in Indonesia, which is to support accountability. There are a number of reasons why evaluation practices in these two case study locations were found to be more effective for the accountability role when compared to learning roles. First, in the provincial evaluation practices, the pressure for accountability role is higher, which is confirmed by the existing regulations for this accountability purpose being more extensive and binding. Second, the accountability role has a close relationship with performance measurement of local governments which affects the remuneration management of government officials. Third, achieving accountability is easier and often only involves ticking some boxes, whereas more effort is needed to achieve learning and improvement, including revising the planning documents and having clear plans about how to use evaluation results.

This study argues that effectiveness is a multi-dimensional construct lying along a spectrum, and that it can be measured in several different ways. Therefore, evaluation practice can be lacking in some aspects yet still be beneficial; it can be partially effective despite not being fully effective. In both case study locations, it was reported that evaluation practices support the regional

development planning process mainly for reporting and accountability purposes, but they do not contribute much to learning and improvement. As explained in Chapter 6, the findings in this study show that existing evaluation practices in the two case study provinces are in part effective substantively due to their support for accountability. However, this support for accountability does not necessarily lead to other desirable outcomes in a substantive effectiveness framework, which should support the informed decision-making in planning and ultimately achieving the objectives of learning and improvement. In both case study locations, the effectiveness was hindered by the contextual variables including the poor evaluation culture and inadequate capability of practitioners to produce high-quality evaluation. There were also limitations in the key mechanisms related to this outcome, including inadequate appreciation of the importance of evaluation by the policy makers, especially the local leader or the governor, and the lack of political will to support utilisation of evaluation results in the decision-making process.

The effectiveness framework that classifies effectiveness of evaluation practice into substantive, procedural, transactive, and normative principles emphasises the significance of the multi-dimensional nature of evaluation effectiveness. This study has shown that actors in evaluation should represent diverse groups, which is an essential strategy for a more comprehensive understanding of evaluation practice. The emphasis on diversity of perspectives in this study increases the range of audience, which includes evaluation practitioners and other stakeholders, including policy makers (including planners), academics, local parliament members, community leaders, and members of society.

The effectiveness framework is also useful in portraying how different groups of stakeholders assessed the effectiveness of evaluation practice differently. From the effectiveness point of view, evaluation practices in DI Yogyakarta Province were perceived by the participants to be better than the evaluation practices in West Sumatra Province. Almost all groups of stakeholders in West Sumatra Province perceived that the evaluation practices in this province failed to meet most of the criteria in the effectiveness framework. However, there are clear differences between the state and the non-state stakeholders' views. The evaluators and policy makers, the state stakeholders,

seemed to view the state of the evaluation practices more favourably than the non-state stakeholders. Most of the evaluators reported that they have been trying their utmost to conduct evaluation practices within the available resources. However, they also reported that the resources were very limited. Consequently, the outcomes in the West Sumatra Province were not as they had hoped for. In contrast, in DIY Province, all groups of stakeholders perceived that the evaluation practices in this province partly succeeded in meeting most of the criteria in the effectiveness framework. Therefore, the research revealed a clear difference in the perceived effectiveness of evaluation practices between the two case study locations, with DIY Province being regarded as superior.

In aiming to design an effective evaluation practice, the underlying mechanisms should be identified and understood, while the specific elements of context that are conducive for a particular mechanism to operate should be defined and established. This research has identified the underlying mechanisms and also provided insights into the most significant mechanisms which influence the effectiveness of evaluation practice. Based on the similarities of their characteristics, the identified key mechanisms in this thesis can be classified into three groups, as follows: those entailing relationships, those about capacity, and those involving drivers. These three groups complement each other in producing an effective evaluation practice. However, some mechanisms in each group appear to have a more significant role in influencing the effectiveness of evaluation practice than the other mechanisms. In the relationships group, the *inclusiveness* mechanism is the most significant mechanism. In the capacities group, *commitment* is the mechanism which appears to have more influence on the effectiveness of evaluation practice. Additionally, commitment is considered to be the underpinning mechanism of the other key mechanisms. In the drivers group, *political will* is the one which appears to have more influence on the effectiveness of evaluation practice.

As for the context, this research has identified several significant elements of context which were perceived to trigger different mechanisms. Additionally, some elements of context appear to have greater influence when compared to the others. This study reveals “evaluation culture” as the most

significant element of context which was considered to have a major influence in triggering most of the identified mechanisms to operate in generating effective evaluation practice. Besides strong evaluation culture, a supportive political system, harmonious legal framework, and sufficient capability of practitioners were also perceived to have more significant influences in triggering the important mechanisms when compared to the other elements of the context.

In summary, the effectiveness framework developed in this research is useful in framing and determining the state of evaluation practices in case study locations while the Context-Mechanism-Outcome Configuration is useful in unpacking the elements of context and mechanisms which influence the effectiveness of evaluation practice. Therefore, the effectiveness framework and the CMOc complement each other.

In a practical sense, this research has shown that understanding the interactions of the elements of context, mechanisms, and outcomes can provide a direction for evaluation practitioners and government regulators who are usually responsible for the design or formulation of an evaluation practice. It clarifies under which circumstances a particular evaluation practice could be effective. This study contributes to fill a gap in evaluation in planning studies by unveiling the processes, elements of context, and mechanisms through which evaluation practice works.

This study has applied the realist-informed research method to study the evaluation practice, which has rarely been done, and has employed intervention theory to identify plausible causal configurations of context-mechanism-outcome. This intervention theory improves the generalisability of this study's findings, especially in other similar contexts. Additionally, this study has theorised how evaluations are expected to work in practice and demonstrated the value of theory-based approaches, such as realist evaluation. The study has shown that both theory development and realist evaluation are valuable for advancing knowledge about evaluation practices, despite there being many challenges. Evaluation practices have two significant roles in supporting regional development planning processes and therefore it is important to continue to refine their design through an understanding of how they operate, for whom they provide benefits,

and the circumstances under which the desirable outcomes can be achieved. It is hoped that the intervention theory and effectiveness framework developed in this study will be tested and refined in future studies to contribute to better evaluation practice design, process, and outcomes.

8.2 Implications for Policy and Practice

While evaluation practices have the potential to support regional development planning processes by ensuring accountability and supporting informed policymaking, significant changes in relevant elements of context which trigger the influential mechanisms are required to optimise the role of the evaluation practices in supporting regional development planning at the provincial level in Indonesia. This research offers important lessons and implications for policy and practice, especially for the major stakeholders including the evaluation practitioners and policymakers, to improve the effectiveness of evaluation in supporting regional development planning at the provincial level.

The findings of this research demonstrate that it is important for the evaluation practitioners at the regional level to tailor the evaluation practices to local contexts based on the detailed understanding of CMOc. This approach can enable practitioners to enhance their evaluation practices. For policy makers at the regional level, the findings point to the need to allocate the resources more effectively considering they should know which element of context has the highest significance, and which mechanism makes the greatest contribution. For policymakers at the national level who regulate the evaluation practices at the regional level, it is advised that existing regulations should be revised based on the understanding that only some evaluation practices make a significant contribution to the planning process due to differences in local contexts. Therefore, current one-size-fits-all approaches in national regulation should be modified to account for different contexts at the provincial level. Additionally, there should be an umbrella policy which leads to coordination among responsible ministries in Indonesia. This is based on the empirical

data which has shown that there are conflicting regulations, timelines, and overlapping reporting systems that negatively impact evaluation practices in the case study locations of this study.

Considering the significant influence of commitment, including commitment of local leaders which lead to political will, it is advised that interventions should therefore focus on the improvement and optimisation of existing techniques such as routine meetings led by the governor, routine capacity building on evaluations, and also specification of roles for evaluation. Additionally, to ensure commitment of evaluation practitioners, this study proposes that the national government, specifically the Ministry of National Development Planning/BAPPENAS, should develop a dedicated career path for evaluation practitioners in public institutions, as it already does for functional planners. This would also be useful to address other existing challenges, such as limited capability and quantity of evaluation practitioners and low stability of positions, which creates insecurity for career professionals in this important field and may be a disincentive for others to study this discipline at university.

The fact that the present study has shown the importance of evaluation culture, developing and nurturing an evaluative culture is vital in order to optimise the role of evaluations in supporting regional development planning processes. Both evaluation practitioners and policymakers need to work together in improving evaluative mindsets and how all related stakeholders make evaluation fit into local beliefs and traditions to support the institutionalisation of evaluation. A strengthened evaluation culture will foster the utilisation of evaluation results in the policymaking process. This is because evaluation culture has the role of improving awareness and appreciation for evaluation, and also to ensuring increased application of evaluation results in policymaking.

8.3 Limitations of the Research

There were several limitations in this research, the recognition of which leads to suggestions on important areas for future research. One of the main challenges in this research was the difficulty

of properly conceptualising the identified mechanisms based on the realist perspective, which is very specific and more in-depth compared to the use of the term “mechanism” in other perspectives (as explained in Chapter 2, Section 2.6.1.2). The realist approach is useful in this type of study but requires substantial understanding of its principles and processes. Difficulties defining mechanisms and contexts, and distinguishing between them, hinders development of CMO configurations. This issue was raised by other studies as well, such as Lacouture *et al.* (2015), Marchal *et al.* (2012), Punton *et al.* (2016), and Salter and Kothari (2014).

The next limitation in this study is that it is not a mixed method research, rather it is only qualitative research. Considering the research question of this study, qualitative data collection was considered to be the best approach in order to gain the detailed information on context, mechanisms, and outcomes. The qualitative methods were essential in identifying the elements of context and mechanisms explained by the participants in the study during interviews and focus groups. This is an exploratory study of intervention theories from literature and stakeholders, explored and expanded through qualitative methods of data collection and analysis. However, mixed methods might improve the generalisability of the research findings.

It is also interesting to note that the systematic scoping review was helpful, but made the literature review relatively narrow, since it focused on the literature on evaluation and policy studies. However, this study found that to fully understand evaluation practice in the field of regional development planning, it is necessary to look outside the discourse on evaluation and policy literature, to include literature from psychology and sociology. Therefore, a broader approach in further research that examines other aspects of evaluation in relation to the social science disciplines of cultural anthropology, political science, and economics might reveal useful insights complement the findings of this research in the evaluation and policy area.

Additionally, the study findings might be affected by recall bias and temporality, commonly encountered in retrospective realist research, and in any retrospective study in general. Moreover, this study was undertaken in two case study locations. It is strongly recommended that further

research be undertaken at other case study locations with different ranges of context. Lastly, this research developed the effectiveness framework which includes four dimensions: procedural, transactive, substantive, and normative effectiveness. However, this study only focuses on investigating procedural, transactive and substantive effectiveness. It did not include the normative effectiveness dimension as part of the analysis because more extensive data would have been needed in analysing the normative effectiveness dimension while the time limitation for this study was considered insufficient to cover analysis of all four dimensions.

8.4 Directions for Future Research

While this research presents useful findings in the context of minimal existing evidence about the effectiveness of evaluation practice in supporting regional development planning, it strongly recommends that further research be undertaken in other case study locations. This is essential to test the findings of this study, and collect further evidence regarding the effectiveness of evaluation practices. Additionally, this research was undertaken in Indonesia as one of the developing countries in Asia. It will be interesting to test and compare the findings further in other regional contexts. Furthermore, this study proposes a working definition of an effective evaluation practice in a regional development planning context which could be developed further in other contexts.

The intervention theory of effective evaluation practice in regional development planning developed in this research can be adapted and refined further to identify other important elements of context to activate influential mechanisms in producing effective evaluation practice. The proposed intervention theory in this research is perceived as an ideal model; therefore, further research can test, refine, and also conduct further exploration. Future research could also explore whether some mechanisms underpin others or lead to other mechanisms or are alternatives to other mechanisms, or even whether there is a temporal relationship between mechanisms (for example, mechanism A needs to be present first, then B, then C and D). Additionally, future research could

investigate whether all of the identified mechanisms need to be in place to generate the expected outcome (effective evaluation practice) or whether the existence of the most significant mechanisms would be enough to produce an effective evaluation practice. In summary, further research could test whether all the identified mechanisms in this research were necessary, or if some combinations would have worked while others would not. The findings of this research provide the basis for future research in exploring these possibilities further. Additionally, this research did not include normative effectiveness dimension as part of its analysis. Therefore, further research could focus on investigating this dimension as well.

As mentioned in the previous section, this research was based entirely on qualitative data. Therefore, future studies may consider adopting a quantitative approach using surveys and statistical factor analysis of mechanisms and contexts, as proposed in this study, to identify the most important elements of context and mechanisms. Finally, in order to fully understand evaluation practice in the field of regional development planning, research should seek answers outside of evaluation and policy literature. References from other related disciplines, including psychology and sociology, are needed to improve the effectiveness of evaluation practices.

REFERENCES

- Alexander, E. R. (2016) *Evaluation in Planning: Evolution and Prospects*, Taylor and Francis, New York.
- Alkin, M. C. (2013) *Evaluation roots: A wider perspective of theorists' views and influences*, Sage Publications, Thousand Oaks.
- Alkin, M. C., Daillak, R. and White, P. (1979) *Using evaluations : does evaluation make a difference?* , Sage Publications, Beverly Hills, Calif.
- Alreck, P. L. and Settle, R. B. (2004) *The survey research handbook* (3rd ed. ed.), McGraw-Hill/Irwin, Boston.
- ANU College of Asia and the Pacific (2020), *West Sumatra Province*, CartoGIS Services, <https://asiapacific.anu.edu.au/maponline/base-maps/west-sumatra-province>
- Arnaldo Pellini, B. P., Kharisma Priyo Nugroho, Elisabeth Jackson, Fred Carden (2018) *Knowledge, Politics and Policymaking in Indonesia*, Springer, Singapore.
- Aspinall, E. and Fealy, G. (2003) *Local Power & Politics in Indonesia : Decentralisation & Democratisation*, Institute of Southeast Asian Studies (ISEAS),
- Astbury, B. and Leeuw, F. L. (2010), 'Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation', *American Journal of Evaluation*, 31(3), 363-381.
- Ayres, L. (2008), *Semi-structured interview*, in L. M. Given (ed.), *The SAGE Encyclopedia of Qualitative Research Methods*, SAGE Publications, Inc., Thousand Oaks, California. <http://sk.sagepub.com/reference/research>
- Baker, D. C. and McLelland, J. N. (2003), 'Evaluating the effectiveness of British Columbia's environmental assessment process for first nations' participation in mining development', *Environmental Impact Assessment Review*, 23(5), 2003/08/01/, 581-603.
- Bamberger, M., Mackay, K. and Ooi, E. (2004), 'Influential evaluations: Evaluations that improved performance and impacts of development programs', *Washington, DC: Operations Evaluation Department, The World Bank*,
- Bamberger, M., Rugh, J. and Mabry, L. (2011) *RealWorld evaluation: Working under budget, time, data, and political constraints*, SAGE,
- Bappeda of DI Yogyakarta Province (2019), *Performance Report of the Regional Development Planning Agency of DI Yogyakarta Province Year 2018*, Bappeda of DI Yogyakarta Province, Yogyakarta.
- Bappeda of West Sumatra Province (2019), *Performance Report of the Regional Development Planning Agency of West Sumatra Province Year 2018*, Bappeda of West Sumatra Province, Padang.
- BAPPENAS (2017) *Anugerah Pangripta Nusantara Tahun 2017 (Development Planning Award Year 2017)*, Kementerian PPN/BAPPENAS, Jakarta.
- Barbarie, A. (1998), *Indonesia's national evaluation system*, Evaluation Capacity Development Working Paper Series, World Bank, Operations Evaluation Department, Washington, DC. https://ieg.worldbankgroup.org/sites/default/files/Data/reports/ecd_wp3.pdf.
- Becker, K. L., Renger, R. and McPherson, M. (2015), 'Indicators of buy-in to gauge evaluation success', *Evaluation Journal of Australasia*, 15(2), 12-21.

- Berke, P., Backhurst, M., Day, M., Ericksen, N., et al. (2006), 'What Makes Plan Implementation Successful? An Evaluation of Local Plans and Implementation Practices in New Zealand', *Environment and Planning B: Planning and Design*, 33(4), 581-600.
- Berke, P. R., Godschalk, D. R., Kaiser, E. J. and Rodriguez, D. A. (2006) *Urban land use planning*, University of Illinois Press Urbana, Urbana-Champaign, IL.
- Bond, A., Morrison-Saunders, A. and Howitt, R. (2013). 'Framework for comparing and evaluating sustainability assessment practice', in *Sustainability assessment: pluralism, practice and progress* (Vol. 117, pp. 117-131), ROUTLEDGE in association with GSE Research, Oxon, UK.
- Booth, A. (2005), 'The Evolving Role of the Central Government in Economic Planning and Policy Making in Indonesia', *Bulletin of Indonesian Economical Studies*, 41(2), 197-219.
- Brody, S. D. (2003), 'Are We Learning to Make Better Plans?: A Longitudinal Analysis of Plan Quality Associated with Natural Hazards', *Journal of Planning Education and Research*, 23(2), 191-201.
- Brody, S. D. and Highfield, W. E. (2005), 'Does planning work? Testing the implementation of local environmental planning in Florida', *Journal of the American Planning Association.*, 71(2), 159-175.
- Bryant, C. and White, L. G. (1982) *Managing Development in the Third World*, Westview Press, Boulder, Colorado.
- Brynard, D. J. (1996). 'Planning: the participatory approach', in B. K. ((ed.), *Citizen participation in local government* (pp. 131-142), JL van Schaik, Pretoria.
- Bustelo, M. (2006), 'The Potential Role of Standards and Guidelines in the Development of an Evaluation Culture in Spain', *Evaluation*, 12(4), 437-453.
- Carden, F. (2018), *Policy and Causality: A learning approach*, WORKING PAPER 28, Knowledge Sector Initiative, https://www.ksi-indonesia.org/old/in/file_upload/Policy-and-Causality-A-Learning-Approach-06Feb2018170706.pdf.
- Cashmore, M., Gwilliam, R., Morgan, R., Cobb, D., et al. (2004), 'The interminable issue of effectiveness: substantive purposes, outcomes and research challenges in the advancement of environmental impact assessment theory', *Impact Assessment and Project Appraisal*, 22(4), 2004/12/01, 295-310.
- Central Bureau of Statistics of DI Yogyakarta Province (2018), *DI Yogyakarta Province in Figures Year 2018*, Central Bureau of Statistics of DI Yogyakarta Province, Yogyakarta.
- Central Bureau of Statistics of West Sumatra Province (2019), *West Sumatra in Figures Year 2018*, Central Bureau of Statistics of West Sumatra Province, Central Bureau of Statistics of West Sumatra Province, Padang, West Sumatra.
- Chanchitpricha, C. and Bond, A. (2013), 'Conceptualising the effectiveness of impact assessment processes', *Environmental Impact Assessment Review*, 43, 2013/11/01/, 65-72.
- Chanchitpricha, C., Bond, A. and Cashmore, M. (2011), 'Effectiveness criteria for measuring Impact Assessment tools', *SEA Implementation and practice: making an impact*, 1-13.
- Cheema, G. S. and Rondinelli, D. A. (2007). 'From Government Decentralization to Decentralized Governance', in G. S. Cheema and D. A. Rondinelli (eds.), *Decentralizing Governance: Emerging Concepts and Practices* (pp. 1-20), Brookings Institution Press, Washington, DC.
- Chelimsky, E. (1994), 'Evaluation: Where we are', *Evaluation Practice*, 15(3), 339-345.
- Chen, H.-T. (1996), 'A Comprehensive Typology for Program Evaluation', *Evaluation Practice*, 17(2), 121-130.

- Chen, H.-t. (2005) *Practical program evaluation : assessing and improving planning, implementation, and effectiveness*, SAGE, Thousand Oaks, Calif.
- Cheng, B.-S., Boer, D., Chou, L.-F., Huang, M.-P., et al. (2013), 'Paternalistic Leadership in Four East Asian Societies: Generalizability and Cultural Differences of the Triad Model', *Journal of Cross-Cultural Psychology*, 45(1), 2014/01/01, 82-90.
- Chouinard, J. A. (2013), 'The Case for Participatory Evaluation in an Era of Accountability', *American Journal of Evaluation*, 34(2), 237-253.
- Christie, C. A. (2003), 'What Guides Evaluation? A Study of How Evaluation Practice Maps onto Evaluation Theory', *New Directions for Evaluation*, 2003(97), 7-36.
- Christie, M., Rowe, P., Perry, C. and Chamard, J. (2000). *Implementation of realism in case study research methodology*. Paper presented at the International Council for Small Business, Annual Conference,
- Conyers, D. (1982) *An introduction to social planning in the Third World*, UMI, Ann Arbor, Mich.
- Conyers, D. and Hills, P. (1984) *Introduction to development planning in the Third World*, John Wiley & Sons Ltd, Chichester.
- Cousins, J. B. and Bourgeois, I. (2014), *Organizational Capacity to Do and Use Evaluation*, in I. Bourgeois (ed.), *New Directions for Evaluation* Number 141, Hoboken : Wiley, Hoboken.
- Cousins, J. B., Goh, S. C., Elliott, C., Aubry, T., et al. (2014), 'Government and voluntary sector differences in organizational capacity to do and use evaluation', *Evaluation and Program Planning*, 44, 2014/06/01/, 1-13.
- Dahler-Larsen, P. and Boodhoo, A. (2019), 'Evaluation culture and good governance: Is there a link?', *Evaluation*, 25(3), 277-293.
- Dale, R. (2004) *Development planning : concepts and tools for planners, managers and facilitators*, Zed Books, London.
- Davies, I. C. (1999), 'Evaluation and Performance Management in Government', *Evaluation*, 5(2), 150-159.
- Davies, P. (2012), 'The State of Evidence- Based Policy Evaluation and its Role in Policy Formation', *National Institute Economic Review*, 219(1), R41-R52.
- De Coning, C. (2006). 'The nature and role of public policy', in F. Cloete, Wissink, H & De Coning, C (ed.), *Improving Public Policy: From Theory to Practice* (2nd Edition ed., pp. 3-26), Van Schaik Publishers, Pretoria.
- Deane, K. L., Dutton, H. and Bullen, P. (2020), 'Theoretically integrative evaluation practice: A step-by-step overview of an eclectic evaluation design process', *Evaluation*, 26(1), 98-118.
- Djakababa, N. D. (2010). *Nrimo And The Resilience Of Humanitarian Volunteers in Post-Earthquake Yogyakarta And Central Java*. Paper presented at the The First International Conference Of Indigenous & Cultural Psychology, Yogyakarta,
- Dye, T. R. (2011) *Understanding public policy* (13th ed. ed.), Longman, New York.
- Eberli, D. (2016), 'Tracing the use of evaluations in legislative processes in Swiss cantonal parliaments', *Evaluation and Program Planning*, 2016/09/28/,
- Efendi, D. (2012). *Local politics and local identity: resistance to " liberal democracy" in Yogyakarta special regions of Indonesia*. University Of Hawai'i at Mānoa, Mānoa, USA.
- Faludi, A. (2000), 'The Performance of Spatial Planning', *Planning Practice & Research*, 15(4), 299-318.

- Ferguson, L. (2003). Developing an evaluative culture [electronic version]. Retrieved, from <http://www.aes.asn.au/images/stories/files/conferences/2003/PAPERS/TU%206%20-%20Ferguson.pdf>. DOI:
- Fetterman, D. M. (1994), 'Empowerment evaluation', *Evaluation Practice*, 15(1), 1-15.
- Firman, T. (2009), 'Decentralization Reform and Local-Government Proliferation in Indonesia: Towards a Fragmentation of Regional Development', *Review of Urban & Regional Development Studies*, 21(2-3), 143-157.
- Forss, K., Rebien, C. C. and Carlsson, J. (2002), 'Process Use of Evaluations: Types of Use that Precede Lessons Learned and Feedback', *Evaluation*, 8(1), 29-45.
- Gilbert, N. (2016) *Researching social life* (4th Edition ed.), SAGE, London.
- Gomm, R., Hammersley, M. and Foster, P. (2000) *Case study method: Key issues, key texts*, Sage,
- Government of DI Yogyakarta Province (2018), *DI Yogyakarta Governor's Accountability Report Year 2017*, Government of DI Yogyakarta Province, Yogyakarta.
- Government of DI Yogyakarta Province (2019), *DI Yogyakarta Governor's Accountability Report Year 2018*, Government of DI Yogyakarta Province, Yogyakarta.
- Government of Indonesia (2004), *Law Number 25 of 2004 on the National Development Planning System*, Government of Indonesia, Jakarta.
- Government of Indonesia (2008), *Government Regulation on Stages, Procedures of Preparation, Control and Evaluation of Regional Development Plan* (Vol. 8),
- Government of West Sumatra Province (2019), *West Sumatra Governor's Accountability Report Year 2018*, Government of West Sumatra Province, Padang.
- Green, J. and South, J. (2006) *Evaluation*, Open University Press, Maidenhead, Berkshire.
- Grimaldi, P., Lau, H. and Basso, M. A. (2015), 'There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making', *Neuroscience & Biobehavioral Reviews*, 55, 2015/08/01/, 88-97.
- Guba, E. G. and Lincoln, Y. S. (1989) *Fourth Generation Evaluation*, Sage, Newbury Park, CA.
- Guyadeen, D. and Seasons, M. (2016), 'Plan Evaluation: Challenges and Directions for Future Research', *Planning Practice & Research*, 31(2), 2016/03/14, 215-228.
- Guyadeen, D. and Seasons, M. (2018), 'Evaluation Theory and Practice: Comparing Program Evaluation and Evaluation in Planning', *Journal of Planning Education and Research*, 38(1), 98-110.
- Hadiwinata, B. S. (2003) *The politics of NGOs in Indonesia: Developing democracy and managing a movement*, Routledge, London.
- Harsono, D. (2018). *A monarchy without a kingdom: Yogyakarta's exceptional system of government*. La Trobe University, Melbourne.
- Haryana, A. (2013). *Monitoring and Evaluation System in Indonesia*. Paper presented at the The Third International Conference on National Evaluation Capacities, São Paulo, Brazil, <https://undp-hq-ieo-nec.azurewebsites.net/wp-content/uploads/2018/12/NEC-proceedings-2013.pdf>.
- Hayton, K. (2015), 'Evaluating the evidence – a move to more Realistic Evaluation: A case study of Regional Selective Assistance in Scotland', *Evaluation*, 21(2), 248-262.
- Head, B. W. (2016), 'Toward More "Evidence-Informed" Policy Making?', *Public Administration Review*, 76(3), 472-484.

- Hellmut, W. (2017), 'Utilization of evaluation results in legal policy-making and administration', *Pravoprimerenie*, 1(1), 7-18.
- Hill, H. and Vidyattama, Y. (2016), 'Regional development dynamics in Indonesia before and after the big bang decentralization', *The Singapore Economic Review*, 61(2), 1.
- Højlund, S. (2014), 'Evaluation use in the organizational context – changing focus to improve theory', *Evaluation*, 20(1), 26-43.
- Holzhaecker, R. L., Wittek, R. and Woltjer, J. (2016) *Decentralization and Governance in Indonesia* (Vol. 2), Springer International Publishing, Cham.
- Huang, X. and Kim, S. E. (2020), 'When top-down meets bottom-up: Local adoption of social policy reform in China', *Governance*, 33(2), 343-364.
- Huang, Y. M. B., Alessandro; (2009) *Reshaping Economic Geography in East Asia*, The World Bank, Washington.
- Huffman, D., Thomas, K. and Lawrenz, F. (2008), 'A collaborative immersion approach to evaluation capacity building', *American Journal of Evaluation*, 29(3), 358-368.
- Hull, A., Alexander, E., Khakee, A. and Woltjer, J. (2012) *Evaluation for participation and sustainability in planning* (1st Edition ed.), Routledge, London.
- IOB Evaluation (2009), *Evaluation Policy and Guidelines for Evaluations*, Ministry of Foreign Affairs, The Hague.
- Irawanto, D. W., Ramsey, P. L. and Tweed, D. C. (2012), 'Exploring paternalistic leadership and its application to the Indonesian public sector', *The British journal of leadership in public services*, 8(1), 4-20.
- Johnson, K., Greenesid, L. O., Toal, S. A., King, J. A., et al. (2009), 'Research on Evaluation Use: A Review of the Empirical Literature from 1986 to 2005', *American Journal of Evaluation*, 30(3), 377-410.
- Kąkol, U., Kisilowski, M., Kunikowski, G. and Uklańska, A. (2018), 'Adaptation of Civil Planning and Crisis Management Practices Based on Hofstede's Cultural Dimensions Model', *Management Research & Practice*, 10,
- Kallemeyn, L. M., Hall, J., Friche, N. and McReynolds, C. (2015), 'Cross-Continental Reflections on Evaluation Practice: Methods, Use, and Valuing', *American Journal of Evaluation*, 36(3), 339-357.
- Karina, V. D. (2014). *Pengaruh nilai Budaya Jawa (rukun-hormat) terhadap pemaafan dalam konteks hubungan remaja dengan orang tua (The influence of Javanese cultural values (rukun-respect) on forgiveness in the context of adolescent relationships with parents)*. University of Muhammadiyah Malang, Malang.
- Kazi, M. A. F. (2003), *The Contribution of Realist Evaluation for Practice*, Realist Evaluation in Practice (pp. 158-167), SAGE Publications Ltd, London. <http://methods.sagepub.com/book/realist-evaluation-in-practice>
- Kemitraan (2013), *Executive Report Indonesia Governance Index 2012*, The Partnership for Governance Reform, South Jakarta.
- Khakee, A. (1998), 'Evaluation and planning: inseparable concepts', *Town Planning Review*, 69(4), 359.
- Khakee, A. (2003), 'The Emerging Gap between Evaluation Research and Practice', *Evaluation*, 9(3), 340-352.
- Kimura, F. (2005), 'International Production/Distribution Networks and Indonesia', *The Developing Economies*, 43(1), 17-38.

- Kitchenham, B. (2004), 'Procedures for performing systematic reviews', *Keele, UK, Keele University*, 33(2004), 1-26.
- Kosasih, A. (2013), 'Upaya Penerapan Nilai-nilai Adat dan Syarak dalam Penyelenggaraan Pemerintahan Nagari (Efforts to Implement Customary and Syarak Values in the Implementation of Nagari Governance)', *Humanus*, 12(2), 107-119.
- Krizek, K., Forysth, A. and Slotterback, C. S. (2009), 'Is There a Role for Evidence-Based Practice in Urban Planning and Policy?', *Planning Theory & Practice*, 10(4), 2009/12/01, 459-478.
- Kusak, J. Z. and Rist, R. C. (2004) *Ten Steps to a Results-Based Monitoring and Evaluation System : A Handbook for Development Practitioners*, World Bank Publications, Washington.
- Lacouture, A., Breton, E., Guichard, A. and Ridde, V. (2015), 'The concept of mechanism from a realist approach: a scoping review to facilitate its operationalization in public health program evaluation', *Implementation Science*, 10(1), 1-10.
- Landiyanto, E. A. (2015), *Transformation of the National Monitoring and Evaluation Arrangement in Decentralized Indonesia* (28 January 2016 ed.), MPRA, <https://mpra.ub.uni-muenchen.de/69073/>
- Lasker, R. D., Weiss, E. S. and Miller, R. (2001), 'Partnership Synergy: A Practical Framework for Studying and Strengthening the Collaborative Advantage', *The Milbank Quarterly*, 79(2), 179-205.
- Laurian, L., Crawford, J., Day, M., Kouwenhoven, P., et al. (2010), 'Evaluating the Outcomes of Plans: Theory, Practice, and Methodology', *Environment and Planning B: Planning and Design*, 37(4), 740-757.
- Laurian, L., Day, M., Berke, P., Ericksen, N., et al. (2004), 'Evaluating Plan Implementation: A Conformance-Based Methodology', *Journal of the American Planning Association*, 70(4), 471-480.
- Leeuw, F. L. and Furubo, J.-E. (2008), 'Evaluation systems: What are they and why study them?', *Evaluation*, 14(2), 157-169.
- Lewis, W. A. (1966) *Development planning : the essentials of economic policy*, Allen & Unwin, London.
- Linhorst, D. M. (2002), 'A review of the use and potential of focus groups in social work research', *Qualitative Social Work*, 1(2), 208-228.
- Longhurst, R. (2009). 'Interviews: In-Depth, Semi-Structured A2 - Kitchin, Rob', in N. Thrift (ed.), *International Encyclopedia of Human Geography* (pp. 580-584), Elsevier, Oxford.
- Loomis, J. J. and Dziedzic, M. (2018), 'Evaluating EIA systems' effectiveness: A state of the art', *Environmental Impact Assessment Review*, 68, 2018/01/01/, 29-37.
- Loud, M. L. and Mayne, J. (2014), *Enhancing Evaluation Use: Insights from Internal Evaluation Units*, SAGE Publication, Inc., Thousand Oaks, California. <http://methods.sagepub.com.ezproxy.flinders.edu.au/book/enhancing-evaluation-use>
- Luttrell, A., Briñol, P., Petty, R. E., Cunningham, W., et al. (2013), 'Metacognitive confidence: A neuroscience approach', *International Journal of Social Psychology*, 28(3), 2013/01/01, 317-332.
- MacRae, G. and Hodgkin, D. (2016), 'Beyond the 2006 Yogyakarta earthquake: from sectors to clusters in the international humanitarian system', *Rebuilding Asia Following Natural Disasters: Approaches to Reconstruction in the Asia-Pacific Region*, Cambridge University Press, Cambridge, 261-283.
- Marchal, B., Van Belle, S., Van Olmen, J., Hoérée, T., et al. (2012), 'Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research', *Evaluation*, 18(2), 192-212.

- Marcic, D. (1992) *Organizational Behavior: Experiences and Cases* (3rd edition ed.), West Publishing Co., St. Paul, MN.
- Martin, S. and Sanderson, I. (1999), 'Evaluating Public Policy Experiments', *Evaluation*, 5(3), 245-258.
- Martz, W. (2013), *Organizational effectiveness evaluation checklist*, <http://comm.eval.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=028f9f00-830d-4110-864c-fdc2f5817a04&forceDialog=0>
- Masterson-Algar, P., Burton, C. R., Rycroft-Malone, J., Sackley, C. M., et al. (2014), 'Towards a programme theory for fidelity in the evaluation of complex interventions', *Journal of evaluation in clinical practice*, 20(4), 445-452.
- Mastop, H. and Faludi, A. (1997), 'Evaluation of Strategic Plans: The Performance Principle', *Environment and Planning B: Urban Analytics and City Science*, 24(6), 815-832.
- Maxwell, J. A. (2012) *A realist approach for qualitative research*, SAGE Publications, Inc, Thousand Oaks, California.
- Mayne, J. (2008), 'Building an Evaluative Culture for Effective Evaluation and Results Management', *ILAC Working Paper 8*,
- McDavid, J. C. (2006) *Program evaluation & performance measurement : an introduction to practice*, SAGE, Thousand Oaks, California.
- McDonald, B., Rogers, P. and Kefford, B. (2003), 'Teaching People to Fish? Building the Evaluation Capability of Public Sector Organizations', *Evaluation*, 9(1), 9-29.
- McNamara, G., O'Hara, J., Boyle, R. and Sullivan, C. (2009), 'Developing a Culture of Evaluation in the Irish Public Sector The Case of Education', *Evaluation*, 15(1), 101-112.
- Meer, F.-B. V. D. and Edelenbos, J. (2006), 'Evaluation in Multi-Actor Policy Processes', *Evaluation*, 12(2), 201-218.
- Mertens, D. M. (1999), 'Inclusive Evaluation: Implications of Transformative Theory for Evaluation', *American Journal of Evaluation*, 20(1), 1-14.
- Mihalache, R. (2010), 'A Developing Evaluation Culture in Romania: Myths, Gaps and Triggers', *Evaluation*, 16(3), 323-332.
- Miles, M. B., Huberman, A. M. and Saldaña, J. (2014) *Qualitative data analysis : a methods sourcebook* (Edition 3 ed.), SAGE Publications, Inc., Thousand Oaks, California
- Milstein, B., Chapel, T. J., Wetterhall, S. F. and Cotton, D. A. (2002), 'Building Capacity for Program Evaluation at the Centers for Disease Control and Prevention', *New Directions for Evaluation*, 93, 27-46.
- Minister of Home Affairs (2017), *Minister of Home Affairs Regulation No. 86 of 2017 Concerning Tata Cara Perencanaan, Pengendalian dan Evaluasi Pembangunan Daerah, Tata Cara Evaluasi Rancangan Peraturan Daerah tentang Rencana Pembangunan Pembangunan Jangka Menengah Daerah, serta Tata Cara Perubahan Rencana Pembangunan Jangka Panjang Daerah, Rencana Pembangunan Jangka Menengah Daerah, dan Rencana Kerja Pemerintah Daerah (Procedures for Planning, Controlling and Evaluation of Regional Development, Procedures for Evaluation of Draft Regional Regulations concerning Regional Medium-Term Development Plans, as well as Procedures for Amending Regional Long-Term Development Plans, Regional Medium-Term Development Plans, and Regional Government Work Plans)*, Government of Indonesia, Jakarta.

- Molas-Gallart, J. (2012), 'Research Governance and the Role of Evaluation: A Comparative Study', *American Journal of Evaluation*, 33(4), 583-598.
- Moore, B. and Spires, R. (2000). 'Monitoring and Evaluation', in P. Roberts and H. Sykes (eds.), *Urban regeneration: A handbook* (pp. 203-227), Sage Publications, London.
- Mora, C. and Antonie, R. (2012), 'Levers supporting program evaluation culture and capacity in Romanian public administration: The role of leadership', *Society and Economy*, 34(3), 423-432.
- Mueller, G. P. and Hersperger, A. M. (2015), 'Implementing comprehensive plans: indicators for a task-sheet based performance evaluation process', *Journal of Environmental Planning and Management*, 58(11), 2015/11/02, 2056-2081.
- Muhammad, M. (2017) *Perencanaan Pembangunan (Development Planning)*, CV Dua Bersaudara, Makassar.
- Nasution, A. (2017). 'The Government Decentralization Program in Indonesia', in A. D. B. Institute (ed.), *Central and Local Government Relations in Asia*, Edward Elgar Publishing Limited, Cheltenham, UK.
- Nelmaya, N. (2018), '*Kembali ke Nagari Rekonstruksi Masyarakat: Adat Basandi Syarak, Syarak Basandi Kitabullah (Studi Aplikasi dan Formulasi Perda Pemerintahan Sumatera Barat no 9 Tahun 2000)*/Back to Nagari Community Reconstruction: Adat Basandi Syarak, Syarak Basandi Kitabullah (Study on Application and Formulation of West Sumatra Government Regulation No. 9 of 2000)', *Al Hurriyah: Jurnal Hukum Islam*, 11(1), 85-98.
- O'Sullivan, R. G. (2004). 'Evaluation Practice', in *Practicing evaluation* (pp. 2-22), SAGE Publications, Inc., Thousand Oaks, CA.
- OECD (2002) *Glossary of key terms in evaluation and results based management*, OECD, Paris.
- OECD/DAC (2012), *Assessing the Development Effectiveness of Multilateral Organizations: Guidance on the Methodological Approach* OECD, <https://www.oecd.org/dac/evaluation/dcdndep/50540172.pdf>.
- Olejniczak, K. (2013), 'Mechanisms Shaping an Evaluation System—A Case Study of Poland 1999–2010', *Europe-Asia Studies*, 65(8), 2013/10/01, 1642-1666.
- Olejniczak, K., Raimondo, E. and Kupiec, T. (2016), 'Evaluation units as knowledge brokers: Testing and calibrating an innovative framework', *Evaluation*, 22(2), 168-189.
- Oliveira, V. and Pinho, P. (2010a), 'Evaluation in Urban Planning: Advances and Prospects', *Journal of Planning Literature*, 24(4), 343-361.
- Oliveira, V. and Pinho, P. (2010 May-June). Measuring success in planning: Developing and testing a methodology for planning evaluation. *Town Planning Review*, 81, 307-332 <http://go.galegroup.com/ps/i.do?p=EAIM&sw=w&u=flinders&v=2.1&it=r&id=GALE%7CA229842678&asid=6a73517c92263f37661ce296582ac421>.
- Oliveira, V. and Pinho, P. (2011), 'Bridging the gap between planning evaluation and programme evaluation: The contribution of the PPR methodology', *Evaluation*, 17(3), 293-307.
- Owen, J. M. (2003), 'Evaluation culture: A definition and analysis of its development within organisations', *Evaluation Journal of Australasia*, 3(1), 43-47.
- Pagatpatan, C. P. and Ward, P. R. (2017), 'Understanding the factors that make public participation effective in health policy and planning: a realist synthesis', *Australian journal of primary health*, 23(6), 516-530.

- Pambudi, A. N. (2016). *Hubungan Nilai Budaya Jawa Rukun dan Nrimo dengan Subjective Well-Being Suku Jawa (The Relationship between Rukun and Nrimo Javanese Cultural Values with Javanese Subjective Well-Being)*. University of Muhammadiyah Malang, Malang.
- Parker, A. and Tritter, J. (2006), 'Focus group method and methodology: current practice and recent debate', *International Journal of Research & Method in Education*, 29(1), 2006/04/01, 23-37.
- Patton, M. Q. (2002) *Qualitative evaluation and research methods* (3rd ed. ed.), Sage Publications, Thousand Oak.
- Patton, M. Q. (2008) *Utilization-focused evaluation* (4th ed. ed.), Sage Publications, Thousand Oaks, Calif.
- Pawson, R. (2006) *Evidence-based policy : a realist perspective*, Sage, London.
- Pawson, R. (2013) *The science of evaluation : a realist manifesto*, London : SAGE, London.
- Pawson, R. (2018). 'Realist Memorabilia', in N. Emmel, J. Greenhalgh, A. Manzano, M. P. Monaghan and S. Dalkin (eds.), *Doing realist research* (pp. 203-220), SAGE, London.
- Pawson, R. and Tilley, N. (1997a), *An Introduction to Scientific Realist Evaluation*, in E. C. W. R. Shadish (ed.), *Evaluation for the 21st Century: A Handbook* (pp. 405-418), SAGE Publications, Inc., Thousand Oaks, California. <http://methods.sagepub.com/book/evaluation-for-the-21st-century>
- Pawson, R. and Tilley, N. (1997b) *Realistic Evaluation*, SAGE, London.
- Pearsall, B. (2013), 'Evaluating sexual assault nurse examiner programs', *National Institute of Justice Journal*, 272, 41-46.
- Petticrew, M. and Roberts, H. (2008) *Systematic Reviews in the Social Sciences*, Blackwell Publishing Ltd,
- Punton, M., Vogel, I. and Lloyd, R. (2016), 'Reflections from a realist evaluation in progress: scaling ladders and stitching theory',
- Ramírez, R. and Brodhead, D. (2013) *Utilization focused evaluation: A primer for evaluators*, Southbound Sdn. Bhd, Penang, MY.
- Riyadi and Bratakusumah, D. S. (2003) *Perencanaan Pembangunan Daerah: Strategi Menggali Potensi dalam Mewujudkan Otonomi Daerah (Regional Development Planning: Strategies for Exploring Potential in Realizing Regional Autonomy)*, Gramedia Pustaka Utama, Jakarta.
- Roberts, P. (2006), 'Evaluating regional sustainable development: Approaches, methods and the politics of analysis', *Journal of Environmental Planning and Management*, 49(4), 2006/07/01, 515-532.
- Rogers, P. J. (2000). 'Program Theory: Not Whether Programs Work but How They Work', in D. L. Stufflebeam, G. F. Madaus and T. Kellaghan (eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (Second Edition. ed., Vol. 49, pp. 209-232), Springer Netherlands, Dordrecht.
- Salim, D. P. (2015) *The Transnational and the Local in the Politics of Islam: The Case of West Sumatra, Indonesia*, Springer International Publishing, Switzerland.
- Salter, K. L. and Kothari, A. (2014), 'Using realist evaluation to open the black box of knowledge translation: a state-of-the-art review', *Implementation science*, 9(1), 115.
- Sanderson, I. (2000), 'Evaluation in Complex Policy Systems', *Evaluation*, 6(4), 433-454.
- Sanderson, I. (2002), 'Evaluation, Policy Learning and Evidence-Based Policy Making', *Public Administration*, 80(1), 1-22.
- Saunders, M. (2012), 'The use and usability of evaluation outputs: A social practice approach', *Evaluation*, 18(4), 421-436.
- Sayer, A. (2000) *Realism and social science*, SAGE Publications, London.

- Scharpf, F. W. (1997) *Games Real Actors Play. Actor-Centred Institutionalism in Policy Research*, Westview Press, Boulder.
- Scott, W. R. (2014) *Institutions and organizations : ideas, interests, and identities* (Fourth edition. ed.), SAGE Publications, Inc., Thousand Oaks, California.
- Scriven, M. (1991) *Evaluation thesaurus* (4th ed. ed.), Sage Publications, Newbury Park, Calif.
- Seasons, M. (2002a), 'Evaluation and municipal urban planning: practice and prospects', *The Canadian Journal of Program Evaluation*, 17(1), 43-71.
- Seasons, M. (2002b), 'Evaluation and municipal urban planning: practice and prospects', *The Canadian Journal of Program Evaluation*, 17(1), 43.
- Seasons, M. (2003), 'Monitoring and evaluation in municipal planning: considering the realities', *Journal of the American Planning Association.*, 69(4), 430-440.
- Segone, M., (ed.) (2008), *Bridging the gap: The role of monitoring and evaluation in Evidence-based policy making*, UNICEF Evaluation Office, https://www.unicef.org/ceecis/evidence_based_policy_making.pdf
- Shaxson, L. (2016), *Lessons for building and managing an evidence base for policy*, Working Paper 10, Knowledge Sector Initiative, Jakarta. <https://www.ksi-indonesia.org/assets/uploads/original/2020/01/ksi-1580269971.pdf>
- Shulha, L. M. and Cousins, J. B. (1997), 'Evaluation use: Theory, research, and practice since 1986', *American Journal of Evaluation*, 18(1), 195-208.
- Sjafrizal (2016) *Regional Development Planning in the Era of Autonomy (Perencanaan Pembangunan Daerah dalam Era Otonomi)*, Rajawali Pers, Jakarta.
- Stern, E. (2006). 'Contextual challenges for evaluation practice', in I. Shaw, J. C. Greene and M. M. Mark (eds.), *SAGE handbook of evaluation* (pp. 292-314), SAGE, London.
- Stockmann, R. (2011). 'Competing and complementary approaches to evaluation', in R. Stockmann (ed.), *A Practitioner Handbook on Evaluation*, Edward Edgar, Cheltenham.
- Stockmann, R., Meyer, W. and Taube, L. (2020). 'The Institutionalisation of Evaluation in Europe: A Synthesis', in R. Stockmann, W. Meyer and L. Taube (eds.), *The Institutionalisation of Evaluation in Europe* (pp. 483-522), Springer International Publishing, Cham.
- Stottele-Ishmi, D. (2017), *Evaluation Systems Strengthening*, Bappenas and Unicef, Jakarta.
- Stufflebeam, D. L. (2001), 'Evaluation Checklists: Practical Tools for Guiding and Judging Evaluations', *American Journal of Evaluation*, 22(1), 71-79.
- Subkhan, I. (2014), 'GBHN dan Perubahan Perencanaan Pembangunan di Indonesia (GBHN and Changes in Development Planning in Indonesia)', *Aspirasi*, 5(2), 131-143.
- Sutmuller, P. M. and Setiono, I. (2011), 'Diagnostic on evidence-based public policy formulation under decentralisation', *Report to AusAID*, 01-32.
- Talitha, T., Firman, T. and Hudalah, D. (2020), 'Welcoming two decades of decentralization in Indonesia: a regional development perspective', *Territory, Politics, Governance*, 8(5), 2020/10/19, 690-708.
- Taut, S. and Alkin, M. C. (2010). 'The Role of Stakeholders in Educational Evaluation', in P. Peterson, E. Baker and B. McGaw (eds.), *International Encyclopedia of Education (Third Edition)* (pp. 629-635), Elsevier, Oxford.
- Taylor, D. and Balloch, S. (2005) *The politics of evaluation: Participation and policy implementation*, The Policy Press, Bristol, UK.

- Theophilou, V., Bond, A. and Cashmore, M. (2010), 'Application of the SEA Directive to EU structural funds: Perspectives on effectiveness', *Environmental Impact Assessment Review*, 30(2), 2010/02/01/, 136-144.
- Todaro, M. P. (1971) *Development planning models and methods*, Oxford University Press, Nairobi.
- Todaro, M. P. (1994) *Economic development* (5th ed. ed.), Longman, London.
- Torres, R. T. and Preskill, H. (2001), 'Evaluation and Organizational Learning: Past, Present, and Future', *American Journal of Evaluation*, 22(3), 387-395.
- Turner, M. M. and Hulme, D. (1997) *Governance, administration and development : making the state work*, Macmillan Press Ltd, Houndmills, Basingstoke.
- Visser, I., Kusters, C., Guijt, I., Roefs, M., et al. (2014). *Improving the Use of Monitoring & Evaluation Processes and Findings*. Paper presented at the Improving the Use of Monitoring & Evaluation Processes and Findings, Wageningen,
- Waldner, L. S. (2004), 'Planning To Perform: Evaluation Models For City Planners', *Berkeley Planning Journal*, 17(1),
- Wasono, A. and Maulana, M. (2018), *Critical Study on Development Planning and Budgeting in Indonesia*, KSI Working Paper 27, Knowledge Sector Initiative, Jakarta.
- Webster, D. R. and Robinson, I. M. (1985), 'Regional Planning in Canada History, Practice, Issues, and Prospects ', *Journal of the American Planning Association*, 51(1), 1985/03/31, 23-33.
- Weiss, C. H. (1972). 'Utilization of evaluation: Toward comparative study', in *Evaluating action programs: Readings in social action education* (pp. 318-326), Allyn & Bacon, Boston.
- Weiss, C. H. (1998) *Evaluation : methods for studying programs and policies* (2nd ed.), Prentice Hall, Upper Saddle River, N.J.
- Weiss, C. H., Murphy-Graham, E., Petrosino, A. and Gandhi, A. G. (2008), 'The Fairy Godmother—and Her Warts: Making the dream of evidence-based policy come true', *American Journal of Evaluation*, 29(1), 29-47.
- Weissmann, E. (1969). *Regional Development Planning*. Paper presented at La Regionalizacion De Las Politicas De Desarrollo en America Latina, Santiago, Chile,
- Westhorp, G. (2018). 'Understanding mechanisms in realist evaluation and research', in N. Emmel, J. Greenhalgh, A. Manzano, M. P. Monaghan and S. Dalkin (eds.), *Doing realist research* (pp. 41-57), SAGE, London.
- White, H. (2009), 'Theory-based impact evaluation: principles and practice', *Journal of Development Effectiveness*, 1(3), 2009/09/15, 271-284.
- Widmer, T. and Neuenschwander, P. (2004), 'Embedding Evaluation in the Swiss Federal Administration', *Evaluation*, 10(4), 388-409.
- Wilkinson, S. (1998), 'Focus group methodology: a review', *International Journal of Social Research Methodology*, 1(3), 1998/01/01, 181-203.
- Wilkinson, S. (2004). 'Focus Group Research', in D. Silverman (ed.), *Qualitative research: Theory, method and practice* (pp. 177-199), SAGE Publications, London.
- Worthen, B. R. and Schmitz, C. C. (1997), 'Conceptual Challenges Confronting Cluster Evaluation', *Evaluation*, 3(3), 300-319.

- Yağci, M. and Ardiani, N. (2018). 'National Development Planning, Industrial Policy, and Sustainable Growth Challenges in Indonesia and Malaysia: A Comparative Historical Analysis', in M. A. Yülek (ed.), *Industrial Policy and Sustainable Growth* (pp. 213-241), Springer, Singapore.
- Yasih, D. W. P. (2017), 'Jakarta's Precarious Workers: Are they a "New Dangerous Class"?', *Journal of contemporary Asia*, 47(1), 27-45.
- Yin, R. K. (2014) *Case study research : design and methods* SAGE, Los Angeles.
- Zhang, D. (2015), *Do Local Governments in Indonesia Produce Evidence-Based Policies?*, The Australia Indonesia Partnership for Decentralisation (AIPD), Jakarta.

APPENDICES

Appendix 1: Final and Modification Ethics Approval

FINAL APPROVAL NOTICE

Project No.:

7820

Project Title:

Effectiveness of Evaluation Practice in Supporting Regional Development and Planning (Case Studies: DI Yogyakarta Province and Binjai City, Indonesia)

Principal Researcher:

Mrs Dwi Ratih Suryantining Esti

Email:

esti0005@flinders.edu.au

Approval Date:

20 November 2017

Ethics Approval Expiry Date:

20 February 2022

The above proposed project has been **approved** on the basis of the information contained in the application, its attachments and the information subsequently provided.

RESPONSIBILITIES OF RESEARCHERS AND SUPERVISORS

1. Participant

Documentation

Please note that it is the responsibility of researchers and supervisors, in the case of student projects, to ensure that:

- all participant documents are checked for spelling, grammatical, numbering and formatting errors. The Committee does not accept any responsibility for the above mentioned errors.
- the Flinders University logo is included on all participant documentation (e.g., letters of Introduction, information Sheets, consent forms, debriefing information and questionnaires – with the exception of purchased research tools) and the current Flinders University letterhead is included in the header of all letters of introduction. The Flinders University international logo/letterhead should be used and documentation should contain international dialling codes for all telephone and fax numbers listed for all research to be conducted overseas.
- the SBREC contact details, listed below, are included in the footer of all letters of introduction and information sheets.

This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project Number 'INSERT PROJECT No. here following approval'). For more information regarding ethical approval of the project the Executive Officer of the Committee can be contacted by telephone on 8201 3116, by fax on 8201 2035 or by email human.researchethics@flinders.edu.au.

2. Annual Progress / Final Reports

In order to comply with the monitoring requirements of the [National Statement on Ethical Conduct in Human Research \(March 2007\)](#) an annual progress report must be submitted each year on the **20 November** (approval anniversary date) for the duration of the ethics approval using the report template available from the [Managing Your Ethics Approval](#) SBREC web page. *Please retain this notice for reference when completing annual progress or final reports.*

If the project is completed *before* ethics approval has expired please ensure a final report is submitted immediately. If ethics approval for your project expires please submit either (1) a final report; or (2) an extension of time request and an annual report.

Student Projects

The SBREC recommends that current ethics approval is maintained until a student's thesis has been submitted, reviewed and approved. This is to protect the student in the event that reviewers recommend some changes that may include the collection of additional participant data.

Your first report is due on **20 November 2018** or on completion of the project, whichever is the earliest.

3. Modifications to Project

Modifications to the project must not proceed until approval has been obtained from the Ethics Committee. Such proposed changes / modifications include:

- change of project title;
- change to research team (e.g., additions, removals, principal researcher or supervisor change);
- changes to research objectives;
- changes to research protocol;
- changes to participant recruitment methods;
- changes / additions to source(s) of participants;
- changes of procedures used to seek informed consent;
- changes to reimbursements provided to participants;
- changes / additions to information and/or documentation to be provided to potential participants;
- changes to research tools (e.g., questionnaire, interview questions, focus group questions);

- extensions of time.

To notify the Committee of any proposed modifications to the project please complete and submit the *Modification Request Form* which is available from the [Managing Your Ethics Approval SBREC](#) web page. Download the form from the website every time a new modification request is submitted to ensure that the most recent form is used. Please note that extension of time requests should be submitted prior to the Ethics Approval Expiry Date listed on this notice.

Change of Contact Details

Please ensure that you notify the Committee if either your mailing or email address changes to ensure that correspondence relating to this project can be sent to you. A modification request is not required to change your contact details.

4. Adverse Events and/or Complaints

Researchers should advise the Executive Officer of the Ethics Committee on 08 8201-3116 or human.researchethics@flinders.edu.au immediately if:

- any complaints regarding the research are received;
- a serious or unexpected adverse event occurs that effects participants;
- an unforeseen event occurs that may affect the ethical acceptability of the project.

MODIFICATION (No.1) APPROVAL NOTICE

Project No.:

7820

Project Title:

Effectiveness of Evaluation Practice in Supporting Regional Development and Planning

Principal Researcher:

Mrs Dwi Ratih Suryantining Esti

Email:

esti0005@flinders.edu.au

Modification
Approval Date:

12 December 2017

Ethics Approval
Expiry Date:

20 February 2022

I am pleased to inform you that the modification request submitted for project 7820 on the 21 November 2017 has been reviewed and approved by the SBREC Deputy Chairperson. A summary of the approved modifications are listed below. Any additional information that may be required from you will be listed in the second table shown below called 'Additional Information Required'.

Approved Modifications	
Extension of ethics approval expiry date	
Project title change	X
Personnel change	
Research objectives change	
Research method change	
Participants – addition +/- change	
Consent process change	
Recruitment process change	
Research tools change	
Document / Information Changes	
Other (if yes, please specify)	

Additional Information Required
None.

RESPONSIBILITIES OF RESEARCHERS AND SUPERVISORS

1. Participant Documentation

Please note that it is the responsibility of researchers and supervisors, in the case of student projects, to ensure that:

- all participant documents are checked for spelling, grammatical, numbering and formatting errors. The Committee does not accept any responsibility for the above mentioned errors.
- the Flinders University logo is included on all participant documentation (e.g., letters of Introduction, information Sheets, consent forms, debriefing information and questionnaires – with the exception of purchased research tools) and the current Flinders University letterhead is included in the header of all letters of introduction. The Flinders University international logo/letterhead should be used and documentation should contain international dialling codes for all telephone and fax numbers listed for all research to be conducted overseas.
- the SBREC contact details, listed below, are included in the footer of all letters of introduction and information sheets.

This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project Number 'INSERT PROJECT No. here following approval'). For more information regarding ethical approval of the project the Executive Officer of the Committee can be contacted by telephone on 8201 3116, by fax on 8201 2035 or by email human.researchethics@flinders.edu.au.

2. Annual Progress / Final Reports

Please be reminded that in order to comply with the monitoring requirements of the [National Statement on Ethical Conduct in Human Research \(March 2007\)](#) an annual progress report must be submitted each year on **20 November** (approval anniversary date) for the duration of the ethics approval.

If the project is completed *before* ethics approval has expired please ensure a final report is submitted immediately. If ethics approval for your project expires please submit either (1) a final report; or (2) an extension of time request and an annual report.

Student Projects

The SBREC recommends that current ethics approval is maintained until a student's thesis has been submitted, reviewed and approved. This is to protect the student in the event that reviewers recommend some changes that may include the collection of additional participant data.

Your next report is due on **20 November 2018** or on completion of the project, whichever is the earliest. The report template is available from the [Managing Your Ethics Approval](#) SBREC web page. *Please retain this notice for reference when completing annual progress or final reports.*

3. Modifications to Project

Modifications to the project must not proceed until approval has been obtained from the Ethics Committee. Such proposed changes / modifications include:

- change of project title;
- change to research team (e.g., additions, removals, principal researcher or supervisor change);

- changes to research objectives;
- changes to research protocol;
- changes to participant recruitment methods;
- changes / additions to source(s) of participants;
- changes of procedures used to seek informed consent;
- changes to reimbursements provided to participants;
- changes / additions to information and/or documentation to be provided to potential participants;
- changes to research tools (e.g., questionnaire, interview questions, focus group questions);
- extensions of time.

To notify the Committee of any proposed modifications to the project please complete and submit the *Modification Request Form* which is available from the [Managing Your Ethics Approval](#) SBREC web page. Download the form from the website every time a new modification request is submitted to ensure that the most recent form is used. Please note that extension of time requests should be submitted prior to the Ethics Approval Expiry Date listed on this notice.

Change of Contact Details

Please ensure that you notify the Executive Officer if either your mailing or email address changes to ensure that correspondence relating to this project can be sent to you. A modification request is not required to change your contact details.

4. Adverse Events and/or Complaints

Researchers should advise the [Executive Officer](#) immediately on 08 8201-3116 or human.researchethics@flinders.edu.au if:

- any complaints regarding the research are received;
- a serious or unexpected adverse event occurs that effects participants;
- an unforeseen event occurs that may affect the ethical acceptability of the project.

APPENDIX 2: LETTER OF INTRODUCTION (INTERVIEW)



College of Business, Government and Law
Room 330, Social Sciences South
GPO Box 2100
Adelaide SA 5001
Tel: 08 82012302
noore.siddiquee@flinders.edu.au
<http://www.flinders.edu.au/people/noore.siddiquee>
CRICOS Provider No. 00114A

November 1, 2017

LETTER OF INTRODUCTION

Dear Sir/Madam

This letter is to introduce Dwi Ratih S. Esti who is a PhD student in the College of Business, Government and Law at Flinders University. She will produce her student card, which carries a photograph, as proof of identity.

She is undertaking research leading to the production of a thesis on the subject of Effectiveness of Evaluation Practice in Supporting Regional Development Planning.

She would like to invite you to assist with this project by agreeing to be involved in an interview which covers certain aspects of this topic. No more than two hours on one occasion would be required.

Be assured that any information provided will be treated in the strictest confidence and none of the participants will be individually identifiable in the resulting thesis, report or other publications. You are, of course, entirely free to decline to participate in the interview or to answer particular questions.

Since she intends to make a tape recording of the interview, she will seek your consent, on the attached form, to record the interview, to use the recording or a transcription in preparing the thesis, report or other publications, on condition that your name or identity is not revealed.

If you are willing to be involved in this research, please send an email confirming your involvement in the research to the researcher at esti0005@flinders.edu.au with a signed consent form for the individual interview.

Thank you for your attention and assistance.

Yours sincerely

A handwritten signature in black ink, appearing to read 'Noore A. Siddiquee', with a long horizontal stroke extending to the right.

Associate Professor Noore A. Siddiquee

College of Business, Government and Law

This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project number 7820). For more information regarding ethical approval of the project the Executive Officer of the Committee can be contacted by telephone on 8201 3116, by fax on 8201 2035 or by email human.researchethics@flinders.edu.au

APPENDIX 3: LETTER OF INTRODUCTION (FOCUS GROUP)



College of Business, Government and Law
Room 330, Social Sciences South
GPO Box 2100
Adelaide SA 5001
Tel: 08 82012302
noore.siddiquee@flinders.edu.au
<http://www.flinders.edu.au/people/noore.siddiquee>
CRICOS Provider No. 00114A

November 1, 2017

LETTER OF INTRODUCTION

Dear Sir/Madam

This letter is to introduce Dwi Ratih S. Esti who is a PhD student in the College of Business, Government and Law at Flinders University. She will produce her student card, which carries a photograph, as proof of identity.

She is undertaking research leading to the production of a thesis on the subject of Effectiveness of Evaluation Practice in Supporting Regional Development Planning.

She would like to invite you to assist with this project by agreeing to be involved in a focus group discussion which covers certain aspects of this topic. No more than two hours on one occasion would be required.

Be assured that any information provided will be treated in the strictest confidence and none of the participants will be individually identifiable in the resulting thesis, report or other publications. You are, of course, entirely free to decline to participate in the focus group or to answer particular questions.

Since she intends to make a tape recording of the focus group discussion, she will seek your consent, on the attached form, to record the focus group discussion, to use the recording or a transcription in preparing the thesis, report or other publications, on condition that your name or identity is not revealed.

If you are willing to be involved in this research, please send an email confirming your involvement in the research to the researcher at esti0005@flinders.edu.au with a signed consent form for the focus group.

Thank you for your attention and assistance.

Yours sincerely

A handwritten signature in black ink, appearing to read 'Noore A. Siddiquee', with a long horizontal stroke extending to the right.

Associate Professor Noore A. Siddiquee

College of Business, Government and Law

This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project number 7820). For more information regarding ethical approval of the project the Executive Officer of the Committee can be contacted by telephone on 8201 3116, by fax on 8201 2035 or by email human.researchethics@flinders.edu.au

APPENDIX 4: INFORMATION SHEET (INTERVIEW)



Ms Dwi Ratih S. Esti
College of Business, Government and Law

Sturt Road
Bedford Park SA 5042

GPO Box 2100
Adelaide SA 5001

esti0005@flinders.edu.au

CRICOS Provider No. 00114A

INFORMATION SHEET

(for *Interview Respondents*)

Title: 'Effectiveness of Evaluation Practice in Supporting Regional Development Planning'

Researcher(s):

Ms Dwi Ratih S. Esti

College of Business, Government and Law

Flinders University

Ph: +61 8 82012647

Position in Indonesia: staff in the Ministry of National Development Planning/BAPPENAS

Supervisor(s):

A/Prof Noore Siddiquee

College of Business, Government and Law

Flinders University

Ph: +61 8 82012302

A/Prof Gerry Redmond

College of Business, Government and Law

Flinders University

Ph: +61 8 82012699

Description of the study

This research will provide an in-depth overview of effectiveness of evaluation practice in supporting regional development planning in Indonesia using two case studies. This overview includes the perspectives of evaluation practitioners, intended users of evaluations (policy makers), local leaders, academicians, officials of NGOs, and the central bureau of statistics. This project is supported by Flinders University, College of Business, Government, and Law.

Purpose of the study

This project aims to understand the influencing factors, challenges, and limitations of the effectiveness of evaluation practices in supporting regional development planning process.

What will I be asked to do?

You are invited to attend an interview with a researcher who will ask you a few questions about your views about the effectiveness of evaluation practices in supporting regional development planning. Participation is entirely voluntary. The interview will take about 90-120 minutes. The interview will be audio recorded using a digital voice recorder to help with looking at the results. Once recorded, the interview will be transcribed (typed-up) and stored as a computer file.

What benefit will I gain from being involved in this study?

The sharing of your experiences will assist in the improvement of evaluation practice that would effectively support regional development planning.

Will I be identifiable by being involved in this study?

We do not need your name. Any identifying information will be removed, and your comments will not be linked directly to you. All information and results obtained in this study will be stored in a secure way, with access restricted to relevant researchers. However, although no identifying information will be published, due to the location of the interview, anonymity cannot be guaranteed.

Are there any risks or discomforts if I am involved?

The researcher anticipates few risks from your involvement in this study, however, given the nature of the project, some participants could experience emotional discomfort. If any emotional discomfort is experienced please contact Flinders University Counselling Service on (+618) 8201 2118 for support / counselling that may be accessed free of charge by all participants. If you have any concerns regarding anticipated or actual risks or discomforts, please raise them with the researcher.

How do I agree to participate?

Participation is voluntary. You may answer 'no comment' or refuse to answer any questions, and you are free to withdraw from the interview at any time without effect or consequences. A consent form accompanies this information sheet. If you agree to participate please read and sign the form and send it back to the researcher.

How will I receive feedback?

On project completion, outcomes of the project will be given to all participants via email / post / website.

Thank you for taking the time to read this information sheet, and we hope that you will accept our invitation to be involved.

This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project number 7820).

For more information regarding ethical approval of the project only, the Executive Officer of the Committee can be contacted by telephone on (+618) 8201 3116, by fax on (+618) 8201 2035, or by email to human.researchethics@flinders.edu.au

APPENDIX 5: INFORMATION SHEET (FOCUS GROUP)



Ms Dwi Ratih S. Esti
College of Business, Government and Law

Sturt Road
Bedford Park SA 5042

GPO Box 2100
Adelaide SA 5001

esti0005@flinders.edu.au

CRICOS Provider No. 00114A

INFORMATION SHEET (for *Focus Group Participants*)

Title: ‘Effectiveness of Evaluation Practice in Supporting Regional Development Planning’

Researcher(s):

Ms Dwi Ratih S. Esti
College of Business, Government and Law
Flinders University
Ph: +61 8 82012647
Position in Indonesia: staff in the Ministry of National Development Planning/BAPPENAS

Supervisor(s):

A/Prof Noore Siddiquee
College of Business, Government and Law
Flinders University
Ph: +61 8 82012302

A/Prof Gerry Redmond
College of Business, Government and Law
Flinders University
Ph: +61 8 82012699

Description of the study

This research will provide an in-depth overview of effectiveness of evaluation practice in supporting regional development planning in Indonesia using two case studies. This overview includes the perspectives of evaluation practitioners, intended users of evaluations (policy makers), local leaders, academicians, officials of NGOs, and the central bureau of statistics. This project is supported by Flinders University, College of Business, Government and Law.

Purpose of the study

This project aims to understand the influencing factors, challenges, and limitations of the effectiveness of evaluation practices in supporting regional development planning process.

What will I be asked to do?

You are invited to attend a focus group with a researcher who will ask you as part of a group a few questions about your views about the effectiveness of evaluation practices in supporting regional development planning. Participation is entirely voluntary. The focus group will take about 60-120 minutes. The discussion will be audio recorded using a digital voice recorder to help with looking at the results. Once recorded, the discussion will be transcribed (typed-up) and stored as a computer file.

What benefit will I gain from being involved in this study?

The sharing of your experiences will assist in the improvement of evaluation practice that would effectively support regional development planning.

Will I be identifiable by being involved in this study?

We do not need your name. Any identifying information will be removed, and your comments will not be linked directly to you. All information and results obtained in this study will be stored in a secure way, with access restricted to relevant researchers.

Are there any risks or discomforts if I am involved?

Other focus group participants may be able to identify your contributions even though they will not be directly attributed to you. The researcher will attempt to gain verbal agreement between all participants that they will maintain the anonymity of other participants and the confidentiality of the discussion. The researcher anticipates few risks from your involvement in this study, however, given the nature of the project, some participants could experience emotional discomfort. If any emotional discomfort is experienced please contact Flinders University Counselling Service on (+618) 8201 2118 for support / counselling that may be accessed free of charge by all participants. If you have any concerns regarding anticipated or actual risks or discomforts, please raise them with the researcher.

How do I agree to participate?

Participation is voluntary. You may answer 'no comment' or refuse to answer any questions, and you are free to withdraw from the focus group at any time without effect or consequences. A consent form accompanies this information sheet. If you agree to participate please read and sign the form and send it back to the researcher.

How will I receive feedback?

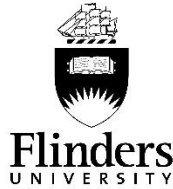
On project completion, outcomes of the project will be given to all participants via email / post / website.

Thank you for taking the time to read this information sheet, and we hope that you will accept our invitation to be involved.

This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project number 7820).

For more information regarding ethical approval of the project only, the Executive Officer of the Committee can be contacted by telephone on (+618) 8201 3116, by fax on (+618) 8201 2035, or by email to human.researchethics@flinders.edu.au

APPENDIX 6: CONSENT FORM (INTERVIEW)



**CONSENT FORM FOR PARTICIPATION IN RESEARCH
(by interview)**

**Effectiveness of Evaluation Practice
in Supporting Regional Development Planning**

I
being over the age of 18 years hereby consent to participate as requested in the interview for the research project on the Effectiveness of Evaluation Practice in Supporting Regional Development Planning.

1. I have read the information provided.
2. Details of procedures and any risks have been explained to my satisfaction.
3. I agree to audio recording of my information and participation.
4. I am aware that I should retain a copy of the Information Sheet and Consent Form for future reference.
5. I understand that:
 - I may not directly benefit from taking part in this research.
 - I am free to decline to participate in the interview and am free to decline to answer particular questions.
 - While the information gained in this study will be published as explained, I will not be identified, and individual information will remain confidential. However, although no identifying information will be published, due to the location of the interview, anonymity cannot be guaranteed.
 - I will be given the opportunity to review and edit my individual interview transcript.

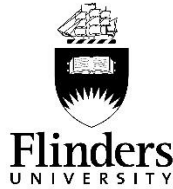
Participant's signature.....Date.....

I certify that I have explained the study to the volunteer and consider that she/he understands what is involved and freely consents to participation.

Researcher's name.....

Researcher's signature.....Date.....

APPENDIX 7: CONSENT FORM (FOCUS GROUP)



**CONSENT FORM FOR PARTICIPATION IN RESEARCH
(by focus group)**

**Effectiveness of Evaluation Practice
in Supporting Regional Development Planning**

I
being over the age of 18 years hereby consent to participate as requested in the focus group for the research project on the Effectiveness of Evaluation Practice in Supporting Regional Development Planning.

1. I have read the information provided.
2. Details of procedures and any risks have been explained to my satisfaction.
3. I agree to audio recording of my information and participation.
4. I am aware that I should retain a copy of the Information Sheet and Consent Form for future reference.
5. I understand that:
 - I may not directly benefit from taking part in this research.
 - I am free to withdraw from the focus group and am free to decline to answer particular questions but I will not be able to withdraw my data or ask that the audio recording be stopped.
 - While the information gained in this study will be published as explained, individual information will remain confidential. However, given the nature of focus group, participant anonymity cannot be guaranteed.

Participant's signature.....Date.....

I certify that I have explained the study to the volunteer and consider that she/he understands what is involved and freely consents to participation.

Researcher's name.....

Researcher's signature.....Date.....

APPENDIX 8: SEMI-STRUCTURED INTERVIEW GUIDE

Evaluation Practitioner: Officials in Evaluation Divisions in RDPAs

Introduce the purpose of the research, ensure that the participant has read and signed the consent form. Ask if he/she has any questions before getting started.

Interview number:

Working experience in the institution:

1. What is the main reason for undertaking evaluation practice?
2. The size of your institution and the unit for evaluation function in particular
 - a. How does the evaluation unit fit with the institutional structure (*reporting, responsibilities, relation to core business*)
 - b. How many people work on evaluation at your institution? What are their roles? What proportion of their time is spent on evaluation?
 - c. How sophisticated would you say members of your institution are in terms of their knowledge skills and ability to practice or oversee evaluation?
 - d. Has your institution initiated any evaluation capacity building activities in the past while? (on the job, workshops/training, education, etc)
3. How many evaluation practices do your institution undertake each year?
 - a. Is there any changes before and after decentralisation era?
 - b. When did your institution first begin undertaking evaluation?
 - c. The types of evaluation (*at what stages in planning process*)
 - d. The focus (*plan document/policy/program*)
 - e. The main purpose (*achievement of goals or target/satisfaction of clients/else*)
 - f. The policies or regulation which guide the evaluation practices
 - g. Are evaluations usually done in-house or are they contracted out (*subject to oversight*)?
4. Please describe for me a "typical" evaluation practice that you were involved in.
 - a. The reason
 - b. Indicators referred to and apply
 - c. Types and sources of data
 - d. Standard approach or methodology
 - Qualitative or quantitative or mixed (*familiar with CBA/utilisation-focused evaluation/logic model/theory of change/outcome mapping/outcome harvesting?*)
 - Reason in choosing the method/approach
 - Is the approach/method influenced by practice elsewhere
 - e. The level of support on resources such as logistics for undertaking evaluation
 - f. Results
 - Quality (*usefulness/comprehensiveness/reliability*)
 - How are the results communicated, and to whom
 - How are the results used? (*support for decisions? Learning about programs, organization, functions, etc.? Symbolic, persuasive and political uses?*)

- Is there any follow-up mechanism/method to ensure that evaluation results, specifically recommendations are used?
5. Have you ever experienced problems in conducting evaluation?
 - a. What are they? (*identify and explain the nature of the problem*)
 - b. How have you solved those problems?
 6. According to you, what constitutes an effective evaluation practice?
 7. What do you think about the present state of evaluation practice in your institution? (*quality, quantity, usefulness, effectiveness, scope*)
 8. In your opinion, what role evaluation practice has in regional development planning?
 9. In your opinion, what factors (contexts and mechanisms) facilitate evaluation practice in your institution specifically, and in your province regional development planning process in general? Which ones are more influential? (mention factors and their reasons)
 10. Conversely, in your opinion, what factors (contexts and mechanisms) impede evaluation practice? Which ones are more influential? (mention factors and their reasons)
 11. How do you know your evaluation practice is/has been successful?
 12. According to you, what would be the consequences of not evaluating your programs/plan documents?
 13. In your opinion, what are the pitfalls/limitations of current evaluation practice?
 14. Do you have any suggestions or ideas on how to improve the effectiveness of evaluation practices in your institution?

Is there anything else you would like to add?

Intended Users: Officials in Planning Divisions in RDPAs

Introduce the purpose of the research, ensure that the participant has read and signed the consent form. Ask if he/she has any questions before getting started.

Interview number:

Working experience in the institution:

1. How would you describe your institution's approach to decision making? (*Evidence-based? Democratic? Collaborative? etc?*)
2. What considerations shape the nature and direction of regional development plans?
3. In your opinion, what role evaluation practice has in regional development planning? (*importance, etc*)
4. How many evaluation results do your working unit utilise each year?
 - a. Is there any changes before and after decentralisation era?
 - b. When did your working unit first begin utilise evaluation results?
 - c. The types of evaluation (*at what stages in planning process*)
 - d. The focus (*plan document/policy/program*)
5. Please describe for me a "typical" planning process which utilise evaluation results that you were involved in.
 - a. The reason to utilise the evaluation results
 - b. Types and sources of data
 - c. Standard approach or methodology
 - Qualitative or quantitative or mixed
 - Reason in choosing the method/approach
 - Is the approach/method influenced by practice elsewhere
 - d. The level of support on resources such as logistics for utilising evaluation
 - e. Stakeholders
 - Who are they (including the evaluators)
 - What are their roles
 - f. Evaluation results
 - Quality level
 - Usefulness
 - How are the results obtained?
 - How are the results used? (*support for decisions? Learning about programs, organization, functions, etc.? Symbolic, persuasive and political uses?*)
6. Can you please indicate the proportion of evaluation results that are utilised for future policies or development plans?
7. Have you ever experienced problems in utilising evaluation results?
 - a. What are they? (*identify and explain the nature of the problem*)
 - b. How have you solved those problems?
8. According to you, what constitutes an effective evaluation practice?
9. What do you think about the present state of evaluation practice in your institution? (*quality, quantity, usefulness, effectiveness, scope*)

10. In your opinion, what factors (contexts and mechanisms) facilitate evaluation practice in your institution specifically, and in regional development planning process in your province in general? Which ones are more influential? (mention factors and their reasons)
11. Conversely, in your opinion, what factors (contexts and mechanisms) impede evaluation practice? Which ones are more influential? (mention factors and their reasons)
12. How do you know the regional development plans/policies/programs are/have been successful if they are not evaluated properly?
13. According to you, what would be the consequences of not evaluating programs/plan documents at the regional/local level?
14. In your opinion, what are the pitfalls/limitations of current evaluation practice?
15. Do you have any suggestions or ideas on how to improve the effectiveness of evaluation practices in supporting regional development planning in your province?

Is there anything else you would like to add?

Officials in the Ministry of National Development Planning and the Ministry of Home Affairs

Introduce the purpose of the research, ensure that the participant has read and signed the consent form. Ask if he/she has any questions before getting started.

Interview number:

Working experience in the institution:

1. How would you describe your institution's approach to decision making? (*Evidence-based? Democratic? Collaborative? etc?*)
2. The size of your institution and the unit for evaluation function in particular
 - a. How does the evaluation unit fit with the institutional structure (*reporting, responsibilities, relation to core business*)
 - b. How many people work on evaluation at your institution? What are their roles? What proportion of their time is spent on evaluation?
3. How often your institution's plans/policies/programs are informed by evaluation from the RDPAs and from the provincial governments in general?
4. Do your institution require evaluation reports regularly from RDPAs specifically and from the provincial governments in general?
 - a. How frequent?
 - b. What do your institution do to those reports?
 - c. When did your institution first begin require this type of reports?
 - d. What are the underlying policies or regulation?
5. Do your institution guide evaluation practice in regional development planning?
 - a. Is there any changes before and after decentralisation era?
 - b. When did your institution first begin producing regulations/guidelines on evaluation practice in regional development planning?
 - c. The types of evaluation (*at what stages in planning process*) which stated in the regulations
 - d. The focus (plan document/policy/program)
 - e. The main purpose (*achievement of goals or target/satisfaction of clients/else*)
 - f. The frequency
6. Please describe for me a "typical" planning process which utilise evaluation results from RDPAs that you were involved in.
 - a. The reason to utilise the evaluation results
 - b. Types and sources of data
 - c. Standard approach or methodology
 - Qualitative or quantitative or mixed
 - Reason in choosing the method/approach
 - Is the approach/method influenced by practice elsewhere
 - d. The level of support on resources such as logistics for utilising evaluation
 - e. Stakeholders
 - Who are they (including the evaluators)
 - What are their roles

- f. Evaluation results
- Quality level
 - Usefulness
 - How are the results obtained?
 - How are the results used? (*support for decisions? Learning about programs, organization, functions, etc.? Symbolic, persuasive and political uses?*)
7. Have you ever experienced problems in utilising evaluation results from provincial governments?
 - a. What are they? (*identify and explain the nature of the problem*)
 - b. How have you solved those problems?
 8. According to you, what constitutes an effective evaluation practice?
 9. What do you think about the present state of evaluation practices in supporting regional development planning process? (*quality, quantity, usefulness, effectiveness, scope*)
 10. In your opinion, what role evaluation practice has in regional development planning at the moment? What role it should have based on your perceptions?
 11. In your opinion, what factors (contexts and mechanisms) facilitate evaluation practice in supporting regional development planning process? Which ones are more influential? (mention factors and their reasons)
 12. Conversely, in your opinion, what factors (contexts and mechanisms) impede evaluation practice? Which ones are more influential? (mention factors and their reasons)
 13. How do you know the regional development plans/policies/programs are/have been successful if they are not evaluated properly?
 14. According to you, what would be the consequences of not evaluating programs/plan documents at the regional/local level?
 15. In your opinion, what are the pitfalls/limitations of current evaluation practice in supporting regional development planning at the provincial level?
 16. Do you have any suggestions or ideas on how to improve the effectiveness of evaluation practices in supporting regional development planning at the provincial level?

Is there anything else you would like to add?

Academics

Introduce the purpose of the research, ensure that the participant has read and signed the consent form. Ask if he/she has any questions before getting started.

Interview number:

Working experience in the institution:

1. Do you have any experience of working with RDPA on matters of evaluation practice? If yes, please share your experience (*the main purpose, type of evaluation, the focus, stakeholders, problems /challenges*).
2. According to you, what constitutes an effective evaluation practice?
3. In your opinion, what role evaluation practice has in regional development planning at the moment? What role it should have based on theories or your perceptions?
4. According to you, are evaluation results utilised in the policy process?
5. What do you think about the present state of evaluation practice in supporting regional development planning at the provincial level? (quality, quantity, usefulness, effectiveness, scope)
6. What do you think about the existing commitment and supports for evaluation in supporting regional development planning at the provincial level?
7. What do you think about the capacity and skills of officials in RDPAs to undertake evaluation?
8. In your opinion, are policies and regional development plans in provincial governments evidence-based?
9. In your opinion, what factors (contexts and mechanisms) facilitate evaluation practice in supporting regional development planning process? Which ones are more influential? (*mention factors and their reasons*)
10. Conversely, in your opinion, what factors (contexts and mechanisms) impede evaluation practice? Which ones are more influential? (*mention factors and their reasons*)
11. According to you, what would be the consequences of not evaluating programs/plan documents at the regional/local level?
12. In your opinion, what are the pitfalls/limitations of current evaluation practice in supporting regional development planning at the provincial level?
13. Do you have any suggestions or ideas on how to improve the effectiveness of evaluation practices in supporting regional development planning at the provincial level?

Is there anything else you would like to add?

Local Leaders

Introduce the purpose of the research, ensure that the participant has read and signed the consent form. Ask if he/she has any questions before getting started.

Interview number:

Working experience in the institution:

1. How would you describe your approach to decision making? (*Evidence-based? Democratic? Collaborative? etc?*)
2. What considerations are used to shape the nature and direction in your policymaking process?
 - a. To what extent it is informed by evaluations?
 - b. What mechanisms/processes are in place to incorporate evaluation results?
3. Do you have any experience in utilising evaluation results? If yes, please share your experience (*the main purpose, type of evaluation, the focus, stakeholders, problems /challenges*).
4. According to you, what constitutes an effective evaluation practice?
5. In your opinion, what role evaluation practice has in regional development planning at the moment? What role it should have based on your perceptions?
6. What do you think about the present state of evaluation practice in supporting regional development planning process in your province? (*quality, quantity, usefulness, effectiveness, scope*)
7. What do you think about the existing commitment and supports for evaluation in regional development planning process in your province?
8. In your opinion, are policies and regional development plans in your province evidence-based?
9. In your opinion, what factors (contexts and mechanisms) facilitate evaluation practice in supporting regional development planning at the provincial level? Which ones are more influential? (*mention factors and their reasons*)
10. Conversely, in your opinion, what factors (contexts and mechanisms) impede evaluation practice? Which ones are more influential? (*mention factors and their reasons*)
11. According to you, what would be the consequences of not evaluating programs/plan documents at the regional/local level?
12. In your opinion, what are the pitfalls/limitations of current evaluation practice in supporting regional development planning at the provincial level?
13. Do you have any suggestions or ideas on how to improve the effectiveness of evaluation practices in supporting regional development planning at the provincial level?

Is there anything else you would like to add?

NGO and Statistics Centre/Bureau

Introduce the purpose of the research, ensure that the participant has read and signed the consent form. Ask if he/she has any questions before getting started.

Interview number:

Working experience in the organisation:

1. Do you have any experience in utilising evaluation results from RDPA or cooperating with RDPA related to evaluation practice? If yes, please share your experience (*the main purpose, type of evaluation, the focus, stakeholders, problems /challenges*).
2. According to you, what constitutes an effective evaluation practice?
3. In your opinion, what role evaluation practice has in regional development planning at the moment? What role it should have based on your perceptions?
4. According to you, what prevents utilisation of evaluation results in policy process?
5. What do you think about the present state of evaluation practice in supporting regional development planning process in your province? (quality, quantity, usefulness, effectiveness, scope)
6. What do you think about the existing commitment and supports for evaluation in supporting regional development planning at the provincial level?
7. In your opinion, are policies and regional development plans in your province evidence-based?
8. In your opinion, what factors (contexts and mechanisms) facilitate evaluation practice in supporting regional development planning at the provincial level? Which ones are more influential? (*mention factors and their reasons*)
9. Conversely, in your opinion, what factors (contexts and mechanisms) impede evaluation practice? Which ones are more influential? (*mention factors and their reasons*)
10. According to you, what would be the consequences of not evaluating programs/plan documents at the regional/local level?
11. In your opinion, what are the pitfalls/limitations of current evaluation practices in supporting regional development planning at the provincial level?
12. Do you have any suggestions or ideas on how to improve the effectiveness of evaluation practices in supporting regional development planning at the provincial level?

Is there anything else you would like to add?

APPENDIX 9: FOCUS GROUP GUIDE

1. Consent Process

Consent forms for focus group participants are completed in advance by all those seeking to participate.

2. Introduction

a. Welcome

Introduce the facilitator, and send the Sign-In Sheet around to the group while the facilitator are introducing the focus group.

Review the following:

- Introduce the research and the researcher (facilitator)
 - What will be done with this information
 - Why the participants asked to participate
- b. Explanation of the process
- Ask the group if anyone has participated in a focus group before. Explain that focus groups are being used more and more often in social research.

About focus groups

- We learn from you. No right or wrong answers, only differing points of view
- Not trying to achieve consensus, it's about gathering information
- No virtue in long list, looking for priorities

Logistics

- Focus group will last about one hour
- Feel free to move around
- Location of the bathroom and exit door
- Help yourself to refreshments

3. Ground Rules

- Everyone should participate
- Information provided in the focus group must be kept confidential
- Stay with the group and please don't have side conversations
- Turn off cell phones if possible. If you cannot and if you must respond to a call, please do so as quietly as possible and rejoin as quickly as you can.
- We're tape recording, one person speaking at a time
- You don't need to agree with others, but you must listen respectfully as others share their views

Questions:

1. In your opinion, what role evaluation has, if any, in regional development planning?
 - a. To what extent regional development plans are informed by evaluations?
 - b. What mechanisms/processes are in place to incorporate evaluation results?
2. What considerations shape the nature and direction of regional development plans?
3. To what extent they are consistent with the existing guidelines?
4. What do you think about the present state of evaluation practices in supporting regional development planning in your province? (*present highlights from initial interview data*)
5. In your opinion, how conducive is the context for evaluation practice in your province?
6. According to you, what constitutes an effective evaluation practice?
7. What factors (elements of context and mechanisms) can facilitate evaluation practice in supporting policymaking, particularly in regional development planning? Which ones are more influential? (*present highlights from initial interview data*)
8. What factors (elements of context and mechanisms) can impede evaluation practice in supporting policymaking, particularly in regional development planning? Which ones are more influential? (*present highlights from initial interview data*)
9. How can evaluation practices be improved to support regional development planning?