

**Under a Top-Down Rubric Policy: The Perceptions and Actualisations of
Assessment for Learning and Rubric in Higher Education in Hong Kong**

Doris Yin Kei Chong

September 2019

This thesis is submitted in partial fulfilment of the requirements for the degree
of Doctor of Philosophy.

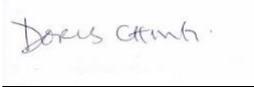
Department of Educational Research,

Lancaster University, UK.

Declaration

This thesis results entirely from my own work and has not been offered previously for any other degree or diploma.

The word-length conforms to the permitted maximum.

Signature:  _____

Abstract

Assessment for Learning has been a popular topic in higher education. However, it appears challenging for teachers and students in Hong Kong to adopt this concept. Given the opportunity of a top-down rubric policy in one university in Hong Kong, this thesis explores the perceptions and actualisations of Assessment for Learning and rubrics of students, educators and managers in this context. It offers a model of a holistic assessment policy that refocuses on the stakeholders, content and processes.

The study is rooted in an interpretivist paradigm and qualitative data are generated through two focus groups of student participants and fourteen semi-structured interviews of educator and manager participants. This research demonstrates that measurement is the major function of assessment and rubrics in this context due to the Confucian-heritage culture, the dominant assessment discourse and suboptimal assessment literacy, while practical knowledge of learning features exist as a result of the participants' applied-discipline backgrounds. This research delves into the barriers and opportunities of actualising Assessment for Learning and rubrics under a mandatory policy. Interestingly, pragmatic and contextual issues are key challenges that create tensions in the actualisation, rather than the concept itself or the disagreement on having a policy. This research somewhat concurs with suggestions from the contemporary literature that assessment practice change is complex and should consider institutional and cultural uniqueness.

Based on the familiarity and acceptance of features of Assessment for Learning because of the applied-discipline background, the thesis therefore offers a context-specific model that engages stakeholders in various processes. It also includes the policy content necessary for improving the enactment of this policy. Other institutions that share similar cultures and backgrounds can benefit from the model when initiating an assessment policy.

Table of Contents

Abstract	i
Acknowledgements.....	xi
Publications derived from work on the Doctoral Programme.....	xii
List of Figures and Tables.....	xiii
Chapter 1: Introduction.....	1
1.1 Background and Rationale of the Study	1
1.2 Personal Interest	4
1.3 Research Aim and Objectives	5
1.4 Research Questions.....	6
1.5 Theoretical Underpinning.....	7
1.6 Summary of Methodology and Methods.....	12
1.7 Contribution to Knowledge.....	13
1.8 Overview of the Thesis Structure	13
Chapter 2: Understanding Assessment Practice in Higher Education ...	18
2.1 Functions and Purposes of Assessment in Higher Education	18
2.1.1 Assessment for Certification and for Quality Assurance	19

2.1.2 Assessment for Learning	20
2.1.3 Assessment for Future/Lifelong Learning – Sustainable Assessment	22
2.1.4 Summary of Section and Implications	23
2.2 The Current Picture of Assessment Practice in Higher Education and the Contributing Factors to this Picture	24
2.2.1 Heavy Focus on Assessment for Certification	26
2.2.2 Examination-Oriented Culture in Hong Kong	29
2.2.3 Insufficient Assessment Literacy	31
2.2.4 Resources, Regulations/Policies and Cultures set Work Priorities.....	33
2.2.5 Summary of Section and Implications	37
2.3 Summary of Chapter and Implications for this Study	38
Chapter 3: The Use of Rubrics in Higher Education.....	39
3.1 Functions and Purposes of Assessment Rubrics.....	40
3.2 Merits of Using Rubrics in Assessment Practice	41
3.2.1 Benefits of Using Rubrics as Grading Tools.....	41
3.2.2 Benefits of Using Rubrics as Learning Tools – Assessment for Learning.....	44
3.2.3 Psycho-social Benefits of Rubrics	47
3.2.4 Summary of Section and Implications	48

3.3 Criticisms of Using Rubrics in Assessment Practice	50
3.3.1 Rubric Development Issues.....	50
3.3.2 Rubric Implementation and Utilisation Issues	52
3.3.3 Summary of Section and Implication	54
3.4 Innovative Methods to Optimise Actualisation of Rubrics.....	55
3.4.1 Innovative Examples	56
3.4.2 Summary of Section and Implication	59
3.5 Summary of Chapter and Implications for this Study	60
Chapter 4: The Current Situation of Assessment Policy in Higher Education.....	62
4.1 Intentions of Assessment Policies in Higher Education.....	62
4.2 Reasons Limiting Enactment of Assessment Policies in Higher Education ..	64
4.3 Recommendations for Holistic Assessment Policies.....	67
4.4 Summary of Chapter and Implications for the Study	70
Chapter 5: Methodology and Methods	72
5.1 The Researcher’s Position and Ethical Considerations.....	73
5.1.1 Position of the Researcher	73
5.1.2 Ethical Considerations.....	75

5.2 The Research Paradigm	78
5.3 Methodology	80
5.4 Methods	83
5.4.1 Sampling	83
5.4.2 Data Collection	85
5.4.3 Data Analysis.....	89
5.5 Limitations of Methodology and Methods	90
5.6 Summary of the Chapter	92
Chapter 6: Results of Assessment for Learning	93
6.1 Demographics of the Participants	95
6.2 Theme 1 – Elements in Assessment Facilitating Learning	97
6.2.1 Authenticity	99
6.2.2 Feedback.....	100
6.2.3 Reflection.....	102
6.2.4 Curriculum Mapping.....	103
6.2.5 Summary of Theme 1	104
6.3 Theme 2 – Burden and Mismatch in Current Assessment Practice	105

6.3.1 Burden – Examination-focused and Over-assessment.....	105
6.3.2 Mismatch of Assessment and Objectives.....	109
6.3.3 Summary of Theme 2.....	112
6.4 Theme 3 – Ideal Actualisations of Assessment for Learning.....	113
6.4.1 Re-alignment of Learning Objectives and Assessment Tasks.....	113
6.4.2 Involvement of Continuous and Authentic Assessment.....	115
6.4.3 Promotion of Sustainable Assessment.....	117
6.4.4 Summary of Theme 3.....	118
6.5 Summary of the Chapter.....	119
Chapter 7: Results of Actualisation of Rubrics in this Policy Context .	120
7.1 Theme 4 - Potentials of Rubrics for Learning.....	122
7.2 Theme 5 - Barriers of Rubric Actualisation.....	124
7.2.1 Barriers from Other Stakeholders.....	124
7.2.2 Barriers from the Nature of Rubrics.....	127
7.2.3 Barriers to the Development of Rubrics.....	128
7.2.4 Barriers from the University’s Culture and Support.....	130
7.2.5 Summary of Theme 5.....	132

7.3 Theme 6 - Optimisation of Rubrics in Assessment Practice	133
7.3.1 Increasing Guidance and Flexibility by Teachers	133
7.3.2 Reinventing Rubrics by Students.....	135
7.3.3 Understanding and Sharing of Good Practice by Management	136
7.3.4 Summary of Theme 6	137
7.4 Theme 7 - Promoting Holistic Assessment Practice through Policy	138
7.5 Summary of the Chapter	141
Chapter 8: A Discussion on the Challenges and Potential of Assessment for Learning and Rubrics under the Policy Situation	143
8.1 Implicit Understanding of Assessment for Learning.....	146
8.1.1 The Dominant Assessment Practice and Traditional Culture	147
8.1.2 Questionable Level of Assessment Literacy	149
8.1.3 Contextual Characteristics	151
8.2 Barriers and Opportunities of Assessment for Learning – Explanations from Traditional and Contextual Perspectives	153
8.2.1 Constituents and Implications of Barriers.....	153
8.2.2 Contributors and Implications of Opportunities.....	156

8.3 Summary and Implications of Conceptions, Barriers and Opportunities of Assessment for Learning.....	159
8.4 Feelings of Powerlessness in Actualising Rubrics in the Context of a Top-Down Policy.....	160
8.4.1 Potentials of Rubrics for Learning Despite Stakeholders’ Experience Level.....	161
8.4.2 Constituents and Implications of Rubric Actualisation Barriers	163
8.4.3 Contributors and Implications for Ideal Actualisation of Rubrics	167
8.4.4 Summary and Implications of Rubric Actualisation in the Research Context	171
8.5 Enhancing Policy Success through a Holistic and Deliberate Approach.....	172
8.6 Summary of Discussions	175
Chapter 9: Contributions to Knowledge and Conclusions	177
9.1 Conclusion of the Findings	177
9.2 Contribution to Knowledge.....	179
9.2.1 Contribution to the Literature on Assessment for Learning, Rubrics and Assessment Policy in Higher Education.....	180
9.2.2 Contribution to the Institution.....	181
9.2.3 Application to Other Local and/or Regional Institutions.....	184
9.3 Study Limitations.....	187

9.4 Recommendations for Future Studies	189
9.5 Conclusion	191
References	194
Appendix 1 - The Rubric Policy Paper	209
Appendix 2 - Interview Questions	212

Acknowledgements

I would like to express my sincere gratitude to many individuals who have helped me through this doctoral journey.

I have been extremely fortunate to have Dr Jan McArthur as my thesis supervisor. She is always supportive, constructive, responsive and kind. She helped and lifted me through occasions when I was feeling frustrated and demotivated with the process. I am truly indebted to her supervision and patience.

Thank you to all the participants who took part in this research. Without their willingness to share their thoughts, I would not have been able to gather these unique findings and contribute to knowledge of the field. They are all inspirational! I am also thankful to my friends (especially Hennie, Veronika, Barbara, Arnold and Joy), cohort 3 classmates (Kay and Alison), colleagues, students and instructors who have motivated me in various research and teaching areas. I could not have thought of these ideas without their stimulation.

Lastly, I am grateful to have a family who always support me in whatever I wish to pursue. Although they do not understand why I want to undertake yet another degree, they never question or judge, which leaves me room to focus. Their love is unconditional.

Publications derived from work on the Doctoral Programme

The following publication has arisen from the work of part one of this PhD programme:

Chong, D. Y. (2019). Benefits and challenges with gamified multi-media physiotherapy case studies: A mixed method study. *Archives of Physiotherapy, 9, 7.*

List of Figures and Tables

List of Figures

Figure 1.1 Framework of the present study (informed by theoretical resources)

Figure 1.2 Use of the theoretical concept of Assessment for Learning in this research

Figure 6.1 Illustration of themes and sub-themes for research questions 1 and 2

Figure 7.1 Illustration of themes and sub-themes for research questions 3 and 4

Figure 9.1 Suggested model of a holistic assessment policy initiative and strategy

List of Tables

Table 6.1 Demographic information of educator and management participants

Table 8.1 Summary of findings and mappings of bodies of literature

Chapter 1: Introduction

1.1 Background and Rationale of the Study

Assessment plays a significant role in higher education. Results from assessment directly dictate the career prospects of students because future employers often rely on higher education assessment results or certification to decide the employability of a graduate (Boud & Falchikov, 2007). As a result, higher education assessment determines and controls a student's future life.

The assessment system in primary and secondary schools in Hong Kong has been characterised as heavily examination-oriented and assessment results are accordingly mainly derived from examinations (Berry, 2011; Brown, Kennedy, Fok, Chan, & Yu, 2009). This scenario is thought to be influenced by the Confucian-heritage culture (Carless, 2011). The situation is similar in higher education institutions where the term 'assessment' almost exclusively refers to examinations or assignments (Ewell, 2006). Despite the growing popularity of formative assessment, it appears challenging for teachers and students in Hong Kong to change from the traditional concept of examination and relate assessment with learning (Berry, 2011; Brown, Hui, Yu, & Kennedy, 2011; Brown & Wang, 2013).

Research to promote formative assessment or Assessment for Learning is not "in paucity" globally and locally. In the Hong Kong higher education context, non-traditional assessment task designs and activities, such as authentic assessment and feedback mechanism, have been

advocated and shown to be effective in embedding and facilitating learning within the framework of assessment (Carless, 2015a). Nonetheless, it is not very well-known how the various stakeholders in higher education in Hong Kong perceive Assessment for Learning. Furthermore, the assessment rubric is one of the tools situated in the context of Assessment for Learning and previous research has demonstrated its ability to mediate learning in assessment (Panadero & Jonsson, 2013). However, not all rubric implementations have been deemed effective (Bennett, 2016; Carless, 2015a). It is therefore crucial to explore the factors underpinning or hindering the effective use of rubrics under the umbrella concept of Assessment for Learning.

Looking at some of the successful rubric implementation examples (Broadbent, Panadero, & Boud, 2018; Fraile, Panadero, & Pardo, 2017; Jones, Allen, Dunn, & Brooker, 2017), it can be seen that apart from practical factors, these initiatives are primarily at the micro-level, which are voluntary and individual- or team-based. It is reasonable to assume that teachers implementing such initiatives possess a fair amount of knowledge of the functions and designs of rubrics or at least a belief in the merits of rubrics.

In addition, these examples are from the Western context, where the Confucian-heritage cultural impact is of lesser concern. In Hong Kong, only limited examples of rubric utilisation are identified in the literature. Carless (2015a) demonstrates a few examples drawn from award-winning teachers but he also criticises how rubrics are actualised in these examples. This may

imply that knowledge of rubric actualisation in Hong Kong is still at its infancy stage and is worthwhile investigating.

Given the current contextual knowledge of Assessment for Learning and rubric from the literature, and the opportunity of a top-down assessment policy (Appendix 1) mandating the use of rubrics in major assessment tasks in one higher education institution in Hong Kong, this study is curious to explore the picture of rubric actualisation and the conception of Assessment for Learning in this setting. The study context is unique because

- rubric utilisation is under a non-voluntary initiative;
- teachers may not possess even fundamental knowledge on rubric design and implementation;
- assessment culture is under the deeply-rooted examination-oriented system where relating assessment and learning may be counter-intuitive; and
- most academic departments are regarded as applied disciplines under the historical position as part of a technical school.

By studying Assessment for Learning and rubric utilisation in this cultural and policy context, the results shed light on the barriers and opportunities when adopting a good assessment practice. This study also investigates how this mandatory policy is interpreted and executed. The literature on assessment policy is very limited at present (Boud, 2007; Carless, 2017); findings on policy interpretation and actualisation will provide

suggestions for refinement of present implementation and directions for future assessment policy initiatives.

1.2 Personal Interest

My personal interest in Assessment for Learning stems from reading about the Learning-Oriented Assessment (LOA) framework proposed by Carless (2007), while I was preparing a proposal for a teaching development project on rubrics. The LOA framework (Carless, 2007, 2015b) provides three conceptualised elements to guide educators to focus on the learning aspect of assessment:

- Using assessment tasks as learning tasks.
- Including students and peers in the assessment process.
- Providing timely and formative feedback.

My personal beliefs and views on assessment align closely with the framework, especially the concept of using assessment tasks as learning tasks. I believe this would not only promote learning but also minimize students' stress and anxiety, as well as tensions between educators and students because of discrepancies in assessment outcomes. These benefits in turn may promote learning engagement, an important factor for life-long learning.

Looking at the three elements in the LOA framework, together with further readings in the umbrella concept of Assessment for Learning (Carless, 2017; McDowell & Sambell, 2014; Sambell, Brown, & Graham, 2017), it

appears that assessment rubrics are a tool to actualise the concept. The opportunity to investigate Assessment for Learning and rubrics arose when the university in which I was working launched a rubric policy in 2016, mandating the use of assessment rubrics for major assignments worth 20% or more of the subject grade for subjects with examinations, and 30% or more for subjects without examinations (refer to Appendix 1 for the policy).

In addition, I obtained a teaching development grant (TDG) funded by the University Grants Committee of Hong Kong in 2017 to redesign an assessment component in one of the subjects I was teaching . This TDG project redesigned the assessment task based on what is proposed in the LOA framework and used multi-stage rubrics to actualise the elements on self-/peer-assessment and feedforward feedback. Hence, it was timely to explore the perceptions and knowledge of various stakeholders on the use of rubrics and the umbrella concept of Assessment for Learning. Since the university-wide implementation of rubrics came from a mandatory policy, I was also interested in understanding the policy implications in practice.

1.3 Research Aim and Objectives

In light of the above descriptions of gaps in the existing Assessment for Learning and rubric literature, my personal research interest and the unique contextual situation, the overall aim of this study is to investigate the perceptions and actualisations of Assessment for Learning and rubrics of students, educators and middle management under a mandatory top-down

policy in one university in Hong Kong. The specific objectives of this research are to:

- critically review relevant literature on assessment practice, rubrics and assessment policy in higher education both locally and globally and locate research gaps in these areas;
- conduct a comprehensive study that investigates perceptions and actualisations of Assessment for Learning and rubrics in light of a top-down and mandatory rubric policy;
- critically analyse participants' perceptions and experiences of rubrics under the concept (or lack of concept) of Assessment for Learning and the influence of policy; and
- reflect on policy interpretation unique to this context and provide suggestions for policy refinement for this context and the implications for other higher education institutes in Hong Kong and Confucian-influenced countries.

1.4 Research Questions

The specific research questions for addressing the aim and objectives of this study are:

- RQ1. What are the perceptions and understanding of Assessment for Learning in students, educators and management?
- RQ2. How do students, educators and management identify barriers and opportunities for the adoption of Assessment for Learning?

RQ3. Can rubrics actualise Assessment for Learning in the eyes of students, educators and management? Why or why not?

RQ4. What are the policy implications according to the experience of educators and management in this policy initiative?

RQ1 and RQ2 explore the background knowledge of participants' understanding of assessment practice in general with a particular focus on Assessment for Learning. It is important to gather this information because participants' conceptions of assessment practice are likely to affect their perceptions and actualisations of rubrics. If participants lack knowledge of Assessment for Learning, they are less likely to associate rubrics with learning, which is what RQ3 explores. RQ3 also investigates the picture of rubric actualisation under this policy initiative and identifies its potential and barriers as such. RQ4 investigates the interpretations and adoption of the rubric policy and provides insights on policy implications. Together, these questions allow this research to address the objectives of Assessment for Learning and the perceptions and actualisation of rubrics in light of the policy, and provide implications for policy refinement and enactment.

1.5 Theoretical Underpinning

This study does not employ a single theory in the research process. Rather, it is grounded on a set of definitions and ideas to inform the research questions, the interview questions, the positions of the data analysis, and the directions of discussion. After gathering an initial knowledge of assessment from textbooks and background literature in the higher education assessment field,

the following bodies of literature are used as the theoretical resources of this study:

- Assessment:
 - Functions/Typologies of Assessment: Assessment for Certification and summative assessment; Assessment for Learning and formative assessment; Assessment for Lifelong Learning and sustainable assessment
 - Confucian-influenced assessment culture
 - Resistance to change or assessment reform
- Rubrics:
 - Benefits and disadvantages of rubrics
 - Examples of rubric design, implementation or utilisation
- Assessment policy:
 - Assessment policy at macro-, meso- and micro-level

Details of how these bodies of literature inform this present study are available in Chapters 2 to 4. It is important to frame theoretical resources because they set the boundary of this research and inform what this research is and is not. Without this boundary, the focus of this research may deviate from the research questions. In other words, I used relevant bodies of literature to form my theoretical framework of the present study. The framework is represented as follows:

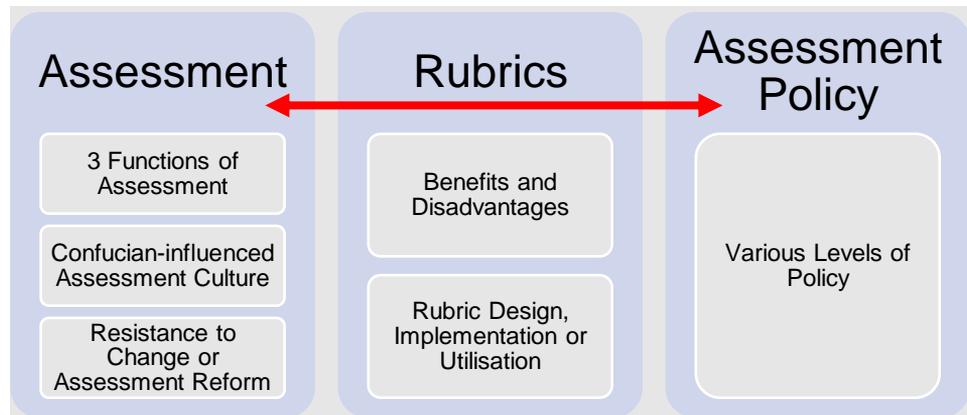


Figure 1.1 Framework of the present study (informed by theoretical resources)

In particular, the concept Assessment for Learning (Boud, 2007; Carless, 2017; Sambell, McDowell, & Montgomery, 2013) under the functions of assessment in the framework has a specific purpose to this study. A detailed description of Assessment for Learning is reviewed in Chapter 2 and is not repeated here. The reason behind using Assessment for Learning as the theoretical underpinning of the present study is that it is a well-studied and evidence-based concept in the field of higher education assessment. It is also the concept in which rubric is situated. The interconnection between Assessment for Learning and rubrics is significant and rubrics are a component within this umbrella concept. In this study context, the rubric policy may also dictate how stakeholders view and practise rubrics and assessment. As such, the three components of this research – assessment, rubrics and assessment policy – are related and likely to exert influence on each other, as depicted by the arrow in Figure 1.1 above.

It is worthwhile to reiterate the boundary of this research which is to explore if rubrics can actualise the concept of Assessment for Learning under this policy situation. While findings on rubric and rubric policy implementation are expected and they may contribute to understanding the relationship between rubrics and Assessment for Learning in this research context, they are not the primary interest of this research. It is also deemed impossible to keep this thesis within the required length if analysing implementation factors is included. Hence, literature on implementation framework is not selected as an area of this research's theoretical underpinning.

Specifically, this study is firmly situated in the practical and theoretical framework of Assessment for Learning. This framework carries the following theory functions (Trowler, 2016) in the present study and is illustrated in Figure 1.2 below:

- It provides definitions and distinguishes various functions of assessment. This fits into the *classification* function of the theory. Assessment for Learning entails a number of features and strategies with a common theme of putting learning at the centre of any assessment. Any assessment tasks or activities that deviate from a focus on learning are classified as other functions in the context of this study.
- It is used to understand and explain participants' perceptions and practices. This is the *explanation* function of theory and works hand-in-hand with the *classification* function. While Assessment for Learning distinguishes assessment activities, it also provides further

explanations of factors needed to adopt this practice. These factors include but are not limited to personal beliefs, backgrounds and experiences. Explanations of participants' findings are therefore grounded in the definition and features of Assessment for Learning.

- It interprets the association between participants' assessment practices and rubric perceptions/actualisations. This describes the *depiction* purpose of theory. As stated previously, rubrics are located within the bigger concept of Assessment for Learning; therefore, it is reasonable to assume that participants' beliefs and experiences with assessment practice influence how they perceive and actualise rubrics. That is, if participants are foreign to the concept of Assessment for Learning, their actualisation of rubrics may lack the learning purpose. On the contrary, if Assessment for Learning is a well-known concept in the study context, a different rubric adoption picture may surface.
- It guides interview questions and informs policy implications based on the findings of the study. *Guidance* is another use of theory and in this function, Assessment for Learning provides a framework of concepts where interview questions are set and the implications of findings are formulated. In particular, policy implications are guided by the definition and features of Assessment for Learning.

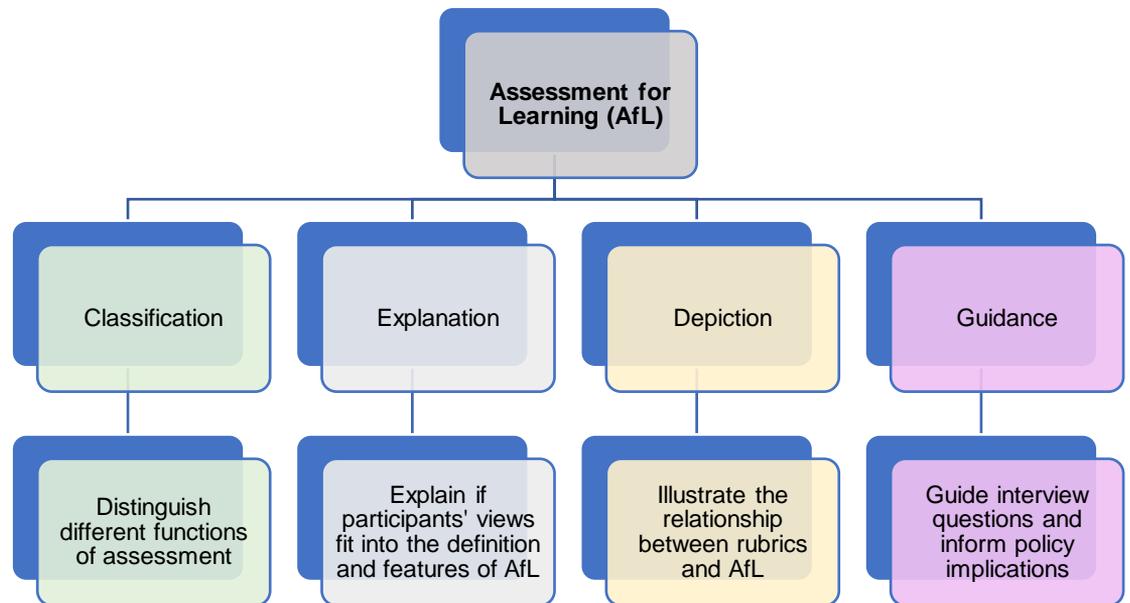


Figure 1.2 Use of the theoretical concept of Assessment for Learning in this research

1.6 Summary of Methodology and Methods

The present study adopts a relativist ontology and an interpretivism paradigm to explore participants' perceptions and experiences of Assessment for Learning and rubric utilisation under a mandatory top-down rubric policy situation in one university in Hong Kong. It investigates in depth a specific phenomenon unique to the research context and the participants (students, teachers and management) all experienced this phenomenon. As such, this study employs qualitative phenomenology as the research methodology. Focus groups and semi-structured interviews are used as the methods to collect qualitative data. Both of these methods allow participants to verbally discuss and give meaning to their perceptions and experiences, thereby enabling the extraction of rich information.

Interview data are audio-recorded, transcribed and systematically analysed to draw meaning units, codes, categories then finally themes. The themes serve as explicit answers for the research questions. Although the methodology and methods utilised in the present study are suitable to gather in-depth meaning of participants' lived experiences, they are not without limitations, such as subjectivity and bias. Chapter 5 discusses the research process in greater detail and provides a means for enhancing the credibility of findings and interpretations.

1.7 Contribution to Knowledge

The results of this research contribute to knowledge for assessment practice and policy initiative and are discussed in Chapter 9. A holistic assessment policy model (Figure 9.1) is derived based on unfolding the experiences and perceptions of the participants on Assessment for Learning and rubrics under a mandatory policy. This model suggests important factors, stakeholders and processes necessary for making an assessment policy more optimal. It does not only inform the university where this research is conducted with a refined practice but can be lifted to similar institutions locally and in the Confucian-influenced region. Contextual specificity is discussed for adopting this model.

1.8 Overview of the Thesis Structure

This thesis consists of nine chapters. This chapter (Chapter 1) provides an overview of the thesis which includes the background and rationale of the study, my personal interest, the research aim and objectives, the research questions, the theoretical underpinnings, a summary of the research

methodology and methods, its contribution to knowledge as well as an overview of the structure of this thesis. This chapter serves as the introduction and leads readers into the main sections of the thesis.

Chapters 2 to 4 are the literature review chapters. Specifically, Chapter 2 examines relevant literature on assessment in higher education. The topics include assessment functions/purposes and issues with current assessment practice. Under the section on assessment functions/purposes, details of assessment typologies – Assessment for Certification or summative assessment, Assessment for Learning or formative assessment and Assessment for Lifelong Learning or sustainable assessment – are reviewed. The section on issues with current assessment practice critically looks at the heavy focus on Assessment for Certification and the reasons behind this dominant function, such as the cultural influence in Hong Kong, the lack of assessment literacy and organisational regulation/culture. This chapter serves to provide the current picture of assessment practice in higher education both locally and globally and to identify the underlying rationales of rubric use.

Given that one of the reasons behind using rubrics is to promote Assessment for Learning, Chapter 3 discusses the body of literature on rubrics. It begins with describing the functions of rubrics and then critically debates the benefits and disadvantages of rubrics. The benefits are classified as: grading, learning and psychological benefits. The relationship between rubrics and Assessment for Learning is drawn here. Criticisms of rubrics involving design and actualisation issues are examined. This chapter finishes

with examples of rubric utilisation, targeting some of the criticisms and discusses how this literature informs gaps for the present study.

The last literature review chapter, Chapter 4, is a brief chapter examining the body of literature on assessment policy in higher education. This is a short chapter because literature in this area is scarce and only limited relevant information could be found. They include the purpose of an assessment policy, opinions of such an assessment policy from various stakeholders and suggestions of how an assessment policy should be adopted in the higher education context. This chapter provides grounds for comparison between the adoption of a rubric policy in the study context and other assessment policy initiatives worldwide.

Chapter 5 is the chapter on methodology and methods of this research. Descriptions of ontological and epistemological positions are provided to justify the use of phenomenology in this study. Explanations and arguments are included to debate why phenomenology is used as opposed to other methodologies. This chapter also includes a description of insider research, the methods (focus groups and semi-structured interviews) employed, the participant recruitment process, as well as the data collection and analysis process. Ethical concerns and limitations of the methodology and methods are examined together with suggestions for addressing the limitations and enhancing creditability.

Chapters 6 and 7 are the two results chapters. Chapter 6 focuses on answering research questions on perceptions, experiences, barriers and

opportunities for implementing Assessment for Learning in one particular university in Hong Kong. Three themes are identified for this part of the research: (1) elements of assessment- facilitating learning; (2) the traditional burden and mismatch in current assessment practice and (3) promoting learning with authentic and holistic assessment. Chapter 7 turns to report the findings on rubric perceptions and actualisations under the mandatory top-down rubric policy situation, as well as the implications of the rubric policy for teachers and management. Four themes are identified in rubric and policy areas: (1) the potentials of rubrics as a learning tool; (2) the barriers of using rubrics; (3) the optimisation of rubrics for Assessment for Learning; and (4) the holistic promotion of Assessment for Learning from a policy level.

Chapter 8 is the chapter on discussion of findings. It builds on the themes identified in the previous two chapters and provides interpretations and explanations of findings. Explanations are grounded mostly in the relevant literature in the framework described in Figure 1.1 and comparisons of findings with this literature are discussed. Overall, critical analysis of the findings suggests that the participants carry anecdotal knowledge and experience with Assessment for Learning and rubrics despite the heavily-ingrained examination-oriented culture. However, actualisations of Assessment and Learning are challenging and these barriers are multifactorial ranging from culture to knowledge to practice. The policy initiative is good but implementation is poor. It implies that a more holistic policy model is imperative.

This dissertation finishes with Chapter 9, where its contribution to knowledge and its conclusions are discussed. Drawing upon the study's findings, a holistic assessment policy model is recommended and can be used for policy refinement of this research institute or other similar universities. The chapter finishes with limitations of the study and future research recommendations.

Chapter 2: Understanding Assessment Practice in Higher Education

The goal of this chapter is to review relevant literature on assessment to classify the purposes and functions of assessment and to explain the current situation of assessment practice and factors contributing to this both locally and internationally. This information provides background understanding of assessment, especially the evidence-based concept of Assessment for Learning, in order to inform explanations and discuss the implications of the findings.

The chapter begins with an overview of various functions and purposes of assessment in higher education. It is important to understand the different functions so as to comprehend the various definitions in current assessment systems. The chapter then discusses major factors that lead to the current situation of assessment practice and the challenges hindering assessment reforms. A special discussion is dedicated to the influence of the Confucian-heritage culture as this present study is conducted in Hong Kong where the education system is heavily affected by this culture. The chapter closes with a summary and implications of this knowledge for the present study.

2.1 Functions and Purposes of Assessment in Higher Education

Assessment is a significant component at any level of education. All students go through some form of assessment in their education experience. Experts in assessment in higher education around the world have identified reasons for why assessment is important:

-
-
- it powerfully dictates how students learn (Sambell et al., 2017) because the results of assessment affect students' future direction and careers (Boud & Falchikov, 2007);
 - it influences what teachers do in teaching and learning activities (Carless, 2015a); and
 - assessment is not an option for students and they have to do it no matter what (Carless, 2017).

Given these reasons, assessment brings significant value to and plays critical roles in all education systems.

There are diverse values and roles of assessment. The most commonly accepted four functions of assessment are: certification, quality assurance, learning and lifelong learning (Sambell et al., 2017). These four functions are discussed in the following sub-sections.

2.1.1 Assessment for Certification and for Quality Assurance

The *certification* function is perhaps most familiar to many teachers and students. It entails collecting evidence on students' performance and deciding if students meet a particular standard to progress in the curriculum (Ashwin et al., 2015). The focus of Assessment for Certification in higher education is to identify the results of learning, produce degree awards and classifications of award. Thus, it is also called Assessment of Learning. The assessment result is often represented by a mark or a grade, with summative assessment the most common way to generate this result (Ashwin et al., 2015). In fact, Assessment for Certification, Assessment of Learning and summative

assessment are seen as describing similar entities in higher education assessment practice (Boud, 2007; Sambell et al., 2017).

The mechanism behind Assessment for Certification is measurement or evaluation (Boud, 2007) through a process of judgment (Knight, 2007). Assessment that carries the *certification* function is often perceived as high stakes because it dictates important progress in one's education and ultimately graduation (Ashwin et al., 2015).

The *quality assurance* purpose is somewhat similar to *certification* because the assessment result is used to demonstrate a standard (Sambell et al., 2017). This standard in the *quality assurance* function, however, is for an institution rather than for an individual student. Here, accountability and ranking are the focuses (Sambell et al., 2017). Nonetheless, Assessment for Quality Assurance is also considered as Assessment of Learning because the major function is the measurement of results. This is in contrast to improving learning during the process, which is discussed in the upcoming sub-section.

2.1.2 Assessment for Learning

Learning is the central purpose in the other two functions of assessment. *Assessment for Learning* focuses on immediacy and the short term while *Assessment for Lifelong Learning* puts the centre of attention on the long term and sustainability (Sambell et al., 2017). In general, Assessment for Learning provides opportunities for students to make use of information gathered or information that arises during the assessment process, to self-regulate what they need to learn or how they need to perform towards a pre-set goal

(Sambell et al., 2013). Similar to the interlinking terms of Assessment for Certification and summative assessment, Assessment for Learning is often associated with formative assessment (Ashwin et al., 2015). Formative assessment provides performance information or feedback to students. Feedback is core in this type of assessment and should be systematically given in order to promote learning (Black & Wiliam, 1998; Sadler, 1998).

It is important to note that Assessment for Learning does *not* necessarily exclude summative assessment nor deny the *certification* function (Sambell et al., 2013). If the purpose of the overall assessment is to help students learn, it can still be considered as Assessment for Learning despite the inclusion of summative assessment (Sambell et al., 2013). This is similar to another framework proposed by (Carless, 2007) named learning-oriented assessment. This framework stresses that the *certification* and *learning* functions should substantially overlap where the central focus is to facilitate learning (Carless, 2007). To achieve this central purpose, three components are crucial: (1) the assessment tasks should also serve as learning tasks, (2) students should actively engage in the assessment process and (3) students should be able to receive and appropriately use feedback for future improvement (Carless, 2007, 2015b).

Regardless of different frameworks, the *learning* function is engineered through enabling students to evaluate and reflect on their own progress and outcome so that they can direct their own learning (Ashwin et al., 2015; Sambell et al., 2013). This type of assessment is seen as low-stake because students are given the opportunities to learn and improve before reaching a

final mark/grade (Ashwin et al., 2015). One criticism of this function is that it focuses on the short term and fails to assess the real application, which always happens after the study is completed (Boud, 2007). The other function of assessment, *future/lifelong learning*, serves this purpose and is discussed in the following sub-section.

2.1.3 Assessment for Future/Lifelong Learning – Sustainable

Assessment

The function of *future or lifelong learning*, as the name implies, is to use assessment to facilitate learning beyond the student's immediate needs. Assessment that carries this function is referred to as Assessment as Learning (Ashwin et al., 2015; Sambell et al., 2017) or Sustainable Assessment (Boud, 2000). This type of assessment goes beyond making use of feedback information in the immediate assessment context to developing one's evaluative judgment for lifelong learning after graduation (Boud, 2000; Boud & Associates, 2010). This lifelong assessment ability is crucial because it is an attribute needed continually throughout one's life.

The skills learnt in this type of assessment centre on using and judging feedback, because formative assessment is still the main assessment type. However, students do not only learn how to use such feedback, but actively engage in feedback (Ashwin et al., 2015). During the assessment process, students should be given opportunities to judge and discern their quality of work (Boud, 2000). This active practice in the classroom builds students' own assessment ability required for lifelong learning in the workplace or in society

(Ashwin et al., 2015; Boud, 2000). Thus, the focus is to move away from depending on others' feedback while building capacity and developing judgement about their own learning (Boud, 2000). The ability should then be sustained throughout life.

2.1.4 Summary of Section and Implications

As evident in this section, assessment involves multiple functions and purposes. These functions and purposes do not operate in isolation and therefore assessment can be a very complex entity (Carless, 2015a). Not only are the functions of assessment many but they also compete with each other because of their very different natures (Carless, 2015a). When the focus is put on measurement, the learning function is compromised. In current higher education assessment practice, measurement is still the dominating function (Boud, 2007). However, research evidence supports making learning the central function of assessment (Black & William, 1998; Carless, 2015a; Hawe & Dixon, 2017; Jessop, 2017). Thus, switching the focus from Assessment for Certification to Assessment for Learning is imperative yet challenging.

Boud (2000) stresses that assessment is always a double duty. As mentioned previously, advocating Assessment for Learning does not mean to eradicate the *certification* function completely. The merit lies in finding a balance between the competing functions, yet understanding and implementing evidence-based assessment practices for the benefit of students' learning. To locate this balance, the first step is to understand the concept and current picture of assessment practice and the possible

contributing factors behind this practice. This understanding allows teachers and administrators to more successfully tackle the double duty of assessment. After all, there is a need for administrators, teachers and students to reconceptualise their roles and responsibilities in assessment because of the evolving learning function assessment entailed in this post-delivery education age (Sambell et al., 2017).

2.2 The Current Picture of Assessment Practice in Higher Education and the Contributing Factors to this Picture

Undoubtedly, Assessment for Certification is still the dominating picture of assessment practice in education systems nowadays, including higher education locally and internationally (Boud, 2007; Ewell, 2006). Numerous complaints about this assessment practice have been reported in the literature. From the students' perspective, these complaints involve assessment authenticity, a mismatch of learning objectives, over-assessment, unclear expectations, a heavy focus on recall and scepticism of fairness (Carless, 2015a; Flint & Johnson, 2010).

These issues appear to be inter-related: when an assessment does not mimic real life needs, it is difficult to link with learning objectives (which are usually the application of knowledge in the real world) and may just focus on the regurgitation of knowledge. Also, when expectations are unclear, students' interpretations can be different and thus create a sense of unfairness. The unfairness goes beyond interpretation and students also worry if they are

given the opportunity to demonstrate these competencies (Flint & Johnson, 2010).

In addition to the assessment system itself, students are generally dissatisfied with the current amount and delivery of feedback (Sambell et al., 2017). Feedback is not given at a suitable time to enable improvement, and that feedback is usually vague leaving little value for constructive use (Carless, 2017).

Higher education assessment research experts have shared their views on the reasons behind this dissatisfaction. In particular, the *certification* purpose in higher education institutions appears to be the dominating function (Boud, 2007) and available higher education assessment policies also suggest that Assessment for Certification is the focus of assessment practice (Boud, 2000, 2007). Examples of these policies include those for quality assurance or for streamlining the assessment process (Boud, 2007). It is problematic to have such a strong focus on Assessment for Certification because it encourages students to focus only on grades and disparages the learning function of assessment. Section 2.2.1 further discusses the issues this brings.

The insufficient focus on learning within the assessment system may also stem from inadequate knowledge of the *learning* function (Carless, 2015a). Most teachers may only be aware of the measurement function as it is traditional and is the teachers' own experience as students. They may have insufficient information about other assessment functions and their effects

(Boud, 2007). Hence, teachers may only focus on the assessment function they are familiar with, or believe in (Carless, 2015a).

Specifically in Hong Kong, the deeply rooted examination-oriented culture and the societal value of examinations add another layer of challenge to focusing on the *learning* function (Carless, 2011). Moreover, lack of time due to heavy workloads and prioritising research activities over teaching are frequently limiting factors for assessment change (Deneen & Boud, 2014; Norton, Norton, & Shannon, 2013). Teachers evidently face practical challenges to change a well-accepted and familiar assessment practice. The following sub-sections discuss these factors in more detail.

2.2.1 Heavy Focus on Assessment for Certification

As mentioned previously, Assessment for Certification entails the evaluation of the learning outcome to determine if students reach pre-set standards, often in terms of marks or grades or classifications. Examinations are a classic example of Assessment for Certification (Knight, 2007). In fact, any assessment tasks that are summative in nature are closely associated with Assessment for Certification. Typically, students are offered one attempt to show if they have met the standards.

The issues with this type of one-off examination/assessment have been well documented (Boud, 2007; Carless, 2015a; Sambell et al., 2017). A number of these issues relate to the detrimental effects on learning. With grades being the ultimate focus in Assessment for Certification and common gatekeepers of performance standards (Carless, 2015a), they shape students

as instrumental rather than critical learners (Jessop, 2017). In particular, students may position their strategies towards obtaining high grades and this behaviour may not be conducive to learning (Carless, 2015a; Knight, 2007).

Not only does Assessment for Certification affect students' learning behaviour, Knight (2007) describes in detail the inherent limitations of focusing on measurement and classification functions. He states that achievement in higher education is complex and is socially constructed; it is basically impossible to measure complex achievement using the existing practice of examination-focused assessment. This study implies that the current methods of measurement are meaningless, irrelevant and unreliable. With regard to using assessment results for ranking purposes, Knight stresses that assessment standards are context-specific and stem from one area of the world that may not be equivalent to another. Thus, the meaning of ranking is skewed.

Perhaps the most important issue is the competition between the functions of measurement and learning (Carless, 2015a). When the focus is to produce grades and rankings, the focus on facilitating learning is compromised (Knight, 2007). Students may refrain from asking constructive questions and teachers may limit their time for feedback or other good practices conducive to learning (Carless, 2015a). This is ironic as Assessment for Certification drives out learning while measuring it at the same time (Boud, 2000).

Is Assessment for Certification or examination necessarily a negative process? Several researchers have contested that Assessment for Certification still has potential if the design of the assessment tasks aligns constructively with learning objectives (Broadbent et al., 2018; Sambell et al., 2017). To achieve this alignment, assessment tasks should be authentic and should include the real life application of knowledge (Carless, 2017). This type of assessment task encourages a deep rather than a superficial learning approach. Open-book examinations are an example of this type of assessment, despite the fact that they are still examinations (Sambell et al., 2017). It requires teachers, administrators and even students to be aware of the potential benefits underlying summative assessment and to find ways to improve its design.

Although it is not the objective of this study to explore the perception of Assessment for Certification, this topic is expected to surface due to its dominating function within the current assessment system in Hong Kong. It is interesting to understand how stakeholders in this university view the functions of assessment and if their views are comparable to those in previous research, especially as this university is situated in a culture described as examination-oriented. Understanding participants' perceptions also helps to explain how they interpret and actualise rubrics. The following subsection discusses in detail this cultural influence on assessment practice.

2.2.2 Examination-Oriented Culture in Hong Kong

Differences between the Western and Eastern education systems have been documented with the Western education system characterised as more student-centred and the Eastern education system as more teacher-centred (Ho, 2010). In student-centred practice, students and teachers share relatively equal roles of instructing and learning. Discussions among peers and teachers are common; teachers are the facilitators and students are active in their learning processes. On the other hand, the teacher-centred practice is more one-way where students are passive recipients of information. Teachers are seen as the authoritative figures delivering knowledge. Questioning is perceived as disrespectful.

With regard to assessment practice, the Western, student-centred education system is characterised as formative-focused, while the Eastern, teacher-centred education system is often labelled as summative-focused (Berry, 2011). Summative assessment, as mentioned previously, is closely associated with examinations because it is the most common method used to produce end-of-learning results. Examinations therefore play a dominant role among Eastern teachers and students and are seen as an approach productive for learning (Brown et al., 2009; Carless, 2011).

This culture of teacher-centred and examination-focused learning is described as being part of the Confucian-heritage culture (Berry, 2011; Biggs, 1996). Confucian-influenced/Eastern countries with a long history of this culture include Hong Kong, China, Taiwan, Singapore, Malaysia, Korea,

Japan and Vietnam. In the Confucian-heritage culture, learning and assessment favour collectivism rather than individualism (Carless, 2011). There is a power/hierarchical relationship between students and teachers where the authority of teachers described above is commonly accepted (Carless, 2011). The emphasis of learning is on effort and diligence. Harmony is also a unique feature in this culture and it explains why little communication occurs between teachers and students in order to avoid conflicts.

The Confucian-heritage culture has its historical origins in the ancient Chinese empire. Success in the imperial examination system of various Chinese dynasties was an important step to achieving status, income and power in society (Berry, 2011; Carless, 2011). This system lay the foundation of the emphasis on the results of examination. The by-products of such a system are described as building competition, stressing memorisation, as well as putting a heavy focus on the utilitarian nature of education, on examination success, book knowledge and final assessment (Carless, 2011). These impacts promote the perception that examination is the main goal of education, the ultimate pathway for career advancement and for moving up in status in society (Carless, 2011). As negative as they may sound, these impacts are described as culturally accepted in many Confucian-heritage culture countries in today's world (Brown et al., 2009; Pham & Renshaw, 2014).

Given the historical origin, summative assessment and competitive examination are deeply rooted in the educational practice and expectations of Hong Kong teachers and students (Brown et al., 2011; Brown & Wang, 2013).

When the common assessment practice emphasises the final result, students are conditioned to focus on the accumulation of knowledge rather than the mastery of skills (Carless, 2011). High achievement is valued over meaningful learning and this value is thought to limit the adoption of formative assessment or other assessment for/as learning (Carless, 2011).

The systematic review by Black and Wiliam (1998) prompts some attention to the benefits of formative assessment and reforms have been implemented at primary and secondary school levels in Hong Kong (Berry, 2011) and other Confucian-heritage culture countries (Pham & Renshaw, 2014; Ratnam-Lim & Tan, 2015). However, not all initiatives have been successful and this may be due to inadequate consideration of historical and culturally-ingrained factors (Pham, 2011). This implies that the cultural factor may affect how the participants in this research context perceive and practise assessment, which is also directly linked to their perceptions and the implementation of rubrics.

2.2.3 Insufficient Assessment Literacy

The *certification* function is deeply rooted in many teachers because it is what most of them experienced as students. Because of their own personal exposure, teachers are familiar with this focus and may not be aware of other functions of assessment. Boud (2007) states that the notion of assessment and learning is “not sufficiently well located within the dominant discourse of assessment” (p.14). Thus, it is not surprising that teachers lack information and knowledge of other assessment functions and effects.

Assessment literacy is the term that describes this understanding of information and knowledge of contemporary assessment practice necessary to best measure students' achievements (Smith, Worsfold, Davies, Fisher, & McPhail, 2013; Xu & Brown, 2016). Being assessment-literate means that the individual possesses a conceptual understanding of assessment as well as skills and intellectual abilities in self- and peer-assessment using technical approaches (Price, 2012).

Based on the conceptual definition by Price, Medland (2015) identifies six elements that characterise assessment literacy:

- A community sharing standardised assessment practice.
- A dialogue between all stakeholders when building assessment practice.
- Knowledge and understanding of effective feedback.
- A programme-wide approach that looks at the alignment of assessment.
- Outcome, adoption of assessment that builds self-regulation.
- A shared understanding of assessment standards.

Looking at Price's definition and Medland's elements, stakeholders (students, teachers and administrators) are required not only to understand assessment concepts in-depth but to effectively execute various assessment functions and practices to facilitate learning. Being assessment-literate may be a difficult goal to attain, especially in Hong Kong where the impact of Confucian-heritage culture is huge.

In fact, insufficient assessment literacy has been reported in higher education research and this is observed among both teachers and students worldwide. For example, teachers are inconsistent with the use of assessment terminology and are anxious not to follow conventional assessment practice (Forsyth, Cullen, Ringan, & Stubbs, 2015). Levels of and familiarity with assessment literacy vary greatly among academics (Medland, 2015, 2018; Rezvani Kalajahi & Abdullah, 2016). Students also possess variable levels of ability to grade others' work (Rhind & Paterson, 2015).

The implication of suboptimal assessment literacy is that it limits how stakeholders perceive and engage in assessment. When applied to rubric practice, stakeholders may only use rubrics for quality assurance and to demonstrate the robust nature of a mark instead of facilitating learning. Knowing the definition and the current picture of assessment literacy helps to explain and understand how participants view and actualise assessment and rubrics in this study context.

2.2.4 Resources, Regulations/Policies and Cultures set Work Priorities

Apart from personal professional factors such as assessment literacy and teaching beliefs, environmental factors also influence how teachers design and implement assessment (Bearman et al., 2017). Environmental factors include resources, institutional and/or departmental regulations/policies and cultures (Bearman et al., 2017; Joughin, Dawson, & Boud, 2017). They inevitably play a role in shaping how teachers set their work priorities when

time is so precious. This subsection discusses how these factors contribute to issues of current assessment practice.

Sufficient resources entail various dimensions such as available time and support. It is obviously crucial to have ample resources for any change to occur. This is especially true for changing assessment practice because teachers do not seem to possess adequate assessment literacy to begin with. It is therefore reasonable to assume that teachers need an abundant amount of time and training to improve their assessment knowledge. In fact, faculties perceive changing assessment practice as time-consuming and as an increase in their workload (Bahous & Nabhani, 2015). Increased workload has also been found to decrease a faculty's participation and satisfaction in good assessment practice (McCullough & Jones, 2015). Workload and lack of time are interlinking factors governing assessment change.

Adequate training is another essential factor to drive change, especially for teachers with insufficient assessment literacy. Although training and support seem to be happening and helping (Deneen & Boud, 2014; Sayigh, 2006), time as well as effective communication strategies allocated for such training may be suboptimal (Ebersole, 2009). There is insufficient time for repetitive good assessment practice, which limits engagement for practice change (Ebersole, 2009). Lack of time still seems to be an issue in situations where training is provided.

Other factors that directly affect the amount of time spent on assessment practice are institutional and/or departmental policies, regulations

and cultures (Joughin et al., 2017). Chapter 4 of this thesis discusses in detail the current assessment policy picture in higher education; therefore, this topic is not reviewed in-depth here. In brief, policies and regulations have a powerful influence on teachers' choice of work priorities.

The heavy workload of academics partially explains the lack of time. How a faculty prioritises work also matters. Contemporary higher education ranking mechanisms put a heavy focus on research as compared to teaching. Also, faculty appraisal systems value research output rather than good teaching (Raaper, 2016). This translates to the phenomenon that research activities are prioritised over teaching, especially when the workload is heavy (Bahous & Nabhani, 2015; Macdonald & Joughin, 2009; Tagg, 2012). A faculty often sees no incentives to spend time on teaching improvement, let alone on assessment change which is only one area of teaching (Norton et al., 2013).

In addition, changing conventional assessment practice is perceived as a high stakes process (Deneen & Boud, 2014) and such change may not be welcomed by students (Norton et al., 2013). As a result, a faculty may choose to invest their precious time on research-related tasks that are less risky and more beneficial to their career advancement.

With appropriate incentives and recognition coming from policies and regulations, assessment change can be promoted. Macdonald and Joughin (2009) state that recognition of good assessment practice has a strong influence on improving assessment change. However, a faculty might feel

there is little recognition in the existing system (Raaper, 2016). There is a place for embedding recognitions in policies or guidelines to encourage good assessment practice, and this recognition should apply to teachers in all academic tracks and should consider differences in organisational cultures (McCullough & Jones, 2015; Norton et al., 2013).

Apart from black-and-white regulations and policies, organisational culture exerts a powerful influence on assessment practice and change (Bearman et al., 2017; Macdonald & Joughin, 2009). Organisational culture encompasses implicit and explicit “ways of doing things” at departmental or disciplinary and institutional levels (Bearman et al., 2017). In contrast to policies and regulations, cultures are usually more silent because most of them are not written down. Nonetheless, they dictate to a certain extent how a faculty sees and prioritises assessment (Bearman et al., 2017; Joughin et al., 2017; Macdonald & Joughin, 2009). If a faculty senses a culture of valuing assessment, they are more likely to participate in upholding standards and changing practice (McCullough & Jones, 2015).

This subsection highlights the relationship between resources, regulations/policies and culture and how they impact on teachers’ work priorities and allocation of time. Academics face a heavy workload on a daily basis and priorities constantly compete with each other. The solid regulations/policies or systems that value research over teaching consider assessment change as a lower priority. This may create an explicit and implicit culture of research superiority, with less resources and attention given to teaching. It is interesting to investigate if this complex interplay of

regulations/policies/organisational culture and resources also influences the view and practice of Assessment for Learning and the adoption of the rubric policy in this Confucian-heritage university, given that there is a heavy historical tradition of examinations and an international trend of being awarded ranking according to research output.

2.2.5 Summary of Section and Implications

This section has critically discussed the literature related to the current picture of assessment practice in higher education and explained factors contributing to the scenario. In sum, there is general dissatisfaction about the existing assessment practice resulting from its heavy focus on Assessment for Certification globally and the examination-oriented culture locally in Hong Kong (and other countries influenced by the Confucian-heritage culture). Apart from the cultural factor, insufficient assessment literacy appears to limit teachers and management from overturning the dominant practice. In addition, traditional organisational cultures and regulations and a lack of resources place research over teaching as a priority in many higher education institutions. As such, changes related to teaching improvements receive less attention despite good intentions. These factors all contribute to the current dominant practice of assessment.

Literature knowledge from this section informs the importance of exploring participants' perceptions and experiences of assessment practice in this research setting. The information generated allows an understanding of the patterns and focus of assessment practice in this institution, and helps

identify plausible reasons behind the interpretation of the rubric policy and the actualisation of Assessment for Learning and rubrics.

2.3 Summary of Chapter and Implications for this Study

This chapter first explains and contrasts the various functions of assessment. It then provides an in-depth literature review on why the current assessment practice is still measurement- dominant internationally but more so in Hong Kong due to the cultural influence. While it is imperative to explore if the participants in this research context also hold the same view on assessment practice, knowledge from the literature review sets the underlying assumption that the concept of Assessment for Learning is less understood in Hong Kong. Thus, relating assessment and learning may not come to mind naturally.

This underlying assumption informs putting Assessment for Learning upfront in the research questions; this study can thereby explicitly investigate the perceptions and experiences of Assessment for Learning among the participants. Researching perceptions and experiences of Assessment for Learning is essential as rubrics feature in this concept and the findings help to associate and explain the overall picture of rubric utilisation in this research context. The next chapter discusses relevant literature on rubrics where research gaps are identified.

Chapter 3: The Use of Rubrics in Higher Education

Assessment rubrics are traditionally referred to as a grading guide to enable marking objectivity and consistency (Popham, 1997). They have been advocated as a learning tool in recent years because they relate to the context of Assessment for Learning (Panadero & Jonsson, 2013; Reddy & Andrade, 2010).

Despite its growing popularity, rubrics have not been widely adopted in Hong Kong. Literature on the use of rubrics in Hong Kong is scarce and examples are primarily drawn from award-winning teachers (Carless, 2015a). Taking place in a mandatory top-down rubric policy initiative in one particular university in Hong Kong, this study explores the perception and experience of students, teachers and management on the development and actualisation of rubrics within the research context. Specific attention is given to explore the perceived relationship between rubrics and Assessment for Learning. This relationship is not a new concept (Carless, 2017); however, it may not be a naturally occurring concept in Hong Kong due to the deeply-rooted examination-oriented culture. In addition, the interpretation and adoption of rubric practice as a result of a compulsory top-down policy may be different from a voluntary initiative. It may well be the case that teachers executing this policy may possess less knowledge of rubrics than teachers who voluntarily use the tool. These are all research gaps that this study attempts to fill.

This chapter begins with an overview of the functions and purposes of rubrics in higher education. Understanding what rubrics do enables a critique of their merits and disadvantages. The merits of rubrics are classified in terms

of grading, learning and psycho-social benefits. The relationship between rubrics and Assessment for Learning is highlighted in the merits section. On the other hand, criticisms of rubrics are discussed in the areas of development and utilisation. The chapter then provides examples of rubric actualisations that address some of problems with rubrics. Research gaps are explained at the end of this chapter.

3.1 Functions and Purposes of Assessment Rubrics

Simply put, rubrics are an assessment tool to guide performance scoring (Popham, 1997). This tool is operated on the basis of providing qualitative descriptors of discrete evaluative criteria (Popham, 1997). These allow teachers to grade according to the descriptors. Ideally, evaluative criteria should be discrete and clear to facilitate teachers' and students' understanding and the distinctions of standards and expectations. That is, users should be able to understand the expectations and differences of an A versus a B versus a C grade performance based on the rubric.

Since rubrics explicitly spell out grading criteria and descriptions of these criteria, they are advocated for use as learning tools (Panadero & Jonsson, 2013). Specifically, the descriptive criteria can be used as feedback which informs students of the qualities they are lacking in a specific assessment task (Jonsson & Svingby, 2007; Panadero & Jonsson, 2013). Rubrics can also serve the function of a planning tool if they are given to students in advance (Tay, 2015). Students can make use of the qualitative

descriptors listed in rubrics and plan the strategies and steps to achieve performance goals.

From the above descriptions, rubrics serve both summative and formative functions – for grading and for learning. Using rubrics as learning tools in higher education has received growing attention in recent years, and this is probably related to the powerful position of rubrics in the context of Assessment for Learning (Panadero & Jonsson, 2013). The following section discusses the merits of using rubrics as grading and learning tools, and critically reviews the relationship between rubrics and Assessment for Learning.

3.2 Merits of Using Rubrics in Assessment Practice

The merits of assessment rubrics have been reported in the literature. To receive a more holistic picture of these merits, they are classified into different categories, namely grading, learning and psycho-social benefits. This section discusses these merits and explains reasons for the growing popularity of rubrics.

3.2.1 Benefits of Using Rubrics as Grading Tools

Both teachers and students may have encountered challenges with grading and grades respectively in their education experience. Specifically, students often question how they got a particular grade and may think teachers grade by their impressions (Andrade & Du, 2005). As a result, there is scepticism

that the grades are a subjective reflection of impression rather than an objective and fair measure.

On the other hand, teachers face particular challenges when multiple graders are involved. There seem to be different standards or different interpretations of grades among graders, hence causing inconsistent gradings. It may also be difficult for teachers to clearly communicate their expectations to their students, which leads to differences in grade expectations between teachers and students.

Rubrics have been shown to address some of the aforementioned grading problems. For example, Bell, Mladenovic, and Price (2013) report that students like the idea of having a standard and guidance when rubrics are used. Jonsson (2014) also argues that rubrics facilitate students' self-assessment because the criteria are transparent. In addition, rubrics enhance a shared understanding of assessment tasks which in turn facilitates collaborative learning (Mauri, Colomina, & de Gispert, 2014). These benefits are thought to be mediated through an increased transparency of grading information (Jonsson, 2014; Panadero & Jonsson, 2013). However, rubrics need to be available to students prior to assignment deadlines in order to actualise and maximise these benefits. In other words, how teachers implement rubrics in their courses is equally important as the designs of rubrics.

For teachers, the merits of rubrics also lie in their potential ability to increase grading reliability, consistency and validity (Jonsson & Svingby,

2007; Menéndez-Varela & Gregori-Giralt, 2016). This is especially true when multiple graders are involved in marking a single assignment. When the criteria are explicitly listed, different teachers can refer to the same list of criteria to cross-check the level of performance of their students. This practice should theoretically limit subjective interpretation but allow the use of concrete standards when judging students' performance.

Nonetheless, just using rubrics may not provide these benefits unless the design and utilisation of rubrics are complemented with other measures, as stated by Jonsson and Svingby (2007): “the reliable scoring of performance assessments can be enhanced by the use of rubrics, especially if they are analytic, topic-specific, and complemented with exemplar and/or rater training” (p.130). This point is further elaborated in the criticism and actualisation sections of this chapter.

The benefit of increasing the validity of assessment with the use of rubrics is a more complex and controversial topic because it depends on the design of the particular rubric and whether it is context/course-specific. Menéndez-Varela and Gregori-Giralt (2016) confirm the validity of their rubrics for service-learning projects due to shared understanding and feedback embedded in the assessment process.

On the contrary, however, Jonsson and Svingby (2007) criticise the general validity of rubrics after they reviewed 75 studies. They argue that rubrics may not increase assessment validity but that valid assessment lies in using a valid framework including rubrics. The key point lies in the need for

validating the rubric rather than using a rubric to make assessment more valid. Nonetheless, rubrics have the potential to improve the validity of assessment if they are fit-for-purpose and used appropriately.

3.2.2 Benefits of Using Rubrics as Learning Tools – Assessment for Learning

Beyond grading, rubrics have been advocated as a learning tool in recent years and are situated within the concept of Assessment for Learning.

Considerable research has explored if and how rubrics facilitate learning. As previously mentioned, rubrics can improve the communication of grading standards between teachers and students and thus enhance marking benefits. This communication extends beyond understanding standard criteria but also helps to identify learning goals (Andrade, 2005; Jonsson & Svingby, 2007; Mauri et al., 2014). This is possible when students study the criteria for achieving excellent performance and the criteria in turn become goals for students to strive for.

How does providing transparent and explicit criteria in rubrics facilitate learning? From critically reviewing the relevant literature, learning appears to be achieved through self-assessment/self-regulation, evaluative judgement and feedback utilisation. In fact, rubrics have been described as one of the common tools used for students' self-assessment (Panadero & Jonsson, 2013; Panadero & Romero, 2014).

The underlying mechanism of using rubrics to self-assess is relatively straight-forward: when students look at the criteria in the rubrics, they

evaluate if their own work matches these criteria (Jonsson, 2014; Panadero, Alonso-Tapia, & Reche, 2013; Van Helvoort, 2010). In other words, students judge if their goals are met using rubrics (Andrade, 2005). Through the self-assessment process, students develop their own strategies to reach the performance level they are aiming for (Panadero & Romero, 2014). This action describes self-regulated or self-oriented learning. Self-assessment and self-regulation, therefore, go hand-in-hand and complement each other in the rubric utilisation process.

Self-regulated learning can also occur before self-assessment, usually at the preliminary stage where students plan the steps they will use to approach an assignment (Zimmerman, 2000). Rubrics can be used to facilitate this planning phase when they are provided to students beforehand, as students can make use of the criteria to formulate a plan to achieve these standards (Panadero, Andrade, & Brookhart, 2018). Rubrics can decrease the negative self-regulatory actions (e.g. avoiding behavior and actions for self-regulation). This is especially true with low to moderately complex tasks, thereby favouring self-oriented learning (Panadero et al., 2013; Panadero & Romero, 2014). Co-creating rubrics with students is another way to enhance self-regulation (Fraile et al., 2017). It is thought that co-creation is powerful in activating students' learning strategies.

Moreover, self-assessment builds evaluative judgement because students critically appraise the quality of their own work to differentiate the good and not-so-good components (Tai, Ajjawi, Boud, Dawson, & Panadero, 2018). Evaluative judgement does not automatically emerge with the use of

rubrics; rather, students learn to judge over time by comparing their work with exemplars in the field and by reflecting on feedback embedded in rubrics, in order to identify the components constituting excellent work (Carless, 2015a). Tai et al. (2018) state that training and the repetitive use of rubrics are important steps to develop evaluative judgment in students. Carless (2015a) also stresses that “exposure to criteria has been shown to be insufficient to support students in developing a firm understanding of what is required in assignments” (p.148). This point is further discussed below in the section on criticism. Nonetheless, the literature implies that rubric actualisation is crucial to optimising these benefits and this is one of the goals of this study.

Another way rubrics contribute to learning is by feedback utilisation. By virtue of the design of rubrics, the feedback mechanism is embedded because students can look at their grades and reflect on the quality they are lacking in their work (Andrade & Du, 2005; Bell et al., 2013), as the qualitative descriptions are clearly listed. Both teachers and students find that reflecting on feedback within rubrics is helpful for learning and value this experience (Andrade & Du, 2005; Panadero & Jonsson, 2013). In addition, it is easier for teachers to provide detailed feedback to students through the use of rubrics (Van Helvoort, 2010).

Feedback does not only come from teachers but also from peers. Greenberg (2015) states that with structured rubrics, peer assessment and feedback is enhanced because students can easily follow the criteria when evaluating. This in turn facilitates mutual learning. It is evident from the above

literature that rubrics can be instrumental in guiding the feedback process when used appropriately.

All in all, this sub-section examines the underlying mechanisms of using rubrics to promote learning. Rubrics have been shown to activate students' ability to self-assess, self-regulate, critically judge and use feedback. However, this evidence is mainly reported in the Western context. The questions remain if the same benefits are perceived and experienced in Hong Kong, especially in an institution where using rubrics is not standard procedure or even a non-voluntary practice. Users' perceptions on the relationship of rubrics with Assessment for Learning may be different. In addition, it is unclear if the various stakeholders of rubrics experience the same benefits while using rubrics. This study attempts to answer these questions.

3.2.3 Psycho-social Benefits of Rubrics

The psycho-social benefits of using rubrics are less documented but are evident in the literature. Reducing anxiety is perhaps a major psycho-social benefit of using rubrics. Regardless of how low-stake an assessment is, it is deemed a stressful event because grades are involved. Rubrics can decrease stress and anxiety related to assessment because grading criteria are transparent (Andrade & Du, 2005; Greenberg, 2015; Panadero & Jonsson, 2013).

This increased transparency allows students to understand the criteria beforehand and enables them to self-regulate and self-assess their goals and

performances. Students feel more confident about how they can handle the work and therefore feel less nervous (Andrade & Du, 2005; Panadero & Jonsson, 2013; Reynolds-Keefer, 2010). In addition, students are better able to ask appropriate questions when rubrics are available, further reducing their anxiety during the assessment process (Reynolds-Keefer, 2010). This claim of reduced anxiety is not without controversy because a higher stress level is also mentioned by some students using rubrics (Panadero & Romero, 2014). In addition to simply using rubrics, perhaps their design and utilisation is also an important point to consider.

Another potential psycho-social benefit of rubrics is increased self-efficacy; however, controversial results have been reported. Theoretically, students' increased confidence in their self-regulation may translate to higher levels of self-efficacy and this is the hypothesis for studying the impact of rubrics on self-efficacy. Nonetheless, most studies find no effect of rubrics on self-efficacy (Fraile et al., 2017; Panadero et al., 2013) in university students; while the impact in elementary/secondary school students is mixed (Andrade, Wang, Du, & Akawi, 2009; Panadero, Tapia, & Huertas, 2012). The mixed findings suggest more complex factors, such as rubric implementation, may have contributed to the psycho-social benefits.

3.2.4 Summary of Section and Implications

Rubrics bring solid grading and learning benefits to teachers and students while their psycho-social impact is mixed. As far as grading benefits are concerned, rubrics increase the transparency of grades, enhance the

understanding and expectation of standards, as well as improving the consistency and reliability of gradings. These benefits mediate communication between teachers and students and tackle the issues underlying traditional grading practice.

When expectations are transparent and explicitly listed, students make use of these criteria to set strategies towards their goals. This is self-regulated learning. Students can also compare their work with the criteria and appraise the quality of their work. With appropriate training and the repeated use of rubrics, students can build evaluative judgement during the self-assessment process. Rubrics have also been found to aid the feedback process by allowing students to check and reflect on what is missing in their work; hence promoting learning. Current evidence points towards the powerful effect of rubrics in Assessment for Learning.

Most previous research has proved that rubrics reduce the anxiety of assessment because the expectations embedded in the rubrics decrease the fear of the unknown. However, whether this reduced anxiety translates into increased control and a higher level of self-efficacy remains unclear.

It is worthwhile noting that these merits are reported from studies mainly conducted in a Western culture or by teachers with a more sophisticated knowledge and experience with rubrics. It is therefore important to explore if the same merits hold true in Hong Kong, where the relationship between Assessment for Learning and rubrics is expected to be less well-known.

In addition, perceptions from the management level appear to be lacking in the literature. In the present study, it is imperative to understand management's perceptions of rubrics because they play an important role in policy implementation. This study attempts to explore these knowledge gaps which are unique to the study context. The next section focuses on criticisms of rubrics. Knowing the drawbacks of rubrics allows for a more holistic and critical understanding of this practice, as well as providing a more practical explanation of the challenges of utilising these in higher education.

3.3 Criticisms of Using Rubrics in Assessment Practice

Despite the number of merits of rubrics reported in literature, rubrics are not without criticism. Such criticism can be divided into development and implementation issues. Development issues concern practical challenges when creating rubrics, whereas implementation issues centre on difficulties with utilisation. This section discusses these drawbacks in more detail.

3.3.1 Rubric Development Issues

Developing a good quality rubric is not a simple and straightforward task. Van Helvoort (2010) states that it is very time-consuming to develop a rubric that works. This is in part due to task specificity and to teachers thinking they need to develop a rubric for each different assessment task (Sambamurthy & Cox, 2016). Also, detailing the assessment criteria and qualitative descriptions of these criteria is daunting because such criteria embed tacit knowledge (Carless, 2015a; Sadler, 2005).

Teachers need knowledge and time to create rubrics that are fit for purpose (Van Helvoort, 2010). With competing work priorities which are discussed in the previous chapter, spending time to create good quality rubrics may be given less priority by teachers as compared to tasks that could bring reward and recognition to their careers, not to mention the mastery of the skills required for such creation.

Another rubric development issue relates to validity. Teachers and students are sceptical whether rubrics can increase the validity of assessment. They also doubt whether the rubrics used are valid and accurately measure the intended learning outcomes. Rezaei and Lovorn (2010) suggest that rubrics do not improve assessment validity unless teachers undergo proper training in rubric design.

Specifically, the language used in rubrics is of particular importance because they need to communicate useful information (Reddy & Andrade, 2010). The criteria should be clear, instructional, and align with the learning outcomes set forth in the subject (Andrade, 2005). However, it is laborious to write out criteria that match all the requirements because of the tacit knowledge involved (Carless, 2015a). No clear evidence has proven the effects of rubrics in enhancing assessment validity (Jonsson & Svingby, 2007).

This is in part due to the nature of validity, that it is context- and task-specific. Proving validity in one rubric also does not mean the same in others, where the subjects and learning outcomes change. It reverts back to the point

of whether the particular rubric is designed to fit the purpose, and relevant design is the most time-consuming part of developing and using rubrics.

3.3.2 Rubric Implementation and Utilisation Issues

Not only is it time-consuming to develop quality rubrics, it is also labour intensive to effectively use rubrics (Van Helvoort, 2010). The utilisation of rubrics entails various dimensions such as how and when rubrics are presented and explained to students. Common ways include issuing or posting the rubrics on learning management systems for students to access by themselves before the assignment deadline (Carless, 2015a). Some, but not all teachers may supplement this with an explanation of the rubric criteria.

Regardless of these methods, using rubrics is suboptimal because multiple interpretations of criteria may occur and thus lead to different interpretations of the criteria (Andrade, 2005; Carless, 2015a). Even when the criteria are explicitly listed out, rubrics alone cannot replace good instructions (Andrade, 2005). Carless (2015a) critically reviews Hong Kong award-winning teachers' utilisations of assessment criteria or rubrics and comments that merely presenting criteria to students does not facilitate engagement with quality. There are several reasons for this criticism. Criteria are often too vague, repetitive and not specific to learning outcomes. For example, using generic descriptors such as fair, good, excellent, appropriate or critical, does not provide clear meaning to students on what they entail. Carless (2015a) suggests that "criteria need to be accompanied by activities which enable students to discern quality in the discipline" (p.149) and he suggests using

exemplars for this purpose. One can imagine the time and effort it requires to optimise rubric utilisation.

Furthermore, using rubrics for assessment has some inherent limitations because the criteria themselves are not straightforward and are difficult to present in words (Sadler, 2005). Take the complex thinking skill as an example; it is nearly impossible to assess this type of skill using written criteria because it involves multiple thought processes, which rubrics are unable to list (Bennett, 2016). The controversial part is that even if students are able to demonstrate complex thinking, they may be poorly evaluated because these criteria may not be listed in the rubrics (Carless, 2015a; Sambamurthy & Cox, 2016). This leads to another adverse consequence of limiting students' creativity: students may simply follow the criteria stipulated in the rubrics, and refuse to think beyond the standard requirements for fear of jeopardising their grades (Bennett, 2016).

Notwithstanding the challenges of interpreting criteria, other factors such as pre-conceptions of standards and hidden criteria also make rubric utilisation disadvantageous. Clearly listed criteria cannot prevent teachers from unconsciously exercising their subjectivity during grading (Carless, 2015a). Subjectivity in the form of a pre-conception of standards probably comes from past grading experience. Experienced teachers may enter into their pre-conception and be unable to extract themselves from this to achieve better objectivity when marking (Shay, 2005). Another factor hindering the interpretation of criteria is the belief in a hidden curriculum, which entails criteria not explicitly written out in rubrics. Students often speculate about the

use of a hidden curriculum when teachers grade (Norton, 2004), causing them to guess what these invisible agenda are and interpret the rubrics differently (Carless, 2015a).

3.3.3 Summary of Section and Implication

Two categories of the issues of rubrics surface while reviewing the relevant literature. From a development perspective, it is time-consuming and laborious to develop quality rubrics. For rubrics to be fit for purpose, the language used in rubrics needs to be specific and explicit. It does not only require a tremendous amount of time but also a great deal of knowledge to write these criteria. Without time and effort to design rubrics that match specific task requirements, rubrics may not reach the level of satisfactory validity teachers and students are seeking.

The utilisation of rubrics poses a bigger issue and this is mainly due to the variability of interpretation. Written criteria can be interpreted differently by different individuals and influenced by personal values and experiences. Interpretation poses a bigger problem if the criteria are vague. An additional difficulty is implicit criteria which are often embedded in an assignment. These tacit criteria are usually impossible to write out in words causing subjectivity and hidden agendas with grading. This is a major problem especially when assessing complicated skills. Writing out all criteria for complex skills is unachievable and even if they are written out, they limit the creativity of students and keep them operating 'inside the box'.

On the same note as the benefits of rubrics, it is very necessary to understand how less-experienced rubric users (in this study) perceive and experience the disadvantages of rubrics, as well as how these drawbacks may affect their utilisation. Since the use of rubrics by the majority of the participants in this study is driven by a top-down policy, it is also interesting to see if criticisms or barriers other than those reviewed in the literature can be identified. On the other hand, could some factors be minimised since there *should* be more resources with a policy-driven initiative? These questions are specific to this context and help to explain the actualisation of the rubric policy.

The next section focuses on selected successful examples of rubric actualisation in order to critically explore the factors that lead to success. The lessons learnt from these examples are crucial for understanding and comparing how the participants in this study perceive the optimal implementation of rubrics to facilitate learning.

3.4 Innovative Methods to Optimise Actualisation of Rubrics

Selected examples that demonstrate ways to overcome rubric implementation barriers are reviewed in the following subsection. It is important to note that not all barriers have a solution found in the literature but the examples below provide innovative intervention targeting the major issues of rubric utilisation. The knowledge of effective practice allows for a comparison of findings with the present study.

3.4.1 Innovative Examples

The first example is the study by Fraile et al. (2017) where they investigated the effects of co-creating rubrics with students on self-regulation and self-efficacy. The rationale behind this intervention is that co-creation could possibly increase the autonomy and sense of belonging of students in the use of rubrics and decrease the perception that rubrics are only an assessment instrument.

The authors hypothesise that these benefits can in turn enhance self-regulation, self-efficacy and performance. By comparing students co-creating/using rubrics and students just using rubrics, the authors found that the co-creating/using group possessed partially higher self-regulation ability measured by the thinking-aloud protocol (verbalising what they are thinking), as well as performance in one assessment task. However, the self-perceived effects on self-regulation and self-efficacy were the same between the two groups.

The result is interesting in that the self-perceived and measured effects are partially conflicting. The authors argue that the measured effect is more objective and relevant to self-regulation because it is not affected by students' personal characteristics and awareness. Nonetheless, the results of this study demonstrate that co-creating rubrics is a feasible intervention to address some of the criticisms previously mentioned, such as issues with criteria interpretation and the development of evaluative judgement. Although rubric co-creation has the potential to target these issues, there is no doubt

that co-creation takes time and requires teachers' knowledge and beliefs to materialise. The intervention may not be successful if readiness is lacking. Readiness may be an issue in users new to using rubrics, such as those in the present study.

Jones et al. (2017) created a 5-step rubric pedagogy that targeted various rubric issues, for example, unclear criteria and interpretation mismatch, lack of evaluative judgement and inability to act on feedback. The 5-step pedagogy that spanned 13 weeks involved students deconstructing the rubrics by clarifying ambiguous terms with teachers, reviewing examples and exemplars of work, performing peer review using rubrics, self-assessment using rubrics, and engaging in feedback by listing their strengths, weaknesses and areas for improvement.

Results showed that this intervention was useful and it increased students' confidence in understanding future assessment criteria. The authors also proved that this intervention enhanced students' performance in terms of assignment grades. The findings of this study are exciting; however, details of how this intervention has helped students are missing due to the quantitative nature of the study. Qualitative comments are available from extracts of students' reflective diaries but they are without systematic analysis. In addition, while the 5-step pedagogy is scheduled to take 30 minutes each time, the actual amount of time and effort spent on actualising it is suspected to be longer due to the preparation and follow-up involved. It is probably worse for inexperienced teachers. Time and readiness once again may pose

a challenge in using this pedagogy. The present study uses a qualitative approach to investigate these challenges in novice users.

The two aforementioned examples show innovative and effective ways to address some rubric utilisation issues; however, the strategies may not always be feasible due to time constraints. Broadbent et al. (2018) tackled the problem of time by using exemplars and audio feedback together with rubrics in large class assessment activities. In addition, moderation and training were provided to teachers to enhance the feedback quality and marking consistency. Students were satisfied with these strategies and thought they promoted motivation and understanding of the assessment criteria.

The authors argue that audio feedback is time- and cost-effective as compared to written feedback, although time is needed initially for enforcing this practice. Nonetheless, this is one of the few studies that provides suggestions for a more time-efficient and less laborious practice. It would be more insightful to understand how audio feedback and exemplars complement the use of rubrics, which is lacking in their study. However, this is also a voluntary initiative administered by teachers with a sophisticated knowledge of rubrics, which is different from the participants' background in the present study.

A very recent proposal to transform the utilisation of rubrics is by using the concept of invitation (Bearman & Ajjawi, 2019). The authors suggest three major 'invitations' – producing new descriptors in a collective productive space, enacting the descriptors repeatedly, and reflecting on/comparing

performance with the descriptors. Students' involvement is crucial in these 'invitations' with the goal of engaging them as owners rather than passive users. This idea of metaphor is innovative and perhaps essential to tackle some of the inherent issues of rubrics. However, the idea has yet to be actualised.

3.4.2 Summary of Section and Implication

This section reviews studies that offer a fresh perspective to address some of the rubric actualisation issues, namely unclear interpretation of criteria, ineffective engagement with feedback and underdevelopment of evaluative judgement. Co-creating rubrics with students and using various means such as exemplars and audio feedback together with rubrics are ways to optimise actualisation. Transforming the use of rubrics from only providing transparency to inviting engagement is another mindset shift.

Although innovative and proven to be effective, adopting these interventions requires time, knowledge and practice. This is especially true for inexperienced rubric users. As such, how do novice users and perhaps involuntary users at different stakeholder levels envision the optimisation of rubrics? Would their thoughts be comparable to those published in the literature? In this research context, since rubric utilisation mainly comes from a mandatory policy, perspectives from management are also important to analysis adoption. This research aims to address these gaps.

3.5 Summary of Chapter and Implications for this Study

This chapter critically reviews the benefits and criticisms of rubrics and contemporary examples of rubric utilisations. While the grading benefits of rubrics are rather visible to stakeholders, the learning function is not automatic and students need to engage with the criteria and learn how to reflect on feedback in order to maximise this merit. In addition, good rubric practice is limited by development issues and the inherent drawbacks of rubrics. While research continues to suggest various methods to address these issues, their adoption may require a more advanced understanding of rubrics and assessment as a whole.

Despite a number of studies conducted on various aspects of rubrics worldwide, little is known about how Hong Kong academics perceive and use rubrics. There are few examples in the literature, but information on teachers' perceptions of the functions, usefulness and utilisation of rubrics is missing. This is particularly important in this study context, since teachers in this university are explicitly required to adopt the use of rubrics as a result of a mandatory policy. Understanding in-depth how various stakeholders perceive and engage in the use of rubrics can explain policy implications in this context and shed light on similar practice in other parts of the region.

Management, teachers and students are stakeholders in this study. Linking back to the first part of this study on assessment practice, the stakeholders' stance on Assessment for Learning may have informed how they perceive and engage in rubrics. As such, this would add to the

explanation of their rubric interpretation and its implementation driven by a policy.

Chapter 4: The Current Situation of Assessment Policy in Higher Education

This brief and last literature review chapter focuses on assessment policies in the higher education context. Research on assessment policies in higher education, as compared to assessment practices, is still at a rudimentary stage and most studies are in the area of reflecting on policy implications (Ashwin & Smith, 2015). Since the present study looks at stakeholders' perspectives on policy interpretations and suggestions, only relevant literature surrounding this area is reviewed.

To begin, the background on current assessment policy initiatives and intentions is introduced. The adoption of assessment policies in higher education is then discussed. This information provides insight into the current assessment policy situation within the higher education context worldwide. The chapter finishes with suggestions from the literature for more effective and fit-for-purpose policy implementation strategies.

The knowledge generated from this literature review is crucial – it helps to identify how the rubric policy initiative in this research context is compared to other assessment policy situations worldwide, and how the suggestions from participants' experiences with the rubric policy add to the current body of literature in assessment policy both locally and globally.

4.1 Intentions of Assessment Policies in Higher Education

Policy is usually not a welcoming word. For some people, policy may carry a tacit meaning of an authoritative mandate from higher up. It also implies extra

work or cumbersome processes that policy actors or policy subjects need to actualise (Ashwin, Deem, & McAlpine, 2016). As negative as it may sound, using assessment policy in higher education does provide some positive intentions. At a fundamental level, policy can be used to raise awareness of a particular concern and to change an existing practice, such as transitioning from a conventional to an evidence-based practice (Tijds, David, Vaes, & Kerckhofs, 2012). For students, an assessment policy on high stakes performance standards has the potential to boost their self-regulation and performance because they regard it as an interesting challenge (Kickert, Stegers-Jager, Meeuwisse, Prinzie, & Arends, 2018).

The above two examples are discipline-specific policy intentions – for practice change and for performance improvement. What about the intentions of department- or institution-wide policies? Lambrechts (2015) discusses using audits on sustainable assessment for building policy in this area. The ultimate purpose of developing such policies is for quality assurance at the meso-level and for accreditation purposes at the macro-level. From an even higher level, the government can mandate education policies (including assessment) for higher education sectors for overall quality assurance and funding allocation (Sagarra, Mar-Molinero, & Rodríguez-Regordosa, 2015).

Quality assurance appears to be a common intention for policy initiatives and enactment at an institutional or system-wide level. Boud (2007) reviewed policy statements from various universities worldwide and mentions that the primary focuses of assessment in higher education are the measurement of learning outcomes and quality assurance. One can

reasonably believe that assessment policies at an institutional level intend to serve the *Assessment for Certification* function, with a trivial and subtle aim of promoting Assessment for Learning (if any). The next section looks at the current picture of assessment policy adoption.

4.2 Reasons Limiting Enactment of Assessment Policies in Higher Education

The enactment of assessment policies in higher education is not at the level it should be. It is found that the relationship between a faculty's actual practice and its assessment policy is trivial (Cox, McIntosh, Reason, & Terenzini, 2011), implying that policy exerts little influence on practice. Why is policy enactment limited? Experts have opinions and concerns regarding existing assessment policies, especially those coming from a top-down approach. These criticisms include policies being not holistic and focusing only on selected aspects of assessment, as well as their questionable effectiveness. These comments are discussed below.

As previously mentioned, Boud (2007) reveals that the majority of assessment policies in higher education centre on the measurement function of assessment. That is, they explain the purpose of assessment as measuring learning outcomes but rarely mention the learning function of assessment. This type of policy conveys a strong message to management and teachers, and even students, parents and industrial counterparts that the most important purpose of assessment is *certification*. Experts who advocate Assessment for Learning consider this type of assessment policy as skewing

the holistic picture of assessment and minimising the importance of learning within assessment, when it is proven to be good practice (Boud, 2007; Carless, 2017). Even when elements of Assessment for Learning are mentioned within the policies, they are rather subtle and secondary (Boud, 2007). Structures for executing the learning function are often lacking (Carless, 2017). In Hong Kong, the use of Assessment for Learning is reported as an individual approach rather than a programme- and institutional-wide approach (Ewell, 2006), which suggests holistic policies may be lacking locally.

Another key ingredient of a holistic assessment policy concerns the well-roundedness of policy content. Current policies usually include comprehensive explanations of assessment types and procedures (Meyer et al., 2010); however, the principles guiding assessment practice are often scarce (Fletcher, Meyer, Anderson, Johnston, & Rees, 2012; Meyer et al., 2010). Theoretical concepts of evidence-based practice of assessment is missing and policies merely serve as an operational manual for teachers and management to follow. This is dangerous as teachers and management may not have the knowledge to critically understand the best assessment practice, and rely upon institutions to provide such a background. When only procedures are included and theoretical underpinnings are undermined, assessment practice may turn into a technical skill application rather than an outcome of a critical thought process.

The questionable effectiveness of policies is another likely reason for limiting policy enactment. Cox et al. (2017) find that policies of using

assessment data to inform decision-making merely improve students' experiences in learning (when they are supposed to). This finding suggests that assessment policies are not effective and accountable in what they are intended to do. In addition, assessment policies create confusion for teachers because of the lack of compatibility between the two targets of mastery and scaling of grades (Meyer et al., 2010). On one hand policies stress the importance of linking assessment and learning objectives; on the other hand, policies require teachers to scale grades for a more normal distribution. This conflict forces teachers to compare students with other students, instead of assessing students as to whether they have met the learning objectives.

All in all, assessment policy and assessment practice seem to be disjointed due to the tension between the two aims and the lack of a holistic policy. Relating back to the present study, since the rubric policy is new, it is interesting and imperative to explore how the participants interpret and judge this top-down policy. Their perspectives may help to explain rubric actualisation in this context (research question 3).

It is beyond the scope of this thesis to evaluate or analyse the adoption picture of this rubric policy; however, understanding participants' perspectives on the policy provides insights for policy adoption, refinement and future suggestions of similar initiatives. The next section looks at recommendations for assessment policies in the higher education context, with a particular focus on building holistic assessment policies.

4.3 Recommendations for Holistic Assessment Policies

Given the lack of connection between policy and practice, various suggestions appear in the literature to tackle this issue. This section reviews these suggestions critically and appraises how they fit into this study context.

To formulate a holistic assessment policy, the first step perhaps is to return to the fundamental yet complex nature of assessment in higher education. Meyer et al. (2010) suggest using the Tertiary Assessment Grid to map assessment purposes and dimensions thereby identifying what policies and regulations are lacking and needed. The dimensions they stress are manageability, validity, equity and integrity of assessment for both the *learning* and *certification* purposes. While the framework is comprehensive and provides a well-rounded view of the nature of assessment in higher education, having a regulation or policy in each of the mapped area may be too heavily-loaded and may not be feasible for institutions, especially those that are new to developing assessment policy (i.e. the institution in this current study). In fact, too many policies have been criticised as being overloaded and over-exhaustive (Sindelar & Rosenberg, 2000).

Reflecting on their sustainable assessment policy initiative, Lambrechts (2015) recommends a few lessons for future policy initiatives. Instead of a top-down approach, the author suggests drafting policy vision and planning from the individual and departmental (micro) level and proceeding to the meso- and then macro-level. This bottom-up approach does not stop there; rather, it has

to reintegrate with the quality assurance framework of the institution and be enacted by senior management so as to come back down as a policy.

This bottom-up then top-down approach has the merit of enhancing buy-in thereby engaging teachers for policy adoption. This is perhaps a great lesson that the present context can learn from; however, assessment policy vision requires the assessment literacy of teachers and middle management, which may be a challenge at this stage in this institution. It implies that this approach may not apply to the local context.

In addition to having a policy draft from the faculty, policy content can also come from students as they are on the receiving end of any educational policy and are therefore an important stakeholder group. Poth, Riedel, and Luth (2015) solicited students' views on assessment challenges in their higher education institution and compared views with an assessment policy draft. The aim was to map out policy content that is meaningful to improve students' assessment experiences, such as the principles of fair assessment and the regulations to strengthen the feedback process.

The authors demonstrate a way to generate a more learner-centred policy, the merits of which are likely to increase both students' and teachers' co-operation and adoption of the policy targets their needs as compared to fulfilling an institutional agenda. The authors also provide an appealing idea to adopt a holistic assessment that is context-specific (i.e. addressing the specific needs of that institution). This suggestion of involving students is in fact advocated as an institutional strategy. The term "students as partners" is

used with the notion that students are more than customers in higher education but are part of the transformation team through communication, dialogue and community (Gravett, Kinchin, & Winstone, 2019).

The above examples shed new insight on effective policy approach and design. The role of the middle management group is not to be undermined because they are influential in policy enactment. Middle managers view themselves as “gatekeepers” and “translators” in the policy enactment process (Saunders & Sin, 2015), signifying tensions in their roles because they are close to both senior management and frontline academics. On the one hand they are responsible for ensuring that policy implementation proceeds according to senior management’s wishes, whilst on the other hand they need to convince stakeholders of the buy-in and adoption of the policy by supporting the implementation process. It is interesting to explore whether the management participants in the present study also experience the same tensions in this top-down rubric policy.

Finally, to ensure the quality of policy adoption, it is indispensable to review the outcomes of any policy initiative, in order to examine if the policy objectives are met. A model for policy review is useful for providing structures to look at essential aspects of policy enactment (Harvey & Kosman, 2013). Strengthening research into assessment policy helps to fill the gaps of what is missing, yet at the same time is pivotal in policy design and implementation (Jones, 2014). Although it is not the goal of this study to focus in-depth on policy design, the perceptions of teachers and management participants of

actualising the rubric policy identify additional insights and implications for future policy initiatives.

4.4 Summary of Chapter and Implications for the Study

Literature on higher education assessment policy is limited. The available information reveals that assessment policy is mainly used for quality assurance at the macro-level. At the micro-level, assessment policy can raise awareness of good assessment practice as well as promoting self-regulated learning. Nevertheless, it is most common to see policies coming from an institutional level.

It is interesting to learn that the relationship between assessment policy and practice is incompatible, meaning that practice does not always follow what a policy stipulates. Several factors explain this policy-practice disconnection. First of all, assessment policy is often criticised as not being holistic. Most assessment policies focus on the *certification* function but undermine the *learning* purpose. As such, structures to bring out the learning purpose is lacking. Secondly, assessment policies usually lack an explanation of what assessment is, but often just focus on the technical and procedural arenas of assessment. Policy enactors may not understand the principles behind good assessment practice. In addition, policy and expected practice may be contradictory, leaving teachers with conflicting information on how to approach assessment. These factors hinder the adoption of assessment policies.

To reconnect policy and practice, the literature suggests adopting a framework that caters for various assessment functions and dimensions to increase the well-roundedness of the policy. In addition, policy adoption is more effective if it comes from a bottom-up initiative and reintegrates into a top-down approach. Including students' views in drafting policy is also advocated. Lastly, middle managers carry an important role in mediating policy enactment.

The interest of this study is to explore how teachers and management in this higher education institution interpret and experience the rubric policy, so as to explain the rubric actualisation picture and to provide insight for future policy initiatives. This rubric policy addresses only one part of assessment practice and is a top-down initiative. It would be interesting to find out if the participants share the same perceptions as previously listed in the literature or if there are new perceptions identified.

To date, no previous studies have investigated stakeholders' views on policy enactment in Hong Kong. The findings of this study will add to the body of literature to enrich understanding of policy perceptions in this context because it is culturally-specific. In addition, by linking the bigger concept of Assessment for Learning and rubric actualisation as a result of a mandatory top-down policy, new relationships or conceptualisations between assessment policy and practice may be identified.

Chapter 5: Methodology and Methods

To recap, the overarching aim of this research is to investigate the perceptions and actualisations of Assessment for Learning and rubrics of students, educators and middle management under a mandatory top-down rubric policy in one university in Hong Kong. Policy implication is also explored. As such, the research questions of this study are:

- RQ1. What are the perceptions and understanding of Assessment for Learning in students, educators and management?
- RQ2. How do students, educators and management identify barriers and opportunities for the adoption of Assessment for Learning?
- RQ3. Can rubrics actualise Assessment for Learning in the eyes of students, educators and management? Why or why not?
- RQ4. What are the policy implications according to the experience of educators and management in this policy initiative?

To adequately address the research questions, an in-depth exploration of participants' experience and their perceptions of assessment practice, rubrics and the rubric policy is required. This chapter discusses my position as a researcher, the research paradigm, methodology and methods of collecting and analysing data that are appropriate for addressing the research questions. This structure follows the research process suggested by Denzin and Lincoln (2005) and is deemed to adequately cover information needed to justify the current research study. The chapter ends by addressing the limitations of methodology and methods. Understanding limitations allows the researcher to be more reflective during the entire process.

5.1 The Researcher's Position and Ethical Considerations

This section discusses my position as a researcher in this study as well as the ethical considerations that I anticipated.

5.1.1 Position of the Researcher

The position of the researcher is important as it may influence various parts of the research process, for example, the interpretation of narrative during interviews and/or data analysis (Atkins & Wallace, 2012). While planning the research, I considered this study as insider research, which by broad definition means that the research is carried out at one's working institution and that the researcher understands the culture/language of the institution (Smetherham, 1978). However, Atkins and Wallace (2012) state that a researcher's position should always be on a continuum as there may be times where we are an insider and at other times an outsider. This can happen during the different stages of the research process when the researcher's role, position and/or relationship with informants evolve, in spite of the fact that the research is still conducted at the same institution. This was clearly my case.

I positioned myself as an insider researcher at the beginning because of my familiarity with the institution. I had worked as an academic in the university since June 2012 and I viewed myself as knowing the culture and the university's stance on teaching and learning. Also, I had access to their policies and/or initiatives in teaching and learning which to a certain extent have shaped my research interest. In addition, as with other novice

researchers, conducting research in the working institution offers benefits in recruiting participants and ease of access to key information, as well as that of being a member of the research phenomenon (Atkins & Wallace, 2012; Mercer, 2007).

As the research progressed to the data collection phase, I started to realise some shifts in my position. Despite the fact that participants of all stakeholder levels were from within the university, my understanding of the information they had about various policies or processes was not as familiar to me as I had thought.

There are several reasons for such feelings. Firstly, most of the participants were from different academic departments where the departmental culture and operational system of teaching and learning were not exactly the same as in my home department. Secondly, participants were from different disciplines and there were some nuances in teaching and learning focus and priorities which I had not thoroughly considered beforehand. Third, as a frontline educator I realised I was not as in touch with students' and management's mindsets as I had expected due to role differences.

Together, these differences evolved and placed me more in an outsider perspective where I had to expose myself to some new language and meanings of the research phenomenon. This was slightly challenging as I had to make sure I understood the meanings of my participants' viewpoints. There were also more follow-up questions in some instances. Nonetheless, I do not

think I shifted completely from being an insider to becoming an outsider researcher. I was simply more aware of my shift in perspectives during data collection and adjusted accordingly during the process.

Perhaps the most remarkable change of position started after I left my full-time academic position in June 2018. Although I continued to work in the same university as a part-time member of a project during the data analysis phase, I felt more detached from teaching and learning when I was no longer a frontline educator and became a “passive recipient” of teaching and learning policies and initiatives.

This might have given me some advantages because I could be more objective when I was freed from any power-relationships (Atkins & Wallace, 2012; Mercer, 2007) that I would have otherwise experienced as a frontline educator. All in all, as Mercer (2007) mentions, there are no absolute pros and cons of being an insider or outsider researcher; I have experienced both along the continuum in different circumstances and the merit lies in being aware of each position in order to act accordingly.

5.1.2 Ethical Considerations

This research follows the standard practice of applying for ethical approval and seeking informed consent from all participants prior to data collection. Ethical approval is obtained from both Lancaster University and the university where this research is conducted. Informed consent is explained later under the data collection section (5.4.2).

Perhaps the biggest ethical consideration of this research lies with concerns associated with insider research. These considerations include role identity (e.g. educator vs researcher), boundary conflict, confidentiality, relationships and imbalance in power relations (Atkins & Wallace, 2012; Mercer, 2007). Possible measures used to guard against these challenges are discussed below.

First, confusion with role identity was unavoidable in insider research. During the data collection phase, I constantly reminded myself to disclose my identity to participants, as to when I was an educator and when I was wearing the hat of a researcher. I also exercised reflexivity to maximise the trustworthiness of data (Cousin, 2008), especially when the participants mentioned a point similar to my personal experience. These measures safeguarded boundary conflict when collecting data, and reminded me what my boundary was in each of my roles.

For issues pertaining to confidentiality, participants' identities (e.g. name, gender, departments) were kept strictly confidential and participants could withdraw at any time during data collection if they were not comfortable to continue. I also practised data security and transferred all audio files to the Lancaster Box within three days of data collection, and deleted them from the recording device. This would safeguard against the risk of data leaking.

Issues arising from the insider relationship may happen during the research. Being close to the participants has its pros and cons – they may either feel reluctant to share due to my knowledge of the university system or

they may offer favourable opinions within an overly-rapport relationship (Atkins & Wallace, 2012).

At the beginning of each data collection session, I mentioned to the participant again that I valued their honest opinions and data would only be used for my research, in order to minimise the effects of relationship issues. Also, I shared my knowledge and evidence on Assessment of Learning and rubric practice based on the literature. This might have strengthened participants' intellectual knowledge on the topic and therefore optimised the credibility of data.

Power imbalance is common in insider research (Atkins & Wallace, 2012; Mercer, 2007). With the design of this study, management and senior educators may impose a hierarchical power onto me; whereby I may be seen as more powerful than junior educators and students. To address this potential issue, I employed the measures suggested by Cousin (2008): to disclose personal experience to build up trust between interviewer and interviewee and be reflexive during the process.

Although these steps might not completely alleviate the issues, they at least minimised the problems affecting honesty of opinions. The students from my home department were informed that their opinions would not be counted against their grades or other academic judgment. Students from other academic departments were reassured that their opinions would not be shared with their teachers. These measures should have increased the chance of honest opinions from all participants.

In brief, the ethical considerations of this study mainly stemmed from my insider image and could affect the trustworthiness of data. Measures such as informed consent, reflexivity and stating the theoretical basis of the study helped to enhance trustworthiness. The next section of this chapter discusses the research paradigm of this study, which allows further understanding of the details of this research.

5.2 The Research Paradigm

Research paradigm is an overall theoretical research framework (Mack, 2010) that orients one to approach a research study. It provides assumptions, conceptions and stances to the researcher (Mack, 2010). These elements are important underpinnings of how the researcher approaches the research questions, and guides the subsequent processes of methodology and methods. It is generally accepted that there are three key research paradigms – positivism, critical-realism (or post-positivism) and interpretivism (Grix, 2010).

The basis of the positivism paradigm lies in realism ontology where reality naturally exists and is independent of people's knowledge (Guba & Lincoln, 2005; Mack, 2010). As such, individual and social influence is minimal. Positivists mainly seek to explain a phenomenon; they believe in establishing a causal relationship between objects or predicting an observation (Grix, 2010). The positivism paradigm focuses on empirical evidence rather than personal values and beliefs (Grix, 2010; Guba & Lincoln, 2005).

At the other end of the spectrum is the interpretivism paradigm where relativism dictates reality (Guba & Lincoln, 2005). Interpretivists believe that knowledge is individually and socially constructed, thus subjectivity is normal (Mack, 2010). It is accepted in the interpretivism paradigm that multiple interpretations exist and people can have their own meaning and interpretation of an experience or observation (Yilmaz, 2013). In contrast to positivism, the interpretivism paradigm is mostly interested in understanding a phenomenon, for example why things occur (Grix, 2010; Mack, 2010). The emphasis of this paradigm is on personal values and beliefs rather than on empirical facts (Grix, 2010; Guba & Lincoln, 2005).

The post-positivism paradigm lies in-between positivism and interpretivism as it looks for explanation as well as understanding (Grix, 2010). The ontological belief underpinning post-positivism is critical-realism, which allows for a broader view of facts including the social context. The post-positivists believe in the influence of social powers in determining a causal or predictive relationship (Grix, 2010). The emphasis is placed on both and results are interpreted together. Despite the distinctions between the three paradigms, the boundary is not fixed. The positivists also aim to understand and the reverse holds true for the interpretivists (Grix, 2010). It is the focus of the research that dictates the research paradigm.

This current research employs an interpretivism paradigm because of my personal epistemological belief that knowledge is constructed through individual and social experience. Also, I am interested in understanding, not only listing, the perceptions and experience of Assessment for Learning,

rubrics and the rubric policy of the participants in this university. With this belief and interest, knowledge needs to be constructed through their personal experience and interpretation of these experiences. I value vicarious experience because the diversity allows for rich co-construction of knowledge (Guba & Lincoln, 2005). I also value authentic experience because it will make the results more trustworthy within the context. In contrast, this study will not attempt to explain any relationship or prediction. Also, I do not intend to verify pre-set hypotheses to establish facts. Hence, the positivism and post-positivism paradigms are not suitable to guide the research processes of this study.

5.3 Methodology

This study employs a qualitative design to gather in-depth understanding of Assessment for Learning, rubrics and the rubric policy of this university from students, educators and managers. Qualitative design is concerned with interpreting subjective experience (Grix, 2010). It involves in-depth investigation of knowledge (Grix, 2010), and focuses on lived experience placed in its context (Tracy, 2013). It is used to explore a question where the variables are unknown and the data gathered are primarily in words (Creswell, 2012). The aim of this study corresponds to the description of qualitative design, where the conceptions of Assessment for Learning, rubrics and rubric policy among the participants were unknown at the beginning of the research, and where this knowledge was gained through gathering and analysing participants' lived experience and interpretations of assessment practice within their context.

The specific type of qualitative methodology that this study uses is phenomenology through narrative inquiry (Elliott, 2012; Tracy, 2013). The phenomenological approach focuses on individual, conscious lived experiences and how they interpret or represent presence in their experiences (Giorgi, 1997; Starks & Brown Trinidad, 2007). Researchers delve into their experience and thoughts of “being” (presence), and come up with thick descriptions and particular meanings rather than vague explanations (Giacomini, 2013; Giorgi, 1997). This interpretation of experience is context-specific and provides researchers with an understanding of how participants interpret a situation or a phenomenon of interest and create meaning of that particular situation or phenomenon (Starks & Brown Trinidad, 2007).

Qualitative phenomenological design is an appropriate design to address the research questions in this study. First, I am interested to explore participants’ lived experience and perceptions on Assessment for Learning, rubrics and their interpretation of the rubric policy (the phenomenon of interest). In addition, not only do the participants share their viewpoints and experience, but they also rationalise what the three objects mean to them in real life situations. The meanings are specific and unique to their own experience and perceptions. Only rich contextual information and close analysis of rich information such as this can address the research questions sufficiently.

Narrative interview is the most common form of phenomenology where participants verbally talk about their lived experiences and explain their interpretations of them (Elliott, 2012). This is the type of phenomenology this

study uses. It is deemed appropriate because: (1) participants all possess conscious awareness of the experience of interest (Giorgi, 1997) and (2) participants in this study are all adults and have the ability to express themselves verbally. Also, due to the fact that the participants should carry out most of the talking, narrative interview is thought to have the benefit of empowering participants (Elliott, 2012).

However, Elliott (2012) listed four situations where the narrative process may vary, despite asking relatively similar questions. The four situations are:

- communication may differ in different cultural contexts;
- communication between different researchers and participants may vary;
- the occasions may influence certain narratives; and
- the actual performance of the narrative process.

Despite the fact that I am the only researcher in this study, it is important to be aware of these situations as they may affect the quality of the dialogue and therefore the empowerment of the participants may not always happen. Nonetheless, participants' ability to articulate and the researcher's abilities to interpret and reflect are the commonest challenges of this qualitative methodology (Friesen, 2012; Goulding, 2005).

5.4 Methods

It has been mentioned thus far that this study adopts an interpretivist stance and employs a phenomenology methodology to investigate participants' lived experience and the meaning of Assessment for Learning, rubrics and the rubric policy. This section discusses sampling, data collection and the analytical methods of this study.

5.4.1 Sampling

Purposive sampling was used in this study to recruit student, educator and management participants. Purposive sampling is a common sampling method in qualitative research and it means that the participants are specifically selected based on the researcher's belief in their unique contributions (Creswell, 2012) and experience with the phenomenon of interest (Starks & Brown Trinidad, 2007). This sampling technique is commonly used in qualitative research and this actually makes the research credible as it would allow what the research is intended to explore (Saldana, 2011). Since this research was conducted in the researcher's working institution, the participants invited were also chosen out of a convenience pool, to minimise time and financial burden (Tracy, 2013).

Educator and management participants were invited by email based on our professional relationships or through professional connections, and my knowledge of their willingness to share information on the research topic. Attached in the email invitation were the cover letter and the participant information sheet of the research. This allows participants to understand the

purpose, risks and benefits of the study before committing themselves to participate.

During the sampling process, participants' demographic backgrounds were also taken into consideration; efforts were made to balance out participants' ages, gender, discipline/academic departments and years of experience in higher education. That said, the invitation was not restricted to only educators and management with expertise in assessment practice or the use of rubrics. No other recruitment criteria were set because all educators and management presumably had experience with assessment, rubrics and had encountered the rubric policy in the capacity of this university's employees. In this way, a broader understanding of topics at the university's level could be obtained.

Based on the above descriptions, fifteen educators were recruited, one of whom served as a pilot to test and ensure the quality of the data collection process. Out of the remaining fourteen educator participants, five also performed a concurrent management role at either departmental or faculty level. Detailed demographic information is provided in the next chapter under Results.

For the student participants, one group of students (n=5) from my department was invited by email. Similarly, the cover letter and participant information sheet were provided beforehand. Students were all from the same cohort (Year 2 of their study) and had some experience with using rubrics in a subject that I taught. The selection of students was based on my observation

and interaction with them in class, knowing that they shared similar attributes of expressing opinions. Following the same principle of representativeness within the university, another group of students (n=5) who were in their senior years, was recruited from a different academic department. This group of students was invited by one of the educator participants, and his selection of students was based on the same “willingness to share” principle stated above. As originally planned, a third group of students from yet another department would be invited into this study for representativeness. However, the recruitment was not successful despite efforts made to achieve this.

Depending on the research, the sample size normally ranges from one to ten participants in a phenomenological study (Starks & Brown Trinidad, 2007). However, some larger scale studies have had sample sizes of over 80 (Tight, 2012). As Saldana (2011) states, the “enough” number of participants depends on many factors and there is no one clear standard in qualitative research as compared to quantitative. As long as there are sufficient data to address the research questions, the number of participants is not the main concern. During the data collection process, I discussed with my supervisor about the number of participants sampled/interviewed, and periodically reflected and reported on the data collection. This discussion led us to confirm data richness for sufficient analysis.

5.4.2 Data Collection

Before starting data collection, the purpose and details of the study were explained to the participants. The procedures of data collection, volunteer

participation and data management were specifically elucidated to the participants to ensure comfort in participation. The consent form was signed by each participant signifying agreement to participate.

The method employed to collect qualitative data from educator and management participants was semi-structured interviews. Semi-structured interviews are done on a one-on-one basis, where the interviewers set specific guiding questions (refer to Appendix 2) before the interviews; however, the questions asked during the interview do not need to follow a specific order (Grix, 2010). This allows for flexibility of unexpected lines of enquiry, that is, researchers can dig deeper on specific points according to the comments of the interviewees, the comments that arouse points of interest to the researcher or those that align with the main focuses of the research (Grix, 2010; Murphy, Dingwall, Greatbatch, Parker, & Watson, 1998). In other words, semi-structured interviews allow the researcher to enter into conversations with the interviewees, which the researcher has a genuine interest in (Tracy, 2013). During the semi-structured interviews, the interviewers actively listen and reflect so that complex understanding can be achieved.

Semi-structured interviews were suitable for these two stakeholder groups because they allow participants to share their views privately and in-depth, without the influence of others (Saldana, 2011). Since participants were asked to share their opinions on the university's practice and policy, sensitive or negative comments were expected to arise during the interview. If

the interviews were done in groups, participants might have been reluctant to share their honest opinions.

Also, as mentioned before in section 5.4.1 Sampling, participants were recruited from different academic departments within the university, with different numbers of years of experience in their respective roles. This heterogeneous background might have led to a very different experience of assessment practice and their interpretations on university policies (i.e. a different phenomenon) (Forsey, 2012). Therefore, putting participants in groups was not ideal to gather rich information on the topic.

With the nature of this topic and the research questions, it was also not feasible to collect data using other methods, such as observations. There were many assessment types and events that participants experienced at different time points in their career. Observing various assessment practices would require a tremendous amount of time and effort. More importantly, observing assessment events would not provide information on how educators feel about the assessment itself. My research questions required participants to collect their thoughts and comprehend what those assessment events meant for them; as such interviews were the best way to gather the information (Forsey, 2012).

On the other hand, the focus group was the method of choice for collecting data from the student participants. As the name implies, the discussion is done in groups where there is a focus of topic. Focus group participants should theoretically share similar characteristics and there should

be interaction among participants to facilitate discussion (Robinson, 2012). The interaction part is crucial; rich information may not otherwise be available if the focus group is not used (Kamberelis & Dimitriadis, 2005; Robinson, 2012). This is the main reason why the focus group method should be used.

In this study, the focus group was selected over the semi-structured interview for student participants because participants in their respective groups knew each other and went through similar experiences of assessment and rubrics at around the same time in their educational journey. This created a group dynamic which allowed for a collective construction of meaning (Kamberelis & Dimitriadis, 2005). The group interactive synergies may even cultivate new interpretations among the group and shed new insights beyond what one's memory or perception confined (Kamberelis & Dimitriadis, 2005). Semi-structured interviews would have deprived the researcher of this opportunity and the information gathered would not be as rich; hence, a focus group was the most appropriate method for data collection among the student participants.

Each semi-structured interview lasted for about 1 hour and each focus group lasted for about 75 minutes. All interviews were audio-recorded using a portable recording device. The audio file was transferred to the Lancaster Box within three days of the interview and was deleted from the portable recording device. Some semi-structured interviews were conducted in Cantonese while some were conducted in English, depending on the participant's preference. One focus group was conducted in English while the other one was conducted in Cantonese. Again, this was because of the language preference

within the group. The interviews were transcribed either by the researcher or by external transcribers. In the case of an external transcriber being used, a transcriber confidentiality agreement was signed.

5.4.3 Data Analysis

The main goal of interview data analysis in qualitative research is to “reveal cultural contexts behind the lived experiences of the research participants; the portraits capture the beliefs, the values, the material conditions and structural forces underpinning the socially patterned behaviour of the person that emerged in the interview” (Forsey, 2012, p.374). Since data collected from interviews are rich, data reduction is necessary in order to capture and reveal values and beliefs within the phenomenon (Giorgi, 1997).

Giorgi (1997) provides five concrete steps for data phenomenological data reduction:

- “collection of verbal data;
- reading of the data;
- breaking of the data into some kind of parts;
- organisation and expression of the data from a disciplinary perspective;
- and
- synthesis or summary of the data for purposes of communication to the scholarly community.” (p.245).

In addition, Yilmaz (2013) reinforces the idea that ongoing analysis of data is essential in qualitative approach because “the bottom-up approach to data

analysis with open- coding strategies should be practised to allow themes and patterns to emerge from data” (p.317).

This study adopted the five steps approach to outline the themes. Collection of verbal data was described in detail in the previous section. Reading of data in the form of transcript was thoroughly done; however, only to gather a big picture of the story relevant to the research questions. Repeated reading of transcriptions allows for ongoing analyses if necessary, providing details and accuracy (Hammersley, 2012). As reading continued, data were divided into parts named ‘meaning units’. Meaning units “signified a certain meaning, relevant for the study, and to be clarified further, is contained within the segregated unit” (Giorgi, 1997, p.246). As the name implies, meaning units allowed me to discover meanings in the data as they emerged. These individual meaning units were later systematically organised into codes and categories to link them together (Starks & Brown Trinidad, 2007). Lastly, the categories were further reduced or synthesised into structures or themes for an explicit communication answering the research questions.

5.5 Limitations of Methodology and Methods

Every type of research paradigm and methodology possesses strengths and weaknesses; it is the awareness of these strengths and weaknesses that makes the interpretation of research trustworthy (Yilmaz, 2013). Strengths and justification of the appropriateness of employing the phenomenological qualitative approach via semi-structured interviews and focus groups are

discussed throughout this chapter. The last section of this chapter tackles the limitations of the methodology and methods employed in this study.

Perhaps the biggest weakness of any phenomenological qualitative approach lies in the subjective nature of the data analysis (Starks & Brown Trinidad, 2007). Subjectivity may arise from one's cultural background, disciplinary orientation and personal experiences (Yilmaz, 2013). These may induce personal bias during the interpretation of results. The researcher (I), being the major instrument of data analysis, needs to practice reflexivity and discard personal beliefs and pre-conceptions during the process (Starks & Brown Trinidad, 2007). Reflexivity helps to increase the credibility and trustworthiness of data; a lack of the two elements is always a major criticism of qualitative research.

In addition, I must be clear that data collected via interviews represent thoughts and experiences but not necessarily actions that happened (Tight, 2012). It was mentioned previously that the expression of thoughts and experiences in interviews can be affected by many factors including articulatory abilities. This itself poses another limitation of this study where data is gathered via interviews. Ethnographic studies can potentially bridge this gap; however, it is explained earlier why ethnography was not an appropriate choice to address research questions of the current study. To compensate, interview data are triangulated in this study – by using two interview methods and by involving three participant groups – to allow for the multiple perspectives heard in this study (Saldana, 2011; Yilmaz, 2013).

5.6 Summary of the Chapter

This chapter discusses the methodology and methods employed in this study. This study uses an interpretivism paradigm as a basis for the qualitative inquiry of the four research questions. Only through rich inquiry can these questions be addressed sufficiently. Phenomenology is the choice of methodology because this study is interested to explore the lived experience and the meaning of Assessment for Learning, rubrics and the rubric policy among the participants. Since these experiences do not constitute observable events, interviews allow participants to verbally describe and elaborate their thoughts and hence are the most appropriate method of data collection. Thematic analysis is the method for data analysis. Throughout the research process, I must constantly practise reflexivity in order to bridge some of the limitations with phenomenological qualitative design, to enhance the credibility and trustworthiness of this study.

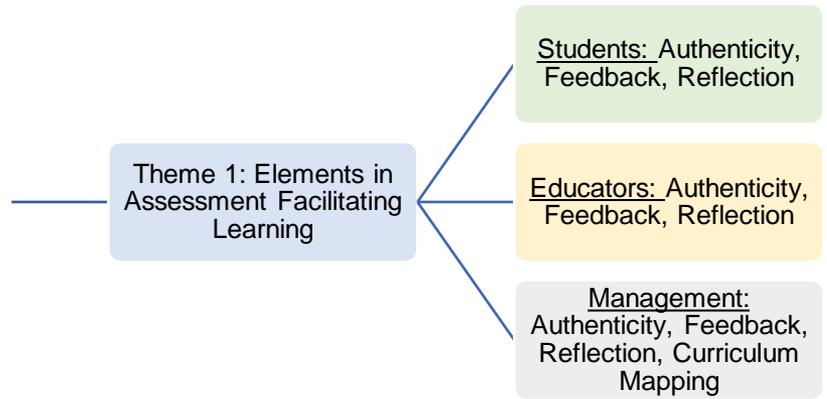
Chapter 6: Results of Assessment for Learning

This chapter is one of the two Results chapters that reports on the findings from focus groups of student participants and the semi-structured interviews of educator and management informants. Before describing the qualitative findings, this chapter begins with a brief summary of participant demographics to illustrate their diverse backgrounds. Next, findings organised in themes that address Research Questions 1 and 2, targeting the concept and the understanding of assessment, in particular Assessment for Learning, are presented. To recap, the first two research questions ask:

- RQ1. What are the perceptions and understanding of Assessment for Learning in students, educators and management?
- RQ2. How do students, educators and management identify barriers and opportunities for the adoption of Assessment for Learning?

The findings from the semi-structured interviews and focus groups identified three themes for these two research questions. Specifically, Theme 1: “Elements in Assessment Facilitating Learning” addresses Research Question 1. Theme 2: “Burden and Mismatch in Current Assessment System” and Theme 3: “ Ideal Actualisation of Assessment for Learning” address Research Question 2. Figure 6.1 below is an illustration of themes and sub-themes for the first two research questions:

RQ1. What are the perceptions and understandings of Assessment for Learning in students, educators and management?



RQ2. How do students, educators and management identify barriers and opportunities for the adoption of Assessment for Learning?

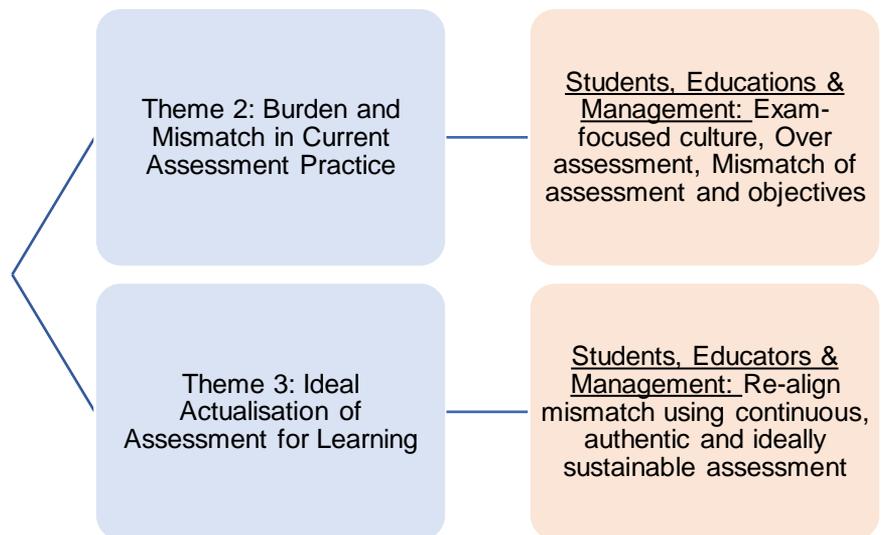


Figure 6.1 Illustration of Themes and Sub-themes for Research Questions 1 and 2

6.1 Demographics of the Participants

This initial and brief section serves to report on the demographics of the participants of this present study. Overall, fourteen participants from the educator/management group joined the study. Among them, 9 participants had a purely educator role at the time of data collection whereas 5 participants were in a management capacity in addition to their educator role. Since it is impossible to completely separate the educator hat for those with a management role, all of their opinions contributed to the educator group during data analysis. The participants with dual roles were interviewed with additional questions sharing their experience and perceptions as managers. As such, there are 14 sets of interviews for the educator group and 5 sets of interviews for the management group.

The fourteen educator/management participants (female=5, male=9) were from 11 departments within the university. Their years of experience as academics ranged from 5 to 30 years. Their years of experience in management ranged from 2.5 to 20 years. Their academic capacities spanned the entire spectrum of teaching and professorial tracks from Instructor to Professor, whereas the management roles spanned from Associate Head/Director to Associate Dean. Table 6.1 displays a summary of the participants' demographics. Individual departments are not listed due to sensitivity of information. The representing departments varied from health-related to business fields.

For the student participants, two groups of five participants joined the focus groups. The two groups were from two academic departments. One group of students was in Year 2 of their study and the other group was in Years 3 and 4 during the time of data collection. The demographics of individual students are not included.

Participant	Gender	Years of Experience in Academia	Years of Experience in Management	Capacity
Ed 1	F	15	NA	Assistant Professor
Ed 2	M	8	NA	Assistant Professor
Ed 3	F	10	NA	Clinical Associate
Ed 4	M	13	NA	Professor
Ed 5	M	5	NA	Teaching Fellow
Ed 6/M 1	F	20	3	Assistant Professor Programme Leader
Ed 7	M	30	NA	Associate Professor
Ed 8	M	7	NA	Teaching Fellow
Ed 9	M	8	NA	Research Assistant Professor
Ed 10	F	18	NA	Instructor
Ed 11/M 2	M	15	2.5	Associate Professor Associate Dean
Ed 12/M 3	M	26	20	Professor Head of Department
Ed 13/M 4	M	25	4.5	Associate Professor Associate Head of Department
Ed 14/M 5	F	22	10	Senior Lecturer Associate Director

Table 6.1 Demographic Information of Educator and Management Participants

6.2 Theme 1 – Elements in Assessment Facilitating Learning

Before investigating the participants' understanding of Assessment for Learning in particular, Theme 1 begins with reporting the first impression participants have when they hear the word "assessment". Not surprisingly, the most frequently used word to describe assessment, despite stakeholder levels, is "evaluation". Most participants think of assessment as evaluating students' study progress, learning, understanding and/or outcomes. Assessment is something needed in the education system to know if students have met a pre-set standard, competency or requirement.

“Assessment is a way to look at whether our students can achieve, like a subject or a program or a professional requirement...if they can achieve the standard or to a certain competency...” (Ed10)

Similar words, such as “measurement”, “judgment” and “competency checking” were used to describe assessment. Nonetheless, the underlying meaning is equivalent to the word “evaluation”. This evaluation is mostly done through assigning a grade or a score to the assessment so that students and educators make sense of where an individual student stands within a group of students.

“...Assessment is used to identify the differences within the whole cohort...categorise students into high and low group...” (Student FG1 S2)

This “categorisation” the student mentioned links to the concept of benchmarking which participants in the management group frequently referred to, although some educators and students are also aware of this benchmarking function. In the management group, apart from benchmarking, quality assurance is also a common function of assessment in their perception. Both benchmarking and quality assurance are important metrics higher education institutions across the globe need to be aware of, in order to raise institutional rankings for prestige and funding purposes. Management uses assessment results from individual subjects and departments to fulfil part of the benchmarking and ranking requirements.

Apart from the major function of evaluation, some participants also mentioned that assessment can provide insight for educators and students to inform about their own teaching and learning respectively.

“The other (function) is to see, what do you have to put in place as well in order to get them (students) there because sometimes it is for the teacher to know, ‘Oh I might not have explained it properly therefore I need to give them more help or more support or more information or whatever.’...” (Ed1)

It is unclear whether this quote illustrates the *certification* or *learning* function; it probably depends upon the depth and direction of individual reflection. Also, the subsequent actions taken after reflection likely distinguish the purpose of this function. If upon reflection teachers or students are able to

identify areas of improvement and take steps to change, then the function of assessment is to facilitate learning.

Despite almost all participants describing their first impression of assessment as evaluation, characteristics of Assessment for Learning evolved as they talked further about their experience of assessment. It is important to note that there is a high degree of resonance among participants across stakeholder levels in terms of the facilitatory elements for learning in an assessment task. This may indicate the commonality and importance of these elements in education despite different disciplines and capacities. The key characteristics are authenticity, feedback, reflection and curriculum mapping, which are described below.

6.2.1 Authenticity

Assessment that includes a scenario mimicking “real-life” skills or knowledge of that particular field is often regarded as authentic and as having a learning element. The actual format of this type of real-life scenario differs among disciplines. For example, educators from Health Science Departments viewed practical examinations or clinical performance assessments as real life. One student group from a Health Science Department also referred to case studies and practical examinations as being ways to study or assess skills required for real clinical practice. The participants from the Department of Law and Management stated that persuasive communication in presentations and essay writings is essential in their field. The participants from the Department of Engineering said building a product is an essential skill students must be

equipped with. Almost all educators thought team work and participations are essential attributes in any workplace and therefore regarded these assessments as real-life and crucial.

Regardless of the discipline, the description of real-life is similar:

“...they can apply the theory learnt...more practical” (Ed5)

“It is hope that students could apply, could translate efficiently if they encounter similar situations, it needs application.” (Ed11M2)

Application seems to be the key in authentic assessment. Students need to translate the skills and knowledge learnt in classroom into real-life situations after they graduate from their respective programmes. Authentic assessment focuses on this aspect rather than the regurgitation of knowledge; hence this type of assessment is regarded as facilitating learning.

6.2.2 Feedback

The authenticity of assessment tasks is only one of the elements the participants mentioned where learning is embedded in the assessment. In fact, the assessment task is an object of the system but how assessment is being used or actualised is an art and should be the key focus. The educator participants shared the view that providing feedback and having students to reflect on their performance are essential elements in these authentic assessments to help students learn. Specific to feedback, the term formative feedback is used by a number of educators. They refer to formative feedback as information given in an assessment task to help students improve and

grow. Most of them state that this is a type of feedback provided before the final grade or score is given in that particular assessment task. For example, Ed3 mentions midterm feedback during a clinical placement where students can use that feedback to work on areas of improvement. Another example shared by Ed9 is handing in a first draft of an essay for feedback before proceeding to the final graded version.

A more innovative way of feedback mentioned by Ed10 is to use in-class non-graded quizzes to assess students' knowledge on a particular topic, where students can discuss with peers after the first quiz attempt and before proceeding to the second trial. The results of in-class quizzes do not count in the final grade but are used as a way for students to learn from educators and peers.

“...they would do peer instructions afterwards (1st attempt), discuss the choice they chose...they will poll again after the discussion...I will then discuss the questions one by one, ask students to answer my questions...” (Ed10)

Apart from peers and educators, feedback can be obtained from other stakeholders essential to the authenticity of the task. In Health Sciences during clinical placements, patients often provide an abundant amount of feedback to students and educators on how students carried out a clinical examination or treatment. This type of feedback is complementary to learning as this is what students will receive in real-life.

A student (FG1 S4) shared their belief that midterm grades with feedback decreased anxiety in the assessment process. They used midterm feedback to improve their work before final submissions. Both groups of students concurred that the timing of feedback is important and it should be timely to allow time for modification of work. In addition, students shared the idea that feedback from teachers is more helpful as it helped them to understand expectations. Explicit yet non-judgmental feedback is instrumental to their learning.

“...even if our responses did not match his/her answers, he/she would still give us constructive feedback and insights of his/her thinking behind....to guide us.” (Student FG1 S2)

6.2.3 Reflection

To make good use of feedback for learning, reflection is crucial in the process as many participants mentioned (Ed1, Ed3, Ed7, Ed9 and both student focus groups). Reflection is a thinking process in itself, where an individual thinks about the situation and tries to reason out the process (Rodgers, 2002). Educators mostly use questionings to facilitate reflection in the assessment process, for example, “how do you think you did?”, “why did you do this?”, “what could be done differently?”. Reflection includes confrontation with others and is often used together with feedback in terms of reflecting on the feedback given (Miedijensky & Tal, 2016).

“There is actually activities where you have to reflect on certain tasks, and then you get confronted with some others....they had to post two

comments on other posts. The activity was such that you post your own reflection but then you gave two feedback to two other posts, so that it was not just ending with posting it but it was actually been retaken.” (Ed1)

In addition, students in Focus Group 1 shared other ways to reflect; an example is through observation. They stated that authentic assessment tasks in themselves force them to reflect and think, and through observing how their peers perform, they learn. Furthermore, in assessment tasks where they had to videotape their own performance (in this case, prescribing an exercise to a client), they reflected on their interactions with clients and improved on their communications in subsequent attempts. Because they knew they were being videotaped, they took time to think through the process beforehand, which is another type of reflection termed ‘reflection for action’ (Olteanu, 2017).

“...even if watching videos, we could know which things we did good and bad.” (Student FG1 S4)

6.2.4 Curriculum Mapping

The management participants specifically mentioned that in order to use assessment tasks for learning, such tasks need to thread through the curriculum to ultimately meet the programme objectives. This is obviously something that could be implemented at a more macro level, beyond learning within an assessment task or within a subject.

“...you need to take a whole-program view...can have a look at the proportion of assessment that comes from exams and assignments and from ...so you could actually do a map of that to see where it's happening....we're asking the question 'is this assessment really valid?'...” (Ed14M5)

This curriculum mapping element is essential for achieving the ultimate learning outcomes of university students. If assessment tasks are not valid and do not build upon one another, it is difficult to reach the subject, programme or even university level objectives. How curriculum mapping provides opportunities for Assessment for Learning is further elaborated in Section 6.4 below.

6.2.5 Summary of Theme 1

Theme 1 reveals that the participants in this study naturally view assessment as evaluation. However, when they further described their experience with assessment, they showed understanding and provided examples of Assessment for Learning. Specifically in their experience, learning happens when the assessment task is authentic and focuses on real-life application rather than simply memorisation of knowledge and skills.

Beyond the assessment task, feedback and reflection of feedback are essential to facilitate learning. The participants very much favoured timely, explicit and non-judgmental feedback. Also, feedback provided at various points of the learning process is best so that students can use feedback for continuous improvement. Reflection upon feedback or through the

observation of one's own performance or the performance of one's peers can also help with learning. At a more macro-level, assessment tasks should be reviewed as a whole so that each assessment builds on another to reach the ultimate programme and the university objectives.

In summary, despite the ingrained evaluative function of assessment in most participants' mindset, they themselves use assessment to facilitate learning, probably without a visible or explicit concept that these are features in the concept of Assessment for Learning.

6.3 Theme 2 – Burden and Mismatch in Current Assessment Practice

Theme 1 illustrates that the participants possess some knowledge and are using some features of Assessment for Learning in this higher education institute in Hong Kong. However, this knowledge is subtle because the participants still naturally perceive evaluation as the primary function of assessment. Continuing on from this, Theme 2 displays the negative views of assessment as the participants talk about frustrations with the current assessment system. In particular, the traditional examination-focused culture, over-assessment and the misalignment of assessment and learning objectives all contribute to their frustrations.

6.3.1 Burden – Examination-focused and Over-assessment

The participants were asked to share the types of assessment they have used or experienced, and to describe the relationship between assessment types and learning. Not surprisingly, examination (especially in the form of multiple-

choice examination) is the most frequently experienced assessment type for participants at all levels. Despite its “popularity”, the participants commented negatively about examinations:

“...examination is too focused on memorisation” (Student FG1 S2 & S5)

“they (examination) only facilitate knowing not learning” (Student FG2 S5)

Educators also regarded examinations as a “one-off assessment where it does not reflect core and individual abilities” (Ed4, Ed6M1, Ed8 & Ed9). Examinations are also a source of stress and anxiety for students (Ed6M1 & Ed9).

Interestingly, the participants provided some explanations for why examinations are a frequent practice:

“I think for the lecturers, they may feel easier to use a paper-based exam (MCQ and short questions), it’s a lot easier for them.” (Student FG2 S2)

“Percentage of assessment (on examination) is set by the department...” (Ed8)

The quote from the educator participant reveals that examinations are required by the departmental policy; whereas in the eyes of the student, examinations may be used out of convenience. In fact, the traditional culture

of examination-oriented assessment in some Confucian-influenced countries, especially when compared to the western culture, is partly because of the pressure to generate academic outcomes (Ratnam-Lim & Tan, 2015).

Examinations are perhaps the easiest and fastest way to inform academic outcomes (Pham, 2011). In addition, they may be the most widely accepted standardised way to show academic results, as compared to other assessment types, even though the participants in this study did not explicitly say so.

In addition to the examination-oriented practice in this context, the amount of assessment also appeared to be an issue:

“I feel that the assessment is far too much, and therefore, we (educators) grade too much... there are groups of 200 students, 250 essays, they write a lot in our course so the grading is very intense.”
(Ed2)

Over-assessment brings negative effects for both students and educators, such as stress and fatigue, and therefore creates a burden to the stakeholders in education:

“...I don't want to give them (students) too much pressure...there are too many homework, tests and exams...it is very stressful and they have immense pressure.” (Ed6M1)

“...They (students) are already very tired with their projects...I don't want to overload them.” (Ed8)

In addition to the emotional burdens created by over-assessment that may hinder learning, over-assessment may also take away time for reflection because students prioritize their time to study for examinations, as illustrated by the following quote:

“I think they (students) are being over-assessed in the subjects. When they come to our classes, they haven’t properly prepared...Every week there is a MC test by some teachers. Every week there’s a group that have presentation. Every two weeks you got to submit something. The students have done so much that they don’t have time to reflect. They don’t have time to slow down.” (Ed7)

As mentioned in Section 6.2.3, the participants identified reflection as an essential element for facilitating learning. Reflection needs time and cognitive effort in order for it to promote deep learning. In an examination-oriented and over-assessment culture, it is reasonable to picture that students focus their time and effort to achieve good grades. The general consensus by students and educators of this sub-theme is that examinations have little benefit in terms of facilitating learning. However, students spend a major proportion of their time preparing for examination-alike assessment tasks, leaving little time and energy to prepare and reflect on what in real life needs to be learnt and how to learn or improve. Together with the lack of feedback mechanisms inherent in examinations as stated by Ed1 and Ed2, students and educators tend to favour the *certification* function of assessment up-front and minimise the *learning* function. These traditions may prevent the more explicit adoption of Assessment for Learning.

6.3.2 Mismatch of Assessment and Objectives

Another issue of the current assessment system the participants identified is the mismatch or misalignment of assessment tasks/types and learning objectives. In every subject there is a set of learning objectives; learning and assessment activities should theoretically be mapped with these objectives in order to promote learning and achieve these objectives essential for the discipline. Unfortunately, in reality this ideal is not always achieved and participants of all stakeholder levels perceived that there is quite a significant mismatch between the two. This mismatch occurs when teachers teach students at one level (e.g. theoretical) and assess them at another (e.g. application); assessment types or tasks do not always reflect the core abilities and essential skills needed for the profession. In addition, there is often a lack of carryover from assessment to learning:

“...assessment drives learning. It has to be in-line of what you would like to achieve because if assessment is not in-line, then you’ll...it’s not a continuum...What you set out as learning tasks, then how you assess it...is my biggest critique on some of the courses...Throughout the course, there was no learning task related to making those connections...If you want them (students) to make the links... then you would have to make them experience and probably fail...” (Ed1)

“...I don’t know if we are cultivating and assessing a student’s ability to reason (which is important to my field and other fields). We just don’t assess it...They (students) are not ask to think about how they reason,

they just stand up and present or they will just give an exam answer.
It's disjointed I think." (Ed7)

"After they (students) passed the exams and went out for clinical placement, they still appeared to be quite 'blank'... even though they did great in exams, they still only perform fairly in placement. And those who did not do great in exams, they really cannot handle some of the skills in placement." (Ed3)

"... I'm actually not sure if the current assessments in our department align with learning objectives..." (Ed13M4)

The above quotes illustrate the mismatch in the assessment system the participants are currently using. The mismatch appears to stem from multiple levels, from teaching/learning tasks to assessment tasks to learning objectives. This indicates that the current assessment practice may not be an accurate or valid reflection of students' abilities. Even if learning may be occurring within an assessment, the direction or end goals of learning do not meet the expectations of educators and the current requirements of the discipline. Given this confusion, it is understandable why adoption of Assessment for Learning is difficult at various levels.

This perception of the misalignment of assessment with learning was not only expressed by the educators and the management group but students also had strong views on this:

“...around exam time, we are told that facts are not that important but application is the key. But how to apply the knowledge, there is no way for us to get the answer. Sometimes we use one to two sample questions to practice, but those are super easy which do not reflect the actual exam questions difficulty level.” (Student FG1 S2)

“I believe that most assessments in this university just stress the ‘knowing’ level where they just ask about your knowledge in the subject.” (Student FG2 S5)

“...the assessments we have is actually – ok if I do well, then I get a good grade, then I remember everything...somehow you only know how to work inside your classroom or work inside your lab, so it’s not even knowing how to work...after the exam, just forget about it then that’s over...” (Student FG 2 S1)

The above opinions from students resonate with those from the educator/management participants, indicating their perceptions and experience are somewhat similar in the current system. Students felt that they were taught at the theoretical or factual knowledge level while being expected to apply in examinations. The assessments do not seem to relate to real life and did not help them with the translation of knowledge.

While the integration and application of knowledge is the ultimate goal of any education, students thought they lacked this experience in the current assessment system. However, this is an interesting finding because participants at all stakeholder levels also named a number of authentic

assessment experiences as mentioned in Theme 1. This may imply that the participants' foci are still on examination when they think of assessment and other non-traditional assessments (and their functions/benefits) do not automatically come to mind.

6.3.3 Summary of Theme 2

Theme 2 illustrates the burden and mismatch in the assessment system in this university in Hong Kong. Examinations are the most commonly used assessment type, despite frustrations from all participants. Various factors contribute to this focus on examinations examination-focused, such as departmental policy and convenience of practice. As a result, this may limit the adoption of more innovative assessment with an authentic focus. In addition, over-assessment is another burden the participants experienced. This may take away students' time and effort for reflection, which may hinder deep learning from happening.

With regard to the mismatch of learning objectives and assessment tasks, participants all experienced that students are taught at one level (e.g. factual) while being assessed at another (e.g. application). The transformation of knowledge into real life application seems to be lacking. Although there seem to be many mismatches expressed by the participants between teaching/learning and assessment, they probably refer to examinations but not to other authentic or innovative assessments as mentioned in Theme 1. Authentic assessment does not automatically feature in the current system. In

the following section, the focus turns to reporting what assessment should be based on if the goal is to promote learning.

6.4 Theme 3 – Ideal Actualisations of Assessment for Learning

Although the traditional examination-oriented assessment practice is heavily ingrained in Hong Kong and creates a burden for adopting Assessment for Learning, participants shared their thoughts on the ideal actualisation of the concept. As criticised as one of the burdens in the current assessment system, the mismatch of learning objectives and assessment tasks is huge; therefore, the two should be realigned in order to promote learning.

In addition, examination is not necessarily very bad when it is continuous and holistic. Specifically, examinations or assessment with a feedback mechanism are recommended. Open-book examinations are regarded as more suitable compared to closed-book examinations because they assess application rather than memorisation. Ideally, assessment should facilitate students' potential for self-regulated and lifelong learning, the concept of Sustainable Assessment (Boud, 2000). The sub-themes are discussed in the following sections.

6.4.1 Re-alignment of Learning Objectives and Assessment Tasks

Theme 2 reported on the mismatch between learning objectives and assessment activities in the current system which limits learning through assessment. In order to adopt Assessment for Learning, it seems imperative

to fix this mismatch. That is, the design of assessment tasks needs to match with learning objectives. This opinion is shared by one of the participants:

“The assessment has to be valid. It also has to be reliable. It’s not just about assessing, it’s about the right kind of assessment to give students opportunities to demonstrate achievement of learning goals.”

(Ed14M5)

Using assessment that is valid and reliable is crucial to minimise the mismatch. For most subjects and programmes in higher education, the ultimate objective is to translate knowledge into real-life practice, i.e. knowledge application. It conveys the message that assessment ought to assess application rather than just memorisation. Authentic assessment is a type of assessment that focuses on assessing application, meaning that it can be a method to minimise the mismatch mentioned above. This point is further elaborated in the next sub-theme in Section 6.4.2.

In addition to using valid and reliable assessment that focuses on assessing knowledge application, there is a need to clearly define the learning objectives and how assessment is matched with the objectives, as mentioned by Ed12M3:

“...if students know what the assessment contents are and how they are link with the course design, it will help them grasp which kind of learning outcomes they should achieve...the clearer the relationship is defined, the higher the chance the teaching/learning activities would fulfil the course objectives” (Ed12M3)

This relationship not only involves linking up learning objectives and assessment tasks but how the two should map with teaching/learning activities. If the ultimate objective is the application of knowledge in the real world, and assessment tasks are designed to assess application, then teaching/learning activities should also centre on application mimicking the real- life situation as much as possible. This should be done at subject levels and map with programme outcomes. This point resonates with using assessment for curriculum mapping and is one of the elements where assessment could promote learning.

6.4.2 Involvement of Continuous and Authentic Assessment

The above sub-section mentioned using valid and reliable assessment to re-align learning objectives and assessment outcomes. The participants expressed the view that continuous and authentic assessment serve this validity purpose and can assess students' learning outcomes more accurately. Continuous assessment means assessment is implemented not only on one occasion but on numerous occasions, in order to accurately gauge students' learning progress. Ed6M1 expressed the view that the feedback process in continuous assessment is instrumental to promoting learning:

“...I would give them evaluation and feedback, to let them know whether they are on the right track, if they are aligning with what I am thinking of or the message I am delivering. Once they have this experience, they move on to do their own research...they will take this

opportunity to make improvement as they have this ‘redo’ process.”
(Ed6M1)

This type of continuous assessment is most often seen in authentic assessment but rarely seen in examinations. It was already reported in Theme 2 that the current one-time examination tradition lacks feedback mechanisms and this factor alone hinders learning from examinations. However, adding continuous assessment with a feedback mechanism to the traditional examination culture is possible (Tanner, 2017) and is definitely an opportunity ahead, especially in this examination-oriented culture where completely abolishing examinations may not be feasible.

Along the same line, the participants commented that assessment should not be over-reliant on examinations but should include different layers for various purposes in order to better match with the learning objectives. Examples include assessing different knowledge layers such as factual and application (Ed1, Ed11M2) and embedding a diversity of assessment types based on the nature of knowledge (Ed12M3). Given this, examinations may have their role in learning, as long as questions include different layers and are fit for purpose. In fact, the participants mentioned the merit of open-book examinations:

“...even though it is also a final exam, it’s very different because it’s open book, open internet, it’s up to you...you actually have to really learn on the subject where you have to understand the concept...then

you can apply the concept to the questions...you have to analyse what is the concept and evaluate so that you can apply..." (Student FG2 S5)

"They (students) don't have...it's not a closed-book, they don't need to memorise the rules or the provisions because there's so much of that, that actually, it is open book, this final exam, it works okay..." (Ed2)

Open-book examinations are not as common as closed-book examinations perhaps because of marking difficulties. Questions are usually analytical and involve application. Students need to search and understand the information before finding an answer or a solution. The cognitive requirements of open book examinations for students are much greater, and mimic real-world requirements. Open-book examinations are therefore regarded as authentic in this situation.

6.4.3 Promotion of Sustainable Assessment

Continuous and authentic assessment provide the opportunity to restructure the current examination-focused assessment practice. Ed1 mentioned assessments should be a trajectory and act as the building blocks in the curriculum. Assessment should ideally be used by students not only for the duration of their study but as a continuous way to monitor one's learning needs and progress, and to seek ways to learn when a deficiency is identified. This self-assessment ability should be beyond the immediate assessment environment and extend to the future, which is the concept of Sustainable Assessment (Boud, 2000).

“...assessment to me is just like, get your hands in there, dig in and get an evaluation out of it. The way that I want to interpret assessment is not a one snapshot thing, which is our biggest failure at the moment. Our assessment is a one snapshot thing in terms of a test, examination and assignment. I would like to support, if possible, an ongoing assessment where there are multiple dips and beyond a semester.” (Ed4)

This practice is rarely implemented in current assessment practice. It takes students, educators and management to understand and appreciate the value and potential of assessment and to break away from the traditional burdens. Reinventing assessment practice to promote lifelong self-assessment and learning is what the participants envision as the ultimate goal of higher education.

6.4.4 Summary of Theme 3

This theme shed lights on the ideal actualisation of Assessment of Learning by tackling existing problems and suggesting assessment types that can serve this role. First, the mismatch between learning objectives and assessment tasks can be minimised by clarifying the relationship between the two and by using valid and reliable assessment types. Continuous and authentic assessment is more accurate in gauging the achievement of learning objectives. Examinations still have their merits if a feedback mechanism is embedded and if there is a balance between examining memorisation versus application. Furthermore, assessment should not be

seen as a single end point of a subject but should thread through the curriculum to meet the programme objectives. With this type of assessment system, learners will use assessment to self-assess and regulate their learning needs during and beyond higher education, to promote lifelong learning.

6.5 Summary of the Chapter

This chapter reports on the findings related to the participants' conceptions and experience of assessment practice and in particular Assessment for Learning. The participants instantly viewed assessment as evaluation even though they mentioned a number of learning features in their assessment experience. They are burdened by the traditional examination-oriented focus and over-assessment practice, as well as frustrated by the mismatch in learning objectives and assessment tasks. As such, the ideal actualisation of Assessment for Learning lies in relieving these burdens, by using authentic forms of assessment versus those focusing on memorisation. The participants also touched on using assessment to teach life-long self-assessment.

The findings in this chapter provide background knowledge on the participants' understanding of Assessment for Learning and serve to explain their perception and utilisation of rubrics, a tool linked to the concept of Assessment for Learning. The next chapter reveals the results related to the concept of rubrics, the experience with rubric actualisation and the interpretation of the rubric policy. All findings are synthesised and discussed in Chapter 8 with explanations and implications provided.

Chapter 7: Results of Actualisation of Rubrics in this Policy Context

This chapter is the second Results chapter of this study. The findings presented in this chapter address Research Questions 3 and 4 targeting rubric actualisation under the rubric policy in one university in Hong Kong. The research questions are:

RQ3. Can rubrics actualise Assessment for Learning in the eyes of students, educators and management? Why or why not?

RQ4. What are the policy implications according to the experience of educators and management in this policy initiative?

Findings display four themes related to the topics of rubrics and the rubric policy. Specifically, Theme 4 “Potential of Rubrics for Learning”, Theme 5 “Barriers of Rubric Actualisations” and Theme 6 “Ideal Actualisations of Rubrics” address Research Question 3. Theme 7 “Promoting Holistic Assessment Practice through Policy” addresses Research Question 4. Below is an illustration of themes and sub-themes for these two research questions:

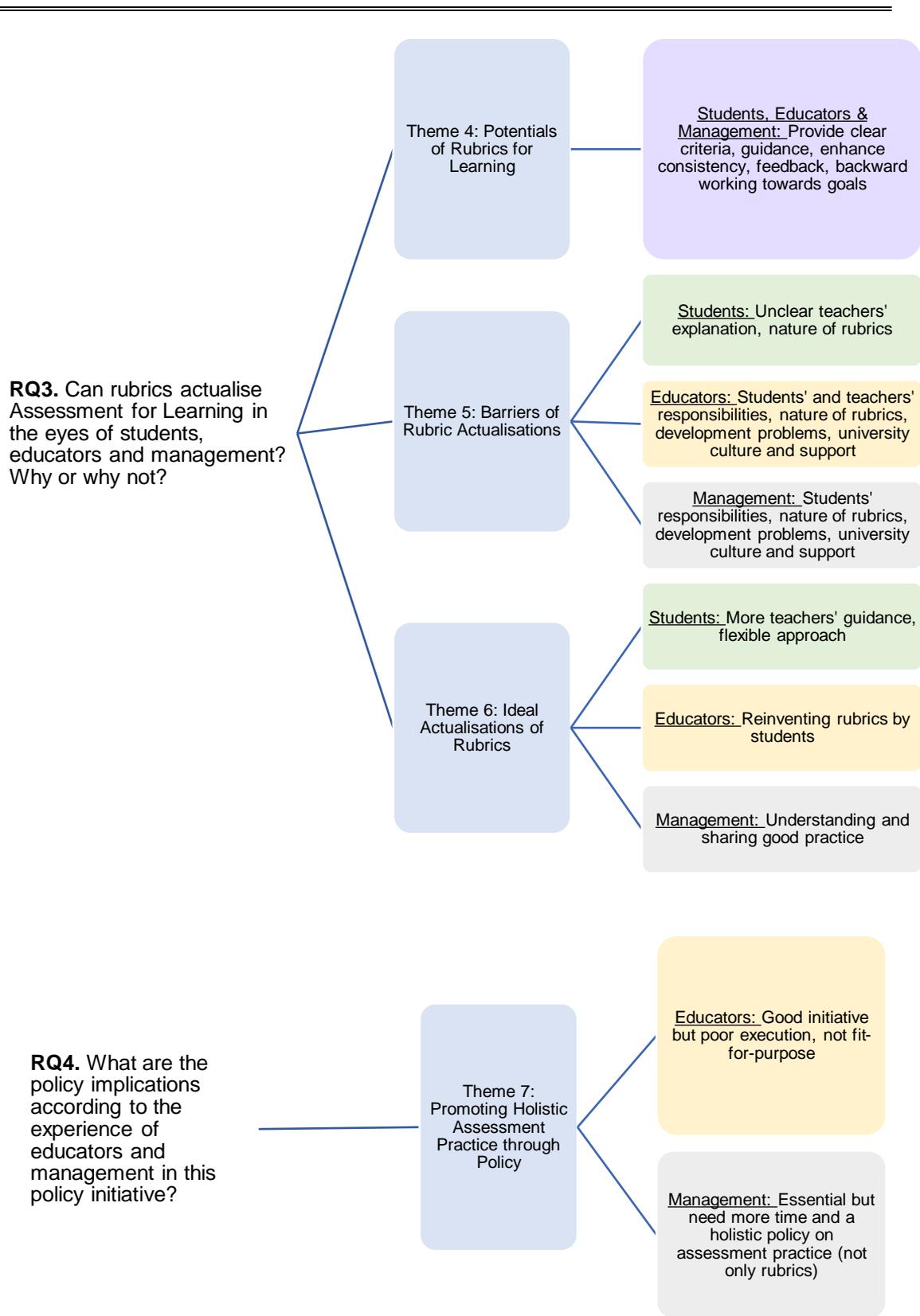


Figure 7.1 Illustration of Themes and Sub-themes for Research Questions 3 and 4

7.1 Theme 4 - Potentials of Rubrics for Learning

The participants were asked to share their views and experiences on using rubrics in assessment. Most of them regard rubrics positively, although more in the realm of rubrics as an assessment tool. That is, they regard rubrics as a grading tool, to evaluate students' performance. This finding is similar to their first impression of assessment where it is used for evaluation.

The benefits of having rubrics as a grading tool include setting expectations, making grading criteria explicit and enhancing grading consistency.

“They (students) can anticipate the grading and for this type of students, apparently, this is very important because they are very assessment oriented and they care a lot about the score...you (teachers) set up the rules upfront and everybody sees that and they don't get surprised after.” (Ed2)

“...No matter what, of course it (grading) is going to be subjective. But where there is a rubric, at least there is a standard of what is the different grade. And therefore like even though when lecturer may not like this student, because this student fulfil the standard, he may still be able to get a higher grade.” (Student FG2 S5)

“...It (rubric) is good for co-teaching, and more consistent when giving grades” (Ed5)

As with the previous findings in Theme 1, the participants revealed some pedagogical functions of rubrics when they talked more deeply about their experience. For example, rubrics can be used as a “backward working” tool for students who make use of rubrics. In other words, students can use the end goal spelled out in rubrics and work backwards to achieve the goals. In addition, rubrics are regarded as a feedback tool to guide students for future improvement. With this, communication is enhanced between students and teachers.

“...From week 13, we’d work backwards...’what does the end look like?’ They (students) ask that to you right at the beginning, so that you have to know where you want to go. Otherwise, you will never get there.” (Ed7)

“I think sometimes it can be good to have a clear rubric...you actually are thinking about how to do it in the best way.” (Student FG2 S2)

“...They (students) realise the things I’m looking for, it helps with communication with students...I will give feedback, apart from verbal feedback I will have a written feedback and attach with the rubric. They (students) will see the feedback and grade together.” (Ed6M1)

The above quotes illustrate that both students and educators experienced some learning benefits of rubrics. What did the participants do to make the learning function possible? Frequently, educators post rubrics on the learning management system for students to read (Ed1, Ed4, Ed5, Ed11M2). Some educators took time to explain rubrics in class (Ed9,

Ed12M3, Ed14M5); while some supplemented rubrics with exemplars in addition to explanations (Ed2, Ed6M1). Students in both focus groups reported that they read, reasoned out and asked for clarification and confirmation from teachers. In essence, they followed what were laid out in the rubrics in order to reach the expectations of a particular assessment task. This is essentially self-regulated learning through assessment. Is the current rubric actualisation ideal and effective? The following theme discusses the barriers the participants faced while using rubrics.

7.2 Theme 5 - Barriers of Rubric Actualisation

Under this theme, the participants identified problems they experienced related to actualising rubrics in general. These problems stem from several areas: other stakeholders, the nature of rubrics, the development of rubrics and the organisational culture of the university. The student participants experienced more issues with rubric nature and thought teachers should play a heavier role to improve rubric execution. On the other hand, the educator and management participants shared issues in the development of rubrics and involving other stakeholders, as well as frustrations with the university's policy and support. The problems are explained in the following sub-sections.

7.2.1 Barriers from Other Stakeholders

One of the barriers of effective rubric actualisation stems from other individuals, both within the same stakeholder group and across other stakeholder groups. Across stakeholder groups, the educator participants said they explicitly explained to students the criteria and grades of the rubrics,

while the student participants thought they needed more debriefing and explanations. Educators thought it is the responsibility of students to read and use the rubrics, while students depended upon teachers to help them interpret them. Clearly, there is an expectation gap between the two groups.

“You got to admit that students will be students. They’re going to say ‘I still don’t understand what you’re assessing.’...I’ve explained it. I’ve got the PDF. I got arrows all over the place. I actually walked it through with you, what these numbers mean, what this means. I give you example and you still ask me questions...even if I answer again, at the end of the semester and student feedback questionnaire, they’re still going to say ‘we don’t get it.’” (Ed7)

“...The briefing before the assignment starts is very important...although we could find information from the rubrics, we don’t know exactly how to do it.” (Student FG1 S2)

The above comments illustrate tensions between the educator and student participants when using rubrics. Previously mentioned in Theme 4, the educator participants frequently posted rubrics on the learning management system and some of them took time to explain them in class. Some educators supplemented this with exemplars. These are common practices and have proven to be effective (Panadero & Jonsson, 2013). Yet, the student participants seem to find them insufficient and think teachers should do more to facilitate the use of rubrics. This tension may suggest a lack of readiness to use rubrics and this is further discussed in Chapter 8.

When it comes to using rubrics as a feedback tool, the educator participants were concerned about their effectiveness because they felt that students do not care:

“The students now always ask for the grading after. They like to know how we will grade it but after we have graded it they just want to know the grade. They will not come back and look at the rubric for example. A few of them will do, I will say 15-20%, no more than that. It meets the purpose of guiding students of what we care about when they prepare their work but the feedback function doesn’t play a good contributive role in my experience.” (Ed2)

“I’ve got the feeling the students don’t care too much. I don’t know why. I suspected they didn’t even read it. I suspect they would only take action if they found that the overall marks that they get eventually is seriously deviated from what they expect. We don’t underestimate our student. They are experts in math calculation.” (Ed4)

These two quotes further display tension in the form of distrust. Some teachers are unsure if students use rubrics seriously and others do not believe that they do. This potentially limits the genuine adoption of rubrics. Furthermore, the inconsistent execution of rubrics among teachers (i.e. within the same stakeholder group) is another layer of the barrier. The educator participants commented that the execution of rubrics is premature because of different adoption patterns and schools of thought both within and across academic departments.

“My criticism. Do we actually follow the rubric when we are marking?...The cracks start to occur when there are subjective interpretation and when some teachers tend to be too lenient or too harsh.” (Ed4)

This quote shows that teachers may not have a uniform execution of rubrics because they interpret the criteria/grades differently. This is somewhat related to the inherent subjectivity of rubrics but is also in contrast to what rubrics are supposed to tackle, i.e. to increase grade consistency. Teachers may value a more linear execution of rubrics in order to avoid confusion for both teachers and students.

7.2.2 Barriers from the Nature of Rubrics

Although there are many benefits of rubrics revealed in Theme 4, the disadvantages of rubrics are equally identified and these drawbacks add to utilisation barriers. In fact, there is a fairly high degree of resonance between the student and educator participants when it comes to the disadvantages of rubrics. Both groups think that rubrics, to a certain extent, limit creativity because the criteria descriptors are pre-set and students are bound to follow these. Although both groups expressed the view that having a rubric is better than not having one, rubrics are still relatively subjective and therefore create frustrations when grades are involved.

“I think there should be a balance between flexibility and also standardisation of grades where I believe different students have different ways of showcasing their understanding of certain subjects

and therefore there should also be flexibility...they may have different formats and different styles..." (Student FG2 S5)

"I do not agree with rubrics because they are too inflexible...they limit students' creativity and growth...they only create students meeting requirements on rubrics." (Ed8)

This drawback of rubrics is not new and rubrics were criticised as limiting students' creativity because students do not want to jeopardise their grades (Bennett, 2016). This may be a more serious issue in an examination-oriented culture like Hong Kong. When students perceive grades as the priority in an assessment, it is human nature to play safe and not risk any chance of submitting something teachers "don't like":

"...The grade difference is huge when not following the guidelines (in rubrics), nobody would dare not to follow." (Student FG1 S1)

This is perhaps one of the inherent limitations rubrics have and has yet to be resolved. Since both the educator and student participants are aware of this limitation, it is perhaps worth including in the debriefing or explanation session to raise awareness of both parties when utilising rubrics.

7.2.3 Barriers to the Development of Rubrics

This sub-theme pertains to the educator and management participants where they found multiple challenges and frustrations when developing rubrics. From a development aspect, the educator and management participants are still relatively confused in terms of how to develop a quality rubric:

“I think rubrics are difficult to set, especially for my subjects, those are hard science...written exam also, the rubrics of written exams are difficult to set.” (Ed10)

“...rubric is a bit difficult to develop, you need to show in details, to what extend you need to achieve in order to get such a grade. On the other hand, it can't be too detailed...so you need to strike a balance between general and specific criteria...I think it is quite challenging, it's not easy to write a good rubric.” (Ed11M2)

It is evident from their comments that rubric development is challenging and they have yet to grasp how to develop rubrics that fit their purposes. This frustration creates another layer of dissatisfaction even though they are aware of the benefits of rubrics as Theme 4 reports. The internal conflict of knowing that something is good to do yet not knowing how to do it well is definitely frustrating. It is worth noting that the university does provide training on rubric development for teachers and most educator participants in this study who have taken the training were satisfied. However, the training focuses on how to develop a rubric but not on what a quality rubric means.

In addition to *not* knowing how to develop a good rubric, developing rubrics is time-consuming and tedious. This may limit educators' commitment to development. For those who need to create rubrics from the beginning, the time commitment in advance is tremendous:

“I don't have time to set a good rubric. My workload is very heavy.”
(Ed8)

“...We need to think about the rubrics, like I have three assessment tasks...if each task need to have a rubric, individual teacher have a lot of work to do...teaching is already quite heavy already...” (Ed12M3)

“Tedious. I think what I wouldn't like when you're developing it (rubric) is you need to go back to the basics...when I said tedious because it goes back...I felt a little bit alone working with the subject team that is probably not quite so much interested in developing rubrics.” (Ed1)

It is a normal behaviour that educators prioritise tasks when the workload is hectic. This is especially true for educators who do not have a good understanding of or an interest in rubric development to begin with as expressed by Ed1. There are other factors that affect workload prioritisation which are discussed in the next sub-theme.

7.2.4 Barriers from the University's Culture and Support

This sub-theme displays a more macro-issue concerning organisational culture and support for rubric actualisation. First, the university's stance on teaching and learning influences how teachers actualise rubrics. Many educators and management participants in this study perceive that the university favours research over teaching. As a result, any initiative related to teaching and learning enhancement receives comparatively little attention. In other words, there are no incentives to adopt an initiative related to teaching and learning. Also, there are no consequences if teachers do not adopt this type of initiative. Together, rubric development and utilisation is not prioritised

when teachers are busy (as discussed in section 7.2.3) as illustrated in the following quotes:

“At the moment I think this is the problem with not getting rubrics or whatever pedagogical tool or approach or theories you want to use in teaching. The university tends to have much more focus on research. Any value system is on research not on teaching.” (Ed1)

“...(developing) rubrics can be hard. You have to think about your assessment tasks and the busy academic or the academic whose focus is not teaching, but more on research finds that an unwelcoming position.” (Ed14M4)

“...Teaching is not a priority in the university, teachers have no incentives to use rubrics...There are also no consequences if they (teachers) don't do it, so they don't...” (Ed13M4)

In addition to a culture that values research over teaching, the educator and management participants perceive that there is insufficient support on rubric development and execution from the university.

“The problem is if you (the university) just give the input without sufficient support, without enough understanding of, sorry for saying that, theoretical underpinnings of learning and teaching, I think it (the policy) doesn't have the impact that it could have.” (Ed1)

The only visible support from the university at the time of this research was staff training on rubric development. However, knowing how to develop a

rubric is not equivalent to understanding the theories and functions of rubrics. The above opinion demonstrates that some fundamental yet important information is missing from the university.

7.2.5 Summary of Theme 5

This theme illuminates the barriers the participants of this study face when actualising rubrics in assessment practice. These barriers span multiple layers, from those of the users to the level of the university. At the user's level, differences exist between the expectations of one's own and others' roles and responsibilities. There is also inconsistent practice within the same stakeholder group. Tension and distrust result. In addition, the inherent limitations of rubrics (e.g. subjectivity and limitation of creativity) cause frustration in using rubrics. Developing rubrics also presents another hurdle due to the teachers' lack of time and knowledge of establishing a good rubric.

At the university level, the training provided only focuses on rubric development but lacks more fundamental knowledge of rubrics. This adds to barriers to rubric utilisation because of this insufficient knowledge base. For organisational culture, the university gives the impression that research is valued over teaching. Unlike research output, there are no incentives or consequences for using or not using rubrics. This is manifested in the prioritisation of research over teaching activities when the workload is heavy and contributes to barriers in adopting rubrics.

With these barriers in mind, the next theme displays optimal actualisation of rubrics in assessment practice. Similarly to the barriers, the

ideal utilisation lies at both the user's and the university level and is reported below.

7.3 Theme 6 - Optimisation of Rubrics in Assessment Practice

The participants at different stakeholder levels shared views on optimising rubric utilisation. The student participants would like to see more teachers' guidance and a consistent yet more flexible approach to help them use rubrics for learning. The educator participants envision students reinventing rubrics for learning. Those participants with a management hat view themselves as having a role to understand and to share good rubric practice within the department. These sub-themes are described in the following sections.

7.3.1 Increasing Guidance and Flexibility by Teachers

As mentioned previously in section 7.2.1, the student participants regard unclear explanation from teachers as the major barrier to using rubrics. Debriefing used by some teachers can be one way to optimise rubric utilisation. This is, however, not enough in the eyes of the student participants. A consistent approach of how rubrics are introduced and treated by teachers is as important:

“...the briefing before the assignment starts is very important. I remember for the briefing for this course is done at the start of the semester for just 5 minutes then the lecture starts. So we actually don't really know what is going on...Although we could find some information from the rubrics, we cannot tell exactly how to do it.” (Student FG1 S2)

“...Some professors treat the rubric very seriously, sometimes they may even explain that in the first class and no class at all on that day because they really want the students to understand how we would be graded but some other professors just uploaded on the blackboard.”
(Student FG2 S1)

A very brief introduction of rubrics, such as simply uploading rubrics on the learning management system, does not allow students to understand how to effectively use rubrics. In addition, time spent may be related to the perceived importance of a task. Students may pay more attention to rubrics during the assessment process when they see that their teachers value rubrics. Teachers should not undermine the time and depth committed to introducing an assessment task and rubrics as it may help students with their learning.

Furthermore, the use of rubrics can be less rigid in the eyes of the student participants:

“...rubric is good, but...it is not a must to follow the rubric, and tell the student that if they want to use another way to pursue the work, please just go ahead and give it a try...if we really cannot think of any alternative, we can just follow the rubric.” (Student FG1 S1)

This comment manifests one of the inherent limitations of rubrics – limiting creativity in an assessment task. This is especially true for innovative students who can always think outside of the box. These students may not always want to follow a standard guideline and may want to have more room

to create their own standards in their learning and assessment process. In fact, the educator participants expressed a similar concept that the ultimate goal is to see students reinventing rubrics (refer to next sub-theme). Adopting a flexible approach seems to have merits in optimising rubric utilisations in the eyes of students.

7.3.2 Reinventing Rubrics by Students

In contrast to the above sub-theme, many educator participants believe that effective utilisation of rubrics depends upon individual students as reported in Theme 5. They expect to see independent and sophisticated use of rubrics by students. If students only care about grades in rubrics, but do not critically utilise rubrics to reach goals and beyond, the learning function of rubrics is minimal. In order to effectively use rubrics for learning, the educator participants envision students “reinventing” rubrics:

“Once you’ve got the rubrics, we want it to sustain, but a really really good rubric will reinvent itself. It will be self-generative to the point of...one day we will decay, but before that point, we are constantly rejuvenating ourselves...reframe and rethink.” (Ed7)

“For the ultimately bright and daring students, I am confident and optimistic that the rubric would not confine their ways and creativity in learning. If a student is really bright...whatever one says is not necessarily the truth...if a student can come up with some additional learning outcome beyond our rubrics, then I’m grateful and moved.” (Ed9)

These quotes share a similar idea to that which the student participants expressed, which is loosening standard requirements and increasing the flexibility of rubrics. The criteria in rubrics are often pre-set by teachers according to learning objectives. However, innovative students may see additional aims and outcomes not written in rubrics. If students are involved in setting criteria in rubrics, they are more engaged in using them (Jones et al., 2017). This is one way to adopt a more flexible approach when developing and actualising rubrics, as suggested in the previous sub-theme in section 7.3.1. Although the two stakeholder groups appear to be blaming each other, they are implying a similar new approach to optimising rubrics.

7.3.3 Understanding and Sharing of Good Practice by Management

The management participants shared another perspective to optimise the utilisation of rubrics, which is by the understanding and sharing of good practice. They view themselves as having the responsibility to assure teachers about understanding rubrics, be it by complying with the university policy, or by adopting a pedagogy they believe in. As previously mentioned in Theme 5, knowledge and time commitment are some of the adoption barriers the educator participants experienced, especially when the workload is heavy and any initiatives related to teaching do not seem to be a priority.

Management's involvement in various aspects may resolve some of these issues.

“From management's perspective, I think we need to try our best to explain clearly to students, to teachers, to management, it (rubric) is

necessary, it is not an extra workload, it is for fairness, justice and consistency...we need to explain why we do so.” (Ed11M2)

“...In management position, if I tell colleagues I received a message from the university that we need rubrics and just pass the ball...I need to know what is rubric, why do we need to implement this and think from colleagues’ perspectives of what query they may have and how I would respond.” (Ed12M3)

The initial action lies in communicating the importance of new initiatives to frontline teachers, rather than just demanding that colleagues follow. It is normal to experience resistance to change in any initiative and if those in management can step into the shoes of teachers, the resistance may be less. Mutual understanding can alleviate tension and this is especially true when teachers have different backgrounds and philosophies:

“...one size doesn’t fit all...we have some recommendation...we could only share good practices...I will find opportunities to talk...” (Ed12M3)

Sharing what they do is a powerful way of influencing change as they lead by example. Considering individual differences can enhance mutual respect when adopting a practice which comes from the top. This, however, may not be enough as priority issues still exist.

7.3.4 Summary of Theme 6

Responding to the barriers of rubric utilisation (Theme 5), this section paints an ideal picture for how rubrics can be used to facilitate learning. The points

are something that the participants envisage because from the experience shared in Theme 5, the current actualisation of rubrics appears to be less than optimal.

The picture and ideas they share include providing more guidance and loosening the standards and rigid criteria of rubrics by teachers. This allows students to think 'outside of the box' when they can. Teachers would like to see students' independent use of rubrics and the reinvention of existing ones. To handle time and commitment issues, the management group commented on the importance of communicating the rationale behind rubrics and the essence of leading by example, rather than demanding teachers to follow. This theme shed lights on what each stakeholder group needs to collectively do in order to effectively optimise rubrics. The next and last theme reports on the roles and implications of policy in this rubric actualisation exercise.

7.4 Theme 7 - Promoting Holistic Assessment Practice through Policy

Tapping into the rubric policy in the institution where this research was conducted, this theme illustrates how the educator and management participants interpret such policy and implication for practice. Interestingly, most participants regard the policy as "helpful" because it increases awareness of rubrics within the university. Also, they expressed the opinion that a policy is needed if the university wants to adopt good practice, otherwise, the assessment practice in general becomes scattered and inconsistent.

“...It (policy) is like a reminder, to remind teachers to think about why it is required to do these assessments.” (Ed5)

“In my experience, if you leave it (rubric) as a voluntary choice, I’m not sure that you will get a uniform approach. No, there are very different approaches here, on how to evaluate.” (Ed2)

“It’s necessary to have a policy. There are many disciplines, staff...there needs to be a coherence...This is good practice...you could not be just like a priest and tell the follower it’s good for you. The effectiveness won’t be there...spreading the central message top down, I think this is the most effective way.” (Ed12M3)

The quotes suggest that having a policy is initially a good and effective way to actualise an initiative uniformly, be it by reminding teachers of their practice or by “forcing” them to adopt a new mindset and approach. Having a good policy initiative is a start; effective execution is equally or more important in this scenario. The educator and management participants mention that despite having good intentions, the execution of this policy is premature and hinders optimal rubric actualisation.

“I think the spirit is good but the execution and the deliverable that came up so far, well, the ones that I’ve seen, they’re terrible...it’s in the execution when they actually use it, then I find it to be terrible...It’s not a holistic approach.” (Ed4)

“If we can set policy that will drive...thinking, the feeling that is the emotion and the acting that is the behaviour of not only our students

but our staff and senior management, that would be awesome...I think we need it (policy) if we use it right.” (Ed7)

These views display an issue of a non-holistic and perhaps not fit-for-purpose policy. It implies that despite this policy, it is unlikely to be achieving its objectives and is therefore unlikely to improve the existing assessment practice. The concern with this non-holistic approach probably points to the fact that rubric is only one element in the concept of Assessment for Learning or within an assessment system. Even with sophisticated design and implementation of rubrics, the entire assessment practice needs to be aligned from assessment task design to tools in order to facilitate learning. A policy on one element of a major practice only tackles one area, in this case the “what”, but is less likely to change the mindset of emphasising the role of learning in the assessment system, the “why” as this participant states:

“I think the rubric policy, it’s one aspect of it. It’s not the answer. It’s not the cure because it doesn’t address those other things that I’ve been talking about – assessment type, assessment mix, the programme, across the programme, and who’s looking at the rubrics holistically...Someone will tick a box to say they have a rubric. Nobody will necessarily look at the quality of that rubric, the application, the validity, whatever...A lot of it will boil down to a box being ticked.”
(Ed14M5)

Altogether, the participants interpret that although this top-down rubric policy is good for raising attention, it is far from helping to improve the current practice because its execution is poor. It implies that rubric practice is still

premature and there is definitely a place to refine the existing policy in order to reach the ultimate goal of promoting Assessment for Learning. If the policy is better-rounded, the barriers of actualisation are likely be minimised.

7.5 Summary of the Chapter

This chapter presents the findings related to the participants' views and experiences of rubric actualisation under the top-down rubric policy in this university. In the same way that they perceive assessment, the participants relate rubrics to evaluation and grading. However, they identify elements in rubrics that facilitate or could promote learning. Nonetheless, the barriers with optimal rubric actualisation are significant and span various domains, including barriers or tensions from other individuals, the nature of rubrics, time and knowledge limitations in development, the value the university places on research as well as insufficient support by the university. The optimal picture of rubric actualisation relates to the learning function, i.e. learning how to use rubrics from teachers, increasing the flexibility of rubrics and increasing knowledge to reinvent rubrics. The management group plays a role in facilitating this process.

The educator and management participants concur that this policy serves to raise the awareness of rubrics in general within the university. However, changes in practice are too premature to comment on because the policy is failed by multiple execution issues. These include the policy itself not being holistic and not fit-for-purpose. All in all, the findings provide insights on

policy refinement and future assessment policy initiatives. The next chapter explains the findings in detail with comparisons to current literature in the field.

Chapter 8: A Discussion on the Challenges and Potential of Assessment for Learning and Rubrics under the Policy Situation

This chapter serves to analyse the findings in depth and to provide explanations and implications for the four research questions of this study. It develops and sustains the argument by comparing and contrasting existing relevant bodies of literature, as well as by critically examining the reasons behind the findings and their implications. As such, this chapter is structured by answering the four research questions. These answers are connected at the end to address the overall aim of this study, which is to explore the perceptions and actualisations of Assessment for Learning and rubrics of students, educators and middle management under a mandatory top-down policy in one university in Hong Kong. Another unique feature of the research context is the heavily-ingrained examination-oriented culture. The policy environment and examination-focused culture both serve as the background when interpreting the findings of this study. For easier understanding of explanations provided for each theme, Table 8.1 is formulated to map bodies of literature used for explanations of findings.

Research Questions	Themes	Main points or sub-themes	Explanations	Bodies of Literature
1. What are the perceptions and understanding of Assessment for Learning in students, educators and management?	Theme 1 – Elements in assessment facilitating learning	<ul style="list-style-type: none"> Naturally regard assessment as measurement Further discussion reveals Assessment for Learning features (authentic assessment, 	<ul style="list-style-type: none"> Dominant discourse of assessment Confucian-heritage culture Lack of assessment literacy Personal experience/beliefs 	<ul style="list-style-type: none"> Section 2.2.1: Heavy focus on Assessment for Certification Section 2.1.2: Assessment for Learning Section 2.2.2: Confucian-heritage culture

Research Questions	Themes	Main points or sub-themes	Explanations	Bodies of Literature
		feedback, reflection on feedback, curriculum mapping)	<ul style="list-style-type: none"> Background of applied disciplines in this university Previous exposure to the Western education culture 	<ul style="list-style-type: none"> Section 2.2.3: Assessment literacy
2. How do students, educators and management identify barriers and opportunities for the adoption of Assessment for Learning?	Theme 2 – Burden and mismatch in current assessment practice	<ul style="list-style-type: none"> Examination-oriented culture Over-assessment Mismatch between learning objectives and assessment tasks 	<ul style="list-style-type: none"> Cultural influence Lack of understanding between learning objectives and assessment tasks 	<ul style="list-style-type: none"> Section 2.2.2: Confucian-heritage culture Section 2.2.3: Assessment literacy Section 2.2.4: Resources and culture
	Theme 3 – Ideal actualisations of Assessment for Learning	<ul style="list-style-type: none"> Realign mismatch Use more continuous and authentic assessment to promote application Embrace sustainable assessment 	<ul style="list-style-type: none"> Background of applied disciplines in this university – focus on application and continuity, i.e. real- life work requirement, e.g. healthcare Lifelong learning is important in these disciplines – may be required for registration 	<ul style="list-style-type: none"> Section 2.1.2: Assessment for Learning Section 2.1.3: Sustainable Assessment
3. Can rubrics actualise Assessment for Learning in the eyes of students, educators and management? Why or why not?	Theme 4 – Potentials of rubrics for learning	<ul style="list-style-type: none"> Naturally regard rubrics as a measurement tool Further discussion revealed learning functions (self-regulation, feedback) Traditional actualisation: post and read, explain, exemplar 	<ul style="list-style-type: none"> Same as Theme 1 	<ul style="list-style-type: none"> Same as Theme 1 Section 3.2: Merits of rubrics

Research Questions	Themes	Main points or sub-themes	Explanations	Bodies of Literature
	Theme 5 – Barriers to rubric actualisations	<ul style="list-style-type: none"> • Other stakeholders or individuals • Nature of rubrics • Development difficulties • University's culture and insufficient support 	<ul style="list-style-type: none"> • Findings on the nature of rubrics, development difficulties and organisational culture/support are comparable to what has been found in the literature • Other stakeholders: <ul style="list-style-type: none"> - policy pressure - readiness issues - expectations differences (alignment of staff) 	<ul style="list-style-type: none"> • Section 3.3: Criticisms of rubrics • Section 2.2.4: Resources and culture
	Theme 6 – Optimisation of rubrics in assessment practice	<ul style="list-style-type: none"> • Guidance and flexibility by teachers (from students) • Reinventing rubrics by students (from teachers) • Sharing good practice by management (from management) 	<ul style="list-style-type: none"> • Students view teachers as the authority and for providing instructions which is an Eastern style of learning; yet wanting flexibility is more a Western style – implying conflicts with learning in the HK culture • Teachers want students to be creative and build evaluative judgement. This is a Western style and fits into sustainable assessment concept • Management's view fits into the middle managers' roles in policy situation – being gatekeepers and translators 	<ul style="list-style-type: none"> • Section 2.2.2: Eastern vs Western teaching and learning style • Section 2.1.3: Sustainable Assessment • Section 3.4: Innovative methods to optimise actualisations of rubrics • Section 4.3: Recommendation for holistic assessment policies (middle managers' roles in policy enactment)
4. What are the policy	Theme 7 –	<ul style="list-style-type: none"> • Agree to have a policy to 	<ul style="list-style-type: none"> • Surface the topic – fits into one of 	<ul style="list-style-type: none"> • Section 4.1: Policy intention

Research Questions	Themes	Main points or sub-themes	Explanations	Bodies of Literature
implications according to the experience of educators and management in this policy initiative?	Promoting holistic assessment practice through policy	address the topic <ul style="list-style-type: none"> • Policy not holistic and not fit-for-purpose • Implementation not well thought-through 	the intensions of policy <ul style="list-style-type: none"> • Usually a top-down policy is not welcomed but not in this context – may be a cultural factor? • Not holistic and implementation issue – similar to what was found in the literature 	<ul style="list-style-type: none"> • Section 4.2: Reasons underpinning limited adoption of assessment policy in higher education

Table 8.1 Summary of Findings and Mapping of Bodies of Literature

8.1 Implicit Understanding of Assessment for Learning

Although the major interest of this study is to explore the concepts and practices related to rubrics, knowing how different stakeholder participants perceive and understand Assessment for Learning is crucial because rubrics are connected to the practice of this concept. Participants' views and knowledge on assessment practice should directly relate to their perspectives on rubrics. As such, the first research question – What are the perceptions and understanding of Assessment for Learning in students, educators and management? – investigates contextual views of the participants and themes extracted provides background information to further explain the findings related to rubric practice, as well as insights for policy implication.

Theme 1 – Elements in Assessment Facilitating Learning – suggests that participants at all stakeholder levels naturally regard assessment as a way to measure or evaluate learning outcomes. However, with explicit

discussions on the learning function of assessment, features of assessment that fit into the concept of Assessment for Learning are identified.

These features include the merit of using authentic assessment, the importance of providing feedback and allowing students to reflect on feedback in an assessment event, as well as the use of assessment for curriculum mapping. This finding suggests that the participants in this study have a certain amount of knowledge of the learning function of assessment; however, this is somewhat dominated by the measurement function. This interesting finding of the implicit understanding of Assessment for Learning is explained by three interactive factors: (1) the dominant assessment practice and traditional culture; (2) the level of assessment literacy and (3) contextual characteristics.

8.1.1 The Dominant Assessment Practice and Traditional Culture

Boud (2007) reports that measurement is still the dominant discourse of assessment in higher education worldwide while Knight (2007) displays that examinations are still the most common type of assessment. Although these two examples from the literature are from the Western context, the dominant discourse with examinations as the assessment focus is also reflected in the findings of this research as it was conducted in Hong Kong.

This finding is not surprising because Hong Kong's higher education system is under the influence of globalisation (Schoeb & Chong, 2019) and therefore cannot escape what is "deemed as dominance". Furthermore in the local context, the Confucian-heritage culture adds another layer to this

dominant discourse because examinations are highly valued in this society even in modern times (Carless, 2011). Various researchers (Berry, 2011; Wang & Brown, 2014) believe that this historical view shapes how students and teachers position and signify assessment in Hong Kong. Berry (2011) analyses the situation in the school system while Wang & Brown (2014) specifically investigate students' views in higher education. Despite taking place in different education sectors, both studies reveal that the acquaintance with examinations in Hong Kong is significant and complex enough to limit assessment reform. On the one hand, stakeholders experience the negative effects of examinations, while on the other hand, they are comfortable with examinations.

This claim is also the finding of this research where the participants repeatedly mention the cultural influence of using examinations as their traditional assessment practice. It is then fair to say that this cultural influence strengthens the dominant discourse and adds to the complexity of seeing measurement as the primary function of assessment. The consequence is that the participants resist abandoning a familiar practice and, therefore, limit change.

Boud (2007) and Carless (2011) state that the attention to quality assurance and procedural details in assessment policies also reinforces the measurement function of assessment. This point is explained by Meyer et al. (2010) who believe that the policy serves to raise awareness of a topic important to the context and therefore sends a powerful message to policy actors about the priorities of an institution.

On the contrary, Brown et al. (2011) state that supportive assessment policies can deliver a positive image of using assessment to facilitate learning. In this research, it is, however, unclear if the rubric policy has a similar influence on the participants' perceptions of assessment because the policy states both the learning and benchmarking functions of rubrics (refer to Appendix 1). Since policy analysis is not the interest of this research, it is impossible to identify if the rubric policy covers or fails to cover the understanding of Assessment for Learning.

8.1.2 Questionable Level of Assessment Literacy

Another factor possibly contributing to the implicit understanding of Assessment for Learning is the participants' level of assessment literacy. To reiterate, assessment literacy is defined as the conceptual understanding of assessment together with skills and intellectual abilities in self- and peer-assessment using technical approaches (Price, 2012).

Assessment literacy has been reported as insufficient worldwide. For example, Forsyth et al. (2015) state that the use of assessment terminology is inconsistent among academics in an university in the United Kingdom. Medland (2018) tested the knowledge of assessment literacy on external examiners and found that the knowledge level varied a great deal. Rezvani Kalajahi and Abdullah (2016) and Rhind and Paterson (2015) also display in their surveys the insufficient and variable level of assessment literacy of lecturers and teachers/students respectively.

Although this present research did not objectively test and survey the participants' knowledge level of assessment literacy and their use of terminology, their qualitative descriptions of assessment somewhat reflect a similar notion of suboptimal assessment literacy. For instance, the participants could mention some features but not the sophisticated features of Assessment for Learning. The knowledge of how these features build on each other is missing as well.

The participants in this research demonstrate little understanding of Medland's (2015) six elements of assessment literacy (a community sharing standardised assessment practice, a dialogue between all stakeholders when building assessment practice, knowledge and understanding of effective feedback, a programme-wide approach that looks at the alignment of assessment and outcome, the adoption of assessment that builds self-regulation, and a shared understanding of assessment standards). Therefore, their knowledge level cannot fit into what is called a conceptual understanding of assessment. This finding may mean that the participants' primary view of the measurement function of assessment is due to the lack of understanding of what assessment entails.

The management participants mentioned curriculum mapping as a feature in Assessment for Learning but without further elaboration of this terminology. Medland (2015) stipulates that effective curriculum mapping takes a programme-wide effort to align learning objectives, teaching activities and assessment tasks. In this study, the terminology of curriculum mapping may be present among the management participants but conceptualisation

and actualization at a programme-wide level are far from happening. Therefore, they too cannot be called assessment-literate.

8.1.3 Contextual Characteristics

So far the chapter has discussed the dominant measurement function, the Confucian-heritage culture and the sub-optimal level of assessment literacy preventing the participants from explicitly understanding Assessment for Learning. However, what is the reason for their implicit understanding? This sub-section attempts to introduce contextual uniqueness as a plausible explanation.

As described in Chapter 1, the institution in this study is regarded as an applied-discipline university where effective application of knowledge is the ultimate goal for most of its programmes. Compared to pure arts and sciences disciplines, applied-disciplines (e.g. physiotherapy, nursing, engineering, design) commonly use case-based, scenario-based or skill-based assessment.

In addition, many applied disciplines mandate internships in their curricula and these involve applying knowledge and skills in real work settings under the supervision of industrial professionals. These internships usually last for a few weeks and are normally graded; professional teachers often give feedback continuously in this period of time to facilitate students to perform at a job-required standard.

As such, concepts such as authentic assessment and feedback are rather familiar to teachers and students in the applied-disciplines (i.e. this research context). The educator participants in this research are not only familiar with authentic assessment but they are also quite skilled at designing a variety of authentic assessment types (e.g. building product prototypes, analysing implicit meanings of scripts) as described in the interviews. Although the participants are rather knowledgeable in authentic assessment, a proper understanding of theories and the terminologies governing their practice appear to be lacking in the absence of formal training.

Another contextual characteristic may not only apply to this university but to universities in Hong Kong in general. There is no doubt that higher education in Hong Kong and other parts of the world is becoming more globalised (Schoeb & Chong, 2019). Administrators, teachers and students often have exposure to assessment elsewhere before working or studying locally. Take the educator participants in this study as an example: 50% of them have prior exposure in a Western education system. Since the Western education system is often associated with formative assessment (Kennedy, Chan, Fok, & Yu, 2008), their previous exposure may also contribute to their understanding of various types of assessment.

Overall, the participants' knowledge of authentic assessment and feedback may come from their applied-discipline backgrounds and from previous exposure to non-traditional assessment. This knowledge is rather implicit unless the learning function is explicitly mentioned. There seems to be an opportunity to raise participants' awareness of their own knowledge of

Assessment for Learning, as well as theoretical knowledge of good assessment practice. Without this fundamental understanding, practice is diverse and unsystematic. The next section discusses the participants' views on the barriers and opportunities for Assessment for Learning. Their views provide insights on strategies to increase assessment literacy.

8.2 Barriers and Opportunities of Assessment for Learning – Explanations from Traditional and Contextual Perspectives

Research question 2 explores the barriers and opportunities of Assessment for Learning in this research context. Despite the participants in this study using authentic assessment in their regular practice, they expressed huge frustrations about adopting assessment tasks with a learning function. Theme 2 reveals these barriers which include traditional burdens and mismatches in the current assessment system. Building on the participants' strong knowledge application background, the opportunities they identified under Theme 3 centre on further promoting authentic assessment, during and even beyond the course of study. These points are explained in the following sub-sections.

8.2.1 Constituents and Implications of Barriers

All participants in this study pinpointed traditional examinations (but not authentic assessment) and over-assessment as major hurdles in adopting Assessment for Learning. The participants also questioned the validity of and expressed dissatisfaction with such systems, as with those who reported in the previous literature in the Hong Kong context (Carless, 2017; Wang &

Brown, 2014). This suggests that similar burdens apply to those participants with or without a strong background in authentic assessment. One might have thought the well-reported criticisms about Hong Kong are mature enough for teachers and administrators to abandon the examination-oriented practice; however, this is not the case. Assessment reform in Hong Kong has been challenging and the examination-focused culture has been blamed for limiting assessment change in primary and secondary schools (Berry, 2011; Brown et al., 2009). The findings of this part of the study reflect that the same challenge also exists in higher education contexts in Hong Kong.

Another burden shared by the participants in this study is over-assessment. According to Price, Carroll, O'Donovan, and Rust (2011), over-assessment means the imbalance between formative and summative assessment and signifies the bias towards summative assessment. Summative assessment is thought to be more resource-intensive because of the administrative workload required which therefore causes fatigue in teachers (Price et al., 2011).

This point, however, is somewhat in contrast to the findings of this study where most educator participants regard formative assessment as more time-consuming due to the need to produce constructive feedback. On the other hand, the student participants think the over-reliance on multiple-choice examinations from teachers is due to the ease of grading it. In addition, almost all participants refer to over-assessment as the overall number of assessment tasks rather than the balance between summative and formative assessment.

These differences may be due to the context of applied disciplines, where the participants greatly value authentic assessment. They hope for a reduction in summative assessment or examinations in general in order to focus on authentic assessment to the discipline and students. Despite the differences in views of labour intensity of formative versus summative assessment, the view on over-assessment concurs with Price et al. (2011) that “the dominance of summative results is probably what lies behind the cries of over-assessment” (p.488).

Another burden found in this study is the conception of mismatch, which implies the issue of assessment validity and in turn questions teachers’ assessment literacy. In particular, most participants are concerned about what their examinations are actually measuring. This validity issue is in line with previous research, where university students in Hong Kong were concerned about the inaccuracy of assessment (Brown & Wang, 2013) and primary/secondary school teachers found assessment irrelevant (Brown et al., 2011). This research supplements these findings in that this concern also exists among university teachers in Hong Kong, implying a universal issue of assessment validity.

From the participants’ descriptions of their frustrations, mismatch is not only an issue with assessment practice but is a broader concern of teaching pedagogies. Students are not often examined at which they are taught, and teaching contents/activities do not always align with learning objectives. This scenario may imply a general lack of assessment literacy, as being able to link assessment and outcome is one of the six assessment-literate elements

(Medland, 2015). Although the participants in this research provide authentic assessment examples, the overwhelming concern of assessment mismatch from both the educator and student participants may indicate that teachers at large still possess insufficient assessment literacy and therefore hinder a wider adoption of Assessment for Learning.

In sum, the participants in this research express their frustrations at the current assessment practice, the examination-oriented culture, over-assessment and the learning objective/assessment mismatch. In their view, these burdens limit changes in assessment practice and the adoption of Assessment for Learning. Some of the burdens found in this study, such as the examination-oriented culture and the inaccuracy of assessment, resonate with previous research and confirm the case for this one university in Hong Kong. Over-assessment, while a subtle concept in assessment literature, surfaces in this research and adds an additional layer of heaviness to the burdens the participants carry. Nonetheless, the participants anticipate hope by identifying opportunities for Assessment for Learning. The next sub-section turns to this focus.

8.2.2 Contributors and Implications of Opportunities

The traditional burdens of the assessment system in Hong Kong do not leave the participants feeling hopeless about the current assessment situation; in fact, they envision multiple ways of enhancing the assessment practice in their immediate context and nearly all of their suggestions align with features of Assessment for Learning. As illustrated in Theme 3, Ideal Actualisations of

Assessment for Learning, the participants at all stakeholder levels wish to see a realignment of learning objectives and assessment tasks, as well as the evolution of more authentic, continuous and sustainable assessments rather than an over-reliance on one-time examinations. This thought is the same as Carless' wish to scale up Assessment for Learning especially in Hong Kong (Carless, 2017). This finding can be explained similarly as in section 8.1.3, where the applied discipline background of the participants in this research context enable them to value assessments that are concerned with real-life applications.

It is worth highlighting the fact that the participants do not oppose all types of examination and do not suggest abolishing examinations altogether. This may imply that the participants still see the merits of examinations in this historical culture and this argument has been partly demonstrated in previous research, at least from the teachers' perspectives (Brown et al., 2011; Brown et al., 2009).

In this context, the participants welcome examinations that focus on applications, i.e. make students think as compared to merely regurgitating knowledge. In addition, examinations that build on each other and have a feedback mechanism (i.e. holistic and continuous) are helpful for students' learning. This conception from the participants implies that an examination can be transformed to something constructive, despite the fact that it is still an examination. It is then necessary to clearly define what the term 'examination' entails when studies criticise it and to offer solutions to the examination-oriented culture.

The findings of this study suggest that the problem of the current examination-oriented culture lies in using one-time examinations for memory testing, rather than the blanket term of examinations alone. Given that the examination-oriented culture is difficult to change in Hong Kong as discussed in the previous sections, transforming the nature of examinations may be an easier and a more acceptable option in the existing system.

Besides transforming the traditional nature of one-time examinations, the participants in this research state that real-life application assessment should be continuous and be sustained beyond the immediate assessment throughout one's career. This fits into the definition of Sustainable Assessment (Boud, 2000) where assessment should be designed to build students' evaluative judgement for life-long learning. In the context of applied disciplines, it means that students should be given the opportunity to learn how to judge their own and their peer's performance in every authentic assessment, so that they can do the same after graduating and working in real industrial settings as suggested by Boud and Falchikov (2007). The establishment of evaluative judgement is perhaps easiest during internships or service learning modules, through self-assessment and comparing it with the industrial teachers' evaluation. Currently, this comparison may not be formally and systematically done. It is worthwhile exploring the addition of this subtle yet beneficial practice to build evaluative judgement.

8.3 Summary and Implications of Conceptions, Barriers and Opportunities of Assessment for Learning

Sections 8.1 and 8.2 discuss the participants' conceptions, barriers and opportunities of assessment in general and of Assessment for Learning in this one university in Hong Kong. The participants in this study cannot escape from the global view of seeing assessment as measurement, which is complicated by their inadequate assessment literacy and the deeply-rooted examination-oriented culture in Hong Kong. Despite this view of measurement, the participants demonstrate knowledge of Assessment for Learning when explicitly asked. This may be due to their applied discipline backgrounds and previous exposure to Western education systems, where using formative and authentic assessment is a common practice.

Still, conceptual understanding of Assessment for Learning appears to be missing when compared to the definition of assessment literacy by Medland (2015). This indicates an enormous opportunity to incorporate assessment literacy into the higher education sector in Hong Kong. On the other hand, the burdens of the examination-oriented culture and over-assessment limit a focus on Assessment for Learning in this context. Despite these burdens, the participants suggest using more authentic, continuous and sustainable assessment to promote learning and minimise measuring.

The findings answering the first two research questions bring significant opportunities to increase assessment literacy in this one university in Hong Kong. The participants' practical knowledge of using authentic assessment could be the foundation of a general acceptance of this practice

and the university does not need to start from scratch to get approval. Rather, the university can educate the participants in theories underlying Assessment for Learning and focus on the systematic actualisation of practice. In this way, practice is evidence-based and more coherent. The thesis now turns to discussing the findings related to rubric actualisation under the mandatory rubric policy in this research context.

8.4 Feelings of Powerlessness in Actualising Rubrics in the Context of a Top-Down Policy

Research question 3 asks “Can rubrics actualise Assessment for Learning in the eyes of students, teachers and management? Why or why not?”. The answer to this question is not straightforward because the participants have experienced a fairly equal share of potential of and barriers to using rubrics alone and using rubrics to promote learning. This main finding aligns with the pros and cons of rubrics reviewed by Panadero and Jonsson (2013). In particular, the benefits of rubrics (i.e. the ‘yes’ answer to the research question) centre on features related to Assessment for Learning; while the barriers (i.e. the ‘no’ answer) concern pragmatic issues.

The findings of this study interestingly reveal tensions between stakeholder groups during rubric actualisation, which may be related to how the policy is implemented in this context. This apparent tension shows itself again when the participants share their thoughts on the ideal optimisation of rubrics. However, the tension appears to concern procedural issues rather than the values and functions of rubrics. This implies that fixing procedural

and pragmatic problems, such as rubric design and training (Rezaei & Lovorn, 2010), may create a favourable future for using rubrics to facilitate Assessment for Learning. These points are discussed one by one in the following sub-sections.

8.4.1 Potentials of Rubrics for Learning Despite Stakeholders'

Experience Level

Theme 4 identifies the potential of rubrics for learning in the eyes of the participants. Not surprisingly, the participants naturally see rubrics as a grading tool and that rubrics are primarily used for measuring whether students meet the listed standards. This is in resonance with the concept revealed in Theme 1 where the participants naturally regard assessment as measurement.

In fact, the rubric policy statement (Appendix 1) introduces rubrics as a tool to actualise the university's criterion-referenced approach to assessment where students should be *graded* according to the listed criteria. This message alone may have prompted stakeholders to position rubrics as a tool for grading because assessment means 'measurement' to them. When it comes to the utilisation of rubrics in higher education, the policy lists three purposes: (1) for the development of students' own expertise, (2) for internal moderation of student assessment results and (3) for benchmarking academic standards. Two out of three of these purposes fit into the concept of Assessment for Certification, where measurement is the key purpose of assessment. Therefore, in addition to the participants' ingrained concept of

the measurement function of assessment as discussed in section 8.1, it may be that the policy message has influenced the participants' conceptions of rubrics even though no participants mentioned this point explicitly.

As with Theme 1, the participants were aware of the learning features of rubrics when they were explicitly asked to focus on this function. The specific benefits they mention concur with the contemporary literature, which includes using rubrics to facilitate communication between teachers and students (Mauri et al., 2014), to self-regulate the assessment process (Andrade & Du, 2005; Bell et al., 2013), and to facilitate the feedback mechanism (Jonsson & Svingby, 2007). Jonsson (2014) argues that the benefits of rubrics are related to increased transparency of information when rubrics are available to students beforehand. This is also the case in this study where the educator participants increased visibility of information by posting, explaining or supplementing rubrics with exemplars. While the previous literature is all from the Western context and therefore, assumes that their stakeholders are familiar with rubrics and formative assessment, the similar findings from this research imply that the benefits of rubrics are quite universal and apparent and can be experienced by stakeholders despite contextual, cultural and knowledge differences.

In addition to the learning merits of rubrics, the participants also report grading and psychological benefits. For example, rubrics make grading expectations explicit and therefore enhance consistency and reduce nervousness. In a way this also suggests rubrics can enhance learning. These are established benefits of rubrics presented by Panadero and

Jonsson (2013) in their review paper. These findings again suggest that regardless of the level of knowledge and experience and willingness to use rubrics, having rubrics as compared to no rubrics (i.e. previous practice in this research context) can infuse some benefits to existing assessment practice. The next sub-section turns to look at another scenario, which is the barriers the participants experience during rubric actualisation.

8.4.2 Constituents and Implications of Rubric Actualisation Barriers

While the original aim of the research question concerns the challenges of using rubrics *as a learning tool*, barriers with rubric execution in general surfaced during the interviews. It is necessary to attend to these issues because they represent fundamental problems to effectively actualise rubrics, which include highlighting the learning function.

All participants criticise the inherent drawback of rubrics as one of the barriers of effective actualisation. These drawbacks are well reported in the contemporary literature. For example, Bennett (2016) criticises the fact that rubrics limit creativity because students may only follow the standards set forth in the rubrics. He also comments that rubrics fail to measure complex thinking skills because it is not feasible to list them in a grid. Carless (2015a) believes that there are hidden meanings in evaluative criteria, making rubrics subjective. Reddy and Andrade (2010) and Van Helvoort (2010) debate about rubric validity where many factors contribute to or limit grading accuracy, thereby complicating its use. Despite contextual, cultural and knowledge level differences, the participants in this study reveal almost identical concerns with

using rubrics. The finding once again signifies the universal and fundamental issues with rubrics.

With these well-reported disadvantages of rubrics, it is reasonable to assume that no perfect solutions exist. Supplementing rubrics with other means such as exemplars (Carless, 2015a), audio-feedback (Broadbent et al., 2018), or by introducing innovative implementations such as co-creating (Fraile et al., 2017) or demystifying rubrics (Jones et al., 2017) is proven to be more effective in unveiling the learning function of rubrics. The participants in this study do not mention these methods except for two educator/management participants who reveal the use of exemplars in their practice. This may be due to their lack of experience and knowledge with rubrics. Clearly, educating stakeholders on rubric innovation is an area that warrants further development and sheds light on policy refinement for this context .

Apart from the inherent limitations of rubrics, the participants in this study express huge pragmatic barriers when actualising rubrics. Pragmatic issues are illustrated in the previous literature such as the time and knowledge required to develop a fit-for-purpose rubric (Carless, 2015a; Van Helvoort, 2010). The participants in this study thought too that they lack support, time and knowledge to develop rubrics. Considering rubric actualisation in this institution comes from a mandatory top-down policy, one would expect a more well-rounded structure from the institutional level and the participants should have only experienced minimal issues on development, implementation and lack of organisational support. This unexpected finding

suggests that the institutional strategies taken to implement this policy may be too simplistic. Despite the fact that the institution provides workshops for teachers to learn how to develop rubrics, the participants' state that the workshops only focus on how to construct a rubric but without details on the quality of rubrics and how to effectively execute this practice.

In addition, even though a rubric policy means rubrics are an important initiative of the university as some participants mention, it is *just* a teaching-related policy where the priority of fruitful implementation is trumped by research. The stress on competitive research output in higher education culture dictates how teachers spend their time and effort (Bahous & Nabhani, 2015; Macdonald & Joughin, 2009; Raaper, 2016).

This culture described in the literature applies to this institution where the research took place because it is in the international ranking system. Hence, it is not surprising that despite a policy, teachers may still refuse to spend time on the development and implementation of rubrics (Tagg, 2012). This is also one of the criticisms reported in the literature on resistance to change in the area of assessment practice (Deneen & Boud, 2014). Evidently, this institution may need to consider an incentive or recognition system for carrying out teaching-related good practice (Norton et al., 2013), given that the research-first culture is impossible to change.

A very interesting and context-specific finding under barriers of implementation is opposition from other stakeholders in the process. The educator and management participants think students should adopt more

ownership in using rubrics. On the other hand, the students express the view that rubrics are vague and difficult to follow; teachers should provide more guidance and should use rubrics more consistently. The inconsistent practice also exists between teachers as one of the educator participants mentions.

This finger-pointing phenomenon indicates tensions between various stakeholders, and this tension may have arisen from the questionable readiness of rubric implementation. Chapter 1 mentions the policy is from a top-down approach and that some teachers had not heard of rubrics prior to this policy. The previous paragraph also reveals that the workshops provided by the university fail to practically address quality implementation. As such, teachers' readiness for effective rubric actualisation is questionable; yet they have to implement rubrics due to policy pressure. The rushed implementation may possibly lead to immature rubric use and therefore finger-pointing and overall dissatisfaction.

If teachers think they have already exhausted their knowledge on how best to use rubrics, they become frustrated when students express insufficient understanding of rubrics despite explanations. The reality is, teachers may not have recognised that some students are also new to rubrics and therefore do not understand the purpose and use. On the other hand, students may think they are only passive recipients of rubric implementation because there is no choice about whether to use rubrics or not. It is reasonable to assume that they have engagement difficulties if all they receive is a grid online with vague explanations.

As Carless (2015a) states, students need to be taught and given the opportunity to use evaluative criteria in order to feel engaged. Using rubrics, and, further, taking ownership of rubrics, is a process that needs to be nurtured. It appears that the tension arises from a mismatch in expectation of rubric use between students and teachers, and a lack of understanding on how best to utilize rubrics. This is once again a pragmatic and procedural issue rather than an issue with the function and value of rubrics. The identified barriers then create huge opportunities for this institute to use rubrics for the facilitation of learning, if these pragmatic issues can be tackled in a sophisticated manner. The next section discusses what the end or ideal goal looks like for rubric actualisation.

8.4.3 Contributors and Implications for Ideal Actualisation of Rubrics

Thoughts on the ideal actualisation of rubrics follow similar patterns to the barriers discussed in section 8.4.2. This is logical because the participants see the potential of rubrics (discussed in section 8.4.1); however this potential is mostly limited by implementation issues at the time of the study (section 8.4.2). As such, ideal actualisations are translated into ways of addressing these issues.

Relating to the research question as to whether rubrics can actualise Assessment for Learning, the answer is 'yes' if these ideal optimisations can be implemented in real life. However, it is currently far from ideal and therefore the answer is 'no' until at least these hurdles are overcome. Although there exist multiple factors involved in this, the focus of ideal

optimisation seems to centre on the responsibilities of other stakeholders. The end goal is to refocus the use of rubrics for learning. These points are discussed in the following paragraphs.

To recap, the student participants in this research think it is the responsibility of teachers to implement rubrics better. They look for more guidance and flexibility from teachers while utilising rubrics. This finding is intriguing because guidance and flexibility seem to be two opposite features in learning traits. Generally speaking, students who need more guidance from teachers are those who are more novel to the process (Kirschner, Sweller, & Clark, 2006), while those who value flexibility should be more mature in their learning and therefore want more autonomy and independence (Schwartz, n.d.).

Another perspective is that guidance signifies an Eastern-centred learning style while flexibility aligns with a Western-centred style (Ho, 2010). The opposite wishes from the student participants may indicate internal conflict and the questionable readiness of students for rubric utilisation. The student participants on the one hand want more autonomy in using rubrics and do not want rubrics to limit their creativity (one of the disadvantages of rubrics they identified). On the other hand, they desire guidance from teachers because of the fear of a less than optimal grade. They likely perceive teachers as the authority, a unique feature in Eastern-centred learning (Ho, 2010). Together with the value on high grades in this examination-oriented culture (Berry, 2011), this, no doubt contributes to their wish for more teachers' guidance. This internal conflict implies that students may not be ready to

uptake rubrics for fear of a suboptimal grade, which in turn limits ideal actualisation in teachers' perspectives.

This illustrates that the university should address issues at all stakeholder levels, including students, when implementing a new initiative or policy. One way to involve students is to allow their voices to be heard when framing educational policy in higher education. This action has been advocated to improve policy engagement (Poth et al., 2015) and can be considered by this institution in future policy initiatives. In addition, it is imperative to spend time to teach students to use rubrics as discussed previously. Without this knowledge, students may always live with this conflict and this in turn may impede the meaningful utilisation of rubrics. This relates back to the need of increasing assessment (and rubric) literacy in both teachers and students, an area neglected in this policy implementation.

From the educator participants' perspective, the independent adoption of rubrics by students was what they envisioned. The fact that the educator participants demand more independence from students signifies a Western-centred style (Ho, 2010) and this may be due to the teachers' background and experience as discussed in section 8.1.4. However, the educator participants do not seem to be aware of the complexity of rubric utilisation to achieve this goal, especially the readiness of the students to adopt it. What the educator participants describe as up-taking rubrics (e.g. reinventing the rubric) involves building evaluative judgement and is a concept situated in not only Assessment for Learning but Sustainable Assessment (Boud, 2007; Boud & Associates, 2010).

In order to achieve this goal, conceptual knowledge of Sustainable Assessment needs to be delivered to students and teachers who do not already possess this knowledge. Fundamental understanding of Sustainable Assessment is required to nurture students' abilities to use rubrics more effectively, and then criteria should be designed that are meaningful to them.

Fraile et al. (2017) provide a feasible example of how this can be done systematically by involving students in the co-creation of rubrics. Bearman and Ajjawi (2019) also suggest transforming students' engagement with rubrics by inviting students to produce, enact and reflect on descriptors. These examples can possibly be adopted by teachers in this institution as the first step towards reaching their goals. Without this kind of opportunity, it is difficult for students to know how to create their own criteria for an assessment task.

The management participants in this research state that they bear the responsibility to share and monitor good rubric practice. This is the only participant group who have explicitly reflected on their own accountability in this policy initiative. Also, the management participants have discussed actions they have implemented or are trying to implement. This is not only a dream that they would like to see but they have translated their words into actions. Their perspectives somewhat accord with how middle managers generally see themselves in a policy situation such as that described in the literature – being the translator and gatekeeper of policy (Saunders & Sin, 2015).

In this context, the management participants have tried to share their knowledge on rubrics, or at least to convey the message from top management on why rubrics are needed. They have tried to understand and believe in rubrics themselves and set a model of practice so as to engage their subordinates. This belief and/or action serves the translator role.

In addition, the management participants have positioned themselves as being accountable for the success of implementing the rubric policy, even though they might have faced similar challenges when using rubrics themselves. This implies that the management participants might have prioritised the management role over the teachers' role in this policy context. The university should not undermine the efforts and effects the managers contribute and bring to a policy initiative, and should empower them in the process.

8.4.4 Summary and Implications of Rubric Actualisation in the Research Context

Relating back to research question 3, section 8.4 discusses whether rubrics can actualise Assessment for Learning, given the influence of a mandatory policy in an examination-oriented culture. No definitive answer is drawn from the participants' views and experiences because they have encountered an equal share of benefits and drawbacks of rubrics while implementing this policy-driven initiative. Despite the fact that most of them believe in the potential of rubrics for students' learning, meaningful utilisation of rubrics is limited by pragmatic implementation issues. Since implementation has been

driven by a policy in this situation, this implies that the policy itself does not facilitate practical rubric actualisation but merely raises awareness of this concept.

A critical analysis of findings of this part of the study reveals a powerless feeling in this policy situation, as evident when demanding further involvement from other stakeholders. The management participants carry an extra layer of responsibility to promote and monitor the use of rubrics in this situation. Clearly, all stakeholders need to be empowered in this policy context before the initiative can be claimed to be successful and to reach its goals. The next and last section discusses the policy implications from this particular rubric policy experience.

8.5 Enhancing Policy Success through a Holistic and Deliberate Approach

The last research question asks “what are the policy implications according to the experience of educators and managers in this policy initiative?”. This question is addressed by looking at how the educator and management participants interpret the rubric policy and implications for practice. An interesting finding is that the participants in this research context have agreed to a top-down policy because they feel it is needed to unveil a topic.

This thought is partially in line with one of the common policy intentions (i.e. to raise awareness of an important area) reported in the literature (Tijds et al., 2012). However, in contrast to most of the literature where a top-down policy is usually not welcomed (Deneen & Boud, 2014; Firestone, 1998; Terhart, 2013), the participants in this study feel positive about such a policy.

It is interesting that they welcome this mandatory top-down policy even though they have expressed frustration and turmoil with policy implementation.

One explanation may be related to the cultural factor where respecting authority seems to be a normal practice within the Confucian-influenced societies. However, given the diverse backgrounds of the participants as discussed in section 8.1.4, this cultural factor cannot be the sole explanation for respecting authority. Further studies should investigate this interesting scenario in the Hong Kong context. Nonetheless, the positive attitude towards a policy paves the path for using this policy for teaching enhancement.

Even though the participants do not oppose a top-down policy, the participants' frustrations with policy implementation have restricted the success of this initiative. The details of the frustrations have already been discussed under section 8.4.2. While there is no comparable literature in this area, the frustrations experienced by the participants are, in fact, common policy pitfalls identified in the literature, such as not being holistic and not deliberate (Boud, 2007; Carless, 2017).

Specific to assessment policies, the criticisms reported in the previous literature concern the focus of policy on the assessment procedure and on quality assurance rather than on student learning (Carless, 2017). The participants in this study did not mention their exact criticism; however, some participants stated that this policy only focuses on one part of the assessment system (i.e. rubrics) but lacks a holistic picture of what assessment in higher education should be. This point warrants attention from the university's top management during this policy review or when planning a new teaching

policy. It implies that this and perhaps other universities should use the available holistic policy framework as a reference point in order to deliberately carry out a policy (Meyer et al., 2010; Poth et al., 2015).

The premature implementation of this policy was also a huge cause of annoyance among the participants. Despite training provided by the university, the content of the training hardly met the participants' expectations because it only focused on the groundwork of creating rubrics but not on the theories behind them, let alone the implementation strategies. The participants perceived that the policy was launched in a rush and the implementation plan was not meticulous.

This perception may be partially explained by the participants' inadequate assessment literacy (as discussed in sections 8.1.2 and 8.2.1). When they lack knowledge of the diverse functions of assessment and rubrics, they need more time and training to understand the underlying concepts of this practice. However, the university training only focused on rubric development but deprived them of fundamental knowledge of assessment and rubrics. This resulted in rushing a practice where stakeholders do not feel ready, which in turn causes frustration and helplessness during the implementation process.

The educator participants envisioned students taking more responsibility when using rubrics (refer to section 8.4.3). Apart from teaching students evaluative judgment to increase their engagement with rubrics (Carless, 2015a, 2017), students' responsibility or accountability can be enhanced by including them in policy initiatives (Poth et al., 2015). Gravett et al. (2019) suggest using the 'students as partners' approach to facilitate the

transformation of institutional cultures. This approach includes engaging students and staff through communication, dialogue and community. These are strategies this university can consider when reviewing this rubric policy or when drafting new policies. After all, students are important stakeholders of any teaching policy or initiative; their voice should be heard for a more productive implementation.

All in all, this rubric policy experience has provided valuable lessons for refining the existing policy or developing new ones. In particular, the policy needs to be more holistic by including other domains of assessment practice rather than just one assessment tool. Policy implementation should be well planned by taking into consideration stakeholders' assessment literacy level and by including opinions from stakeholders of all levels. These measures may increase policy engagement and enactment. This context seems to have an advantage for policy success because the participants are open to a top-down policy. This somewhat positive attitude helps to set the stage for policy success and this will only come when the above-mentioned problems are tackled.

8.6 Summary of Discussions

This chapter analyses the findings and explains their implications using contemporary literature and contextual characteristics. The implicit understanding of Assessment for Learning and rubrics found among the participants is explained by the dominant discourse of assessment in higher education, the cultural perspective and the general insufficient level of assessment literacy. These are factors discussed in contemporary literature

and are also found to be applicable to the Hong Kong context. However, this study discovers the presence of some knowledge of learning with assessment and rubrics which may be related to the participants' applied discipline background and globalised experience. This new finding contributes to understanding the knowledge and practice of the assessment of stakeholders in this one university in Hong Kong. The understanding of Assessment for Learning and rubrics still appears to be inadequate and this analysis concurs with Carless' call (2017) for scaling up this concept.

The findings on rubric actualisation also contribute to understanding the barriers of executing good practice in this context. The well-reported merits and barriers of rubrics are also perceived by the participants in this context, meaning that these characteristics are rather universal despite the users' experience level. However, the actual experience of the participants (novice users) is complicated by inadequate knowledge, readiness issues and immature policy implementation in this institution. As such, they feel powerless in adopting the practice. This study discovers and explains these context-specific problems that need to be rectified before a more fruitful actualisation can occur.

The positive attitude towards policy revealed by the educator and management participants is an unexpected finding and this analysis contributes to understanding that policy may be a good way to unveil the importance of a topic in this institution. Nonetheless, a holistic policy and a deliberate implementation plan are critical for success. Critical analysis reveals that stakeholders need to be empowered to enact a policy and a teaching and learning initiative.

Chapter 9: Contributions to Knowledge and Conclusions

Knowing the interpretations of the study findings, this chapter focuses on how this research contributes to knowledge in the higher education assessment field, for this university and other institutions alike in promoting Assessment for Learning through policy. The chapter first provides a conclusion of how the research questions are answered. It then highlights the contribution to knowledge in the literature at both the local level and beyond by suggesting a model of holistic assessment policy and initiative (Figure 9.1). The chapter closes with study limitations and recommendations for future research.

9.1 Conclusion of the Findings

This present research aims to explore the perceptions and experiences of Assessment for Learning and rubrics under the context of a mandatory top-down rubric policy and an examination-oriented culture in one university in Hong Kong. Under this over-arching aim are four research questions addressing the concepts, barriers and opportunities with Assessment for Learning and rubrics, as well as the experiences of the two under a policy initiative. The results of this research add to the body of literature by outlining the struggles stakeholders faced when actualising Assessment for Learning and rubrics in a specific context in Hong Kong, despite the benefits they have experienced and a policy they appreciate, at least on the surface. The overall implication is a reflection on how the practice of Assessment for Learning can be better promoted through policy in a deeply-rooted examination-oriented culture. The recommendations do not only apply to the university where this

research is conducted, but no doubt also to similar universities within Hong Kong or other Confucian-influenced countries.

The first research question concerns the participants' understanding of Assessment for Learning and the findings display that their conceptual knowledge is limited and implicit due to the dominant view of assessment function (for measurement), a lack of assessment literacy and the influence of the Confucian-heritage culture. Nonetheless, knowledge and practice of Assessment for Learning is present to some extent and is likely because of the background and exposure of the participants, in particular in the applied disciplines. It implies that Assessment for Learning knowledge needs to be scaled-up and be more explicit for stakeholders in order for them to be cognisant of the concept.

The second research question explores the participants' thoughts and experiences on the barriers and opportunities of Assessment for Learning. A heavy focus on examinations, over-assessment and an assessment-objective mismatch were identified as barriers by the participants. This suggests that the deeply-rooted examination-oriented culture and insufficient assessment literacy influence how stakeholders practise Assessment for Learning. Despite the lack of conceptual understanding of assessment practice among the participants, their methods of using assessment to facilitate learning match with some of the features of Assessment for Learning and Sustainable Assessment (Boud, 2000; Carless, 2015a; Sambell et al., 2013). This again may be related to their strong applied discipline backgrounds and it reinforces

the point that the concept of Assessment for Learning needs to be made explicit among stakeholders.

The third research question investigates if rubrics can actualise Assessment for Learning, in particular under the mandate of a top-down rubric policy. The participants identified the potential of rubrics as a learning tool but also expressed huge barriers in developing and using rubrics for learning. They envisioned ideal ways of actualising rubrics, such as the engagement of students, guidance and the flexibility of teachers, as well as mediation of management. The participants clearly displayed powerlessness in this policy experience and their ideal picture of how the rubric policy can be actualised is instrumental in building empowerment structures in this context.

As such, the fourth and last research question addresses the lessons learnt from this policy initiative and what an ideal assessment policy should look like for facilitating good assessment practice. This policy, even though it is top-down, is welcomed by the participants because it can assist teachers in good practice. However, this policy is fragmental and policy implementation is too rushed and unsophisticated, causing frustration and blame among stakeholders. These analytical findings provide insights for policy refinement and suggestions for future policy initiatives, which are discussed in the following sections.

9.2 Contribution to Knowledge

This study does not only contribute to the literature in the field of Assessment for Learning, rubrics and assessment policy in higher education, it also offers

recommendations for the university where this research is conducted and for other higher education institutions in Hong Kong and Confucian-influenced countries. The following subsections discuss the contributions accordingly.

9.2.1 Contribution to the Literature on Assessment for Learning, Rubrics and Assessment Policy in Higher Education

Several new findings emerge from this research and add to the body of literature on Assessment for Learning, rubrics and assessment policy. Firstly, this study reveals tensions between stakeholders in their experiences with rubric actualisation. This finding is only available through including three stakeholder groups as the participants in this research, a design rarely used in previous studies. Interestingly, dialogues display that the participants are merely reflecting on their own responsibilities in this rubric actualisation experience and in the ideal utilisation of rubrics, with the exception of the management participants who state their extra layer of duty in convincing other participants to accept the policy. This adds to the knowledge they need to tackle the complex human interaction when implementing an unfamiliar assessment initiative. It also contributes to the literature that this complexity can be in part addressed by enhancing assessment literacy in all stakeholders.

Secondly, this study contributes to the knowledge that context matters in an assessment-specific initiative. While Deneen and Boud (2014) affirm how context-specific culture shapes resistance to assessment change in Hong Kong, this present study adds to the knowledge the facilitation factors

for change, such as the possible contribution of the applied disciplines background and the acceptance of policies in general. This contribution to knowledge is critical when considering future assessment policy initiatives locally and regionally - the context-specific factors need to be investigated carefully and dynamically as they inform the fine line between resistance to and facilitation of change.

Assessment policies in higher education in Confucian-influenced countries are rarely reported in the literature (Carless, 2017). When reported, they primarily focus on the procedural and/or quality assurance aspects (Boud, 2007; Carless, 2015a). The findings of this study contribute to the literature on higher education assessment policy with a focus on learning and on Hong Kong. Although policy analysis is not the central attention of this study, reflections on the policy experience allow the disentangling of elements leading to the pitfalls (e.g. fragmental policy) and successes (e.g. policy message) of the rubric policy in this context. This finding informs higher education scholars and administrators of an assessment policy situation in Hong Kong, as well as providing implications of this experience, a knowledge rarely found in the literature.

9.2.2 Contribution to the Institution

This sub-section discusses how the new knowledge generated in this research gives opportunities to the university where this present study was conducted. Firstly and most importantly, it is imperative to scale up the conceptual knowledge of Assessment for Learning (and perhaps Sustainable

Assessment) and rubrics in the university. Stakeholders at the university are already advantaged by having fundamental exposure to some features of Assessment for Learning and Sustainable Assessment because of their applied discipline backgrounds. The starting point should be easier when compared to teachers without this experience in their teaching practice.

The purpose of scaling up the conceptual knowledge of Assessment for Learning and rubrics is two-fold. For one, it helps to build or strengthen the theoretical concept of assessment practice in teachers and students. Secondly, it makes Assessment for Learning practice explicit. The two purposes combined can potentially address the tensions between stakeholders and battle against the deeply-ingrained concept of Assessment for Certification, especially in the Confucian-heritage culture. It is important to note that making Assessment for Learning explicit does not mean to abolish examinations altogether. Rather, a more balanced assessment practice with a goal to promote learning is advocated through an enhanced understanding of the concept of Assessment for Learning .

Since the participants in this context had a positive attitude toward the policy, the university top management can build on this advantage and use a policy to bring out the desired effects. However, from this policy experience, much work is needed in order to achieve this goal in this university. Drawing upon the findings of this research, the following ingredients are necessary for a better assessment policy initiative:

- The policy paper should include information on the fundamental knowledge of assessment functions and common assessment practice, then go deeper into the theoretical concepts of Assessment for

Learning and Sustainable Assessment, as well as the use of rubrics as a learning tool. Input from students, teachers and middle managers should be sought to increase accountability and awareness. A student version of the policy should be available to make students cognisant of this practice.

- In addition to rubrics, other features situated in the context of Assessment for Learning and Sustainable Assessment should be promoted to make assessment a holistic practice.
- Training provided by the university should include the concepts of assessment, activities to actualise practice and the development and implementation of strategies. Also, teachers should be educated on how to teach students to use rubrics in the context of Assessment for Learning, for example, building evaluative judgement through various learning activities. The goal of training should focus on increasing the assessment literacy of teachers and students.
- To empower policy enactment, systems and strategies to empower middle managers' roles should be in place. In addition, the university should consider building an incentive and recognition system for good practice. These acts can potentially enable teachers to understand the value of teaching in the university culture.

The following diagram illustrates these suggested ingredients and the process for a holistic assessment policy initiative and strategies to promote the concept and actualisation of Assessment for Learning.

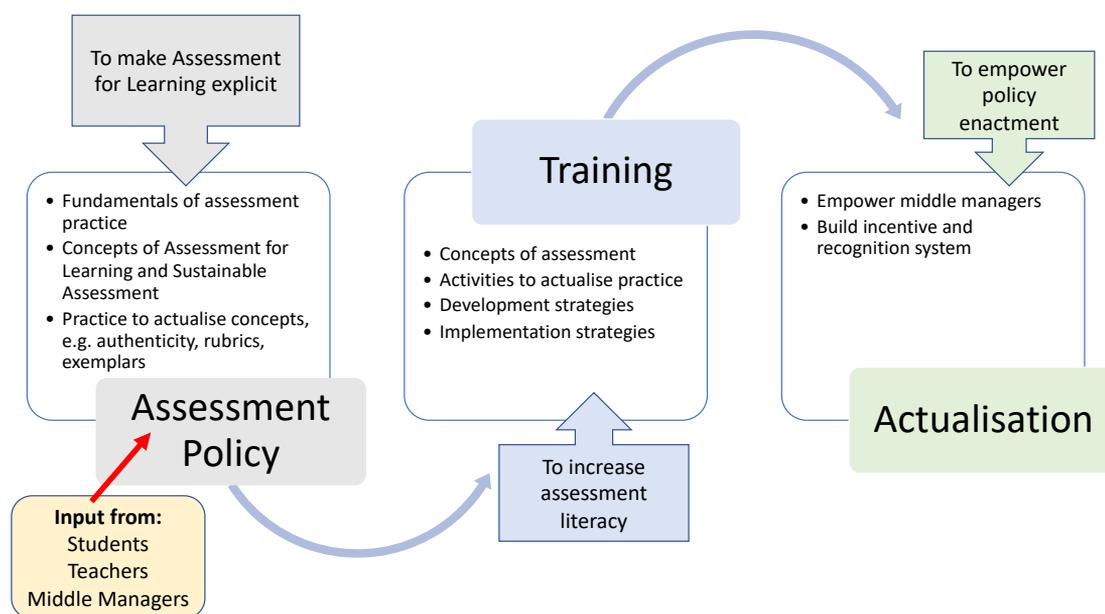


Figure 9.1 Suggested Model of a Holistic Assessment Policy Initiative and Strategy

9.2.3 Application to Other Local and/or Regional Institutions

The new knowledge of this research and the model suggested in Figure 9.1 can be applied to other local universities and/or institutions in the Confucian-influenced region. Despite contextual differences, some similarities of this and other local/regional universities are likely to be present, which allows the adoption of the suggested model to some extent. For example, this research context situates in the Confucian-heritage culture yet may be marginalised by the globalisation of staff and students (Schoeb & Chong, 2019). This unique circumstance is likely to be the case in many universities in Hong Kong and other Confucian-influenced metropolitan cities such as Singapore and in Malaysia (O'Neill & Chapman, 2015). As such, managers, teachers and students may have been influenced by previous exposure to features of

Assessment for Learning. This makes Assessment for Learning easier to receive. In addition, these institutions are still under the influence of the Confucian-heritage culture and may accept a policy the same way as the participants in this research context. That is, they may also welcome a policy even though it comes from the top. Therefore, the suggested holistic assessment policy model as illustrated in Figure 9.1 can be used as a generic guideline for other universities in Hong Kong and Confucian-influenced countries with similar characteristics.

While the model can provide similar institutions with a general direction for an assessment policy initiative, the specificity of how a university approaches a policy differs because context matters. As explained in Chapter 8, the university where this research was conducted has a unique and strong applied discipline background where authentic assessment has been the norm of practice. For this reason, it is assumed that teachers and management possess a certain degree of anecdotal knowledge in this area and training is suggested to focus on making Assessment for Learning knowledge explicit and actualisable.

For other universities in Hong Kong or Confucian-influenced countries where this background and knowledge may not exist, the focus of the policy and the corresponding training may need to be adjusted. It is imperative to understand the stakeholders' baseline knowledge of assessment practice, in order to design training suitable to their needs and intellectual levels. Nonetheless, organisational cultural and contextual issues (Deneen & Boud, 2014; Pham & Renshaw, 2014) need to be considered as the suggested

model is unlikely to address all the aspects of the uniqueness of an institution. Also, workload is a major challenge for academics (Deneen & Boud, 2014). An institution needs to look at these complicated and inter-related factors across the board so as to achieve success in an assessment initiative.

In addition, there are three major ways where other higher education institutions situated in this culture can use the study results to inform institutional policy and practice. Firstly, given the importance of assessment is to all stakeholders in higher education, the benefits of balancing the measurement and learning functions of assessment, and the general acceptance of having a policy to guide practice, institutions should explore their existing assessment practice and consider creating a policy on good assessment practice. As previously mentioned, a policy can be a powerful way to convey an important message and in turn change practice. Institutions should consider having a policy on assessment practice with a learning focus so as to preach for the shift of assessment functions.

Secondly, when it comes to using an assessment policy to guide practice, it is imperative to devise a holistic policy. The results of this study inform that a holistic policy is required to tackle the broad concept of Assessment for Learning. A policy on a small portion of a bigger concept does not appear to serve a good initiative well because it is difficult for stakeholders to visualise the big picture of the ideal practice.

Most importantly, the results of this study inform institutions that implementation of a policy is as important as, if not more important than the

policy itself. Before an institution executes a policy, deliberate effort should be made to educate stakeholders on the rationales behind such policy. In addition, fundamental knowledge that stakeholders need to understand for engaging such policy should be given visible attentions. This is possible through launching institutional-wide forum and mandatory training.

9.3 Study Limitations

There are four major limitations of this present research. First, results from qualitative studies are often not generalisable because of their small scale. However, qualitative studies are often not positioned to claim generalisability. Instead, rich and illuminating insights from the findings are their merits. Subjectivity is a major drawback in a qualitative approach as discussed in the Methodology and Methods chapter. Although reflexivity (Cousin, 2008) has been practiced throughout this present research, the qualitative findings may only apply to those individuals or institutions with similar experience or policy context respectively.

Measures used in this research to tackle this limitation include involving three stakeholder groups as informants, so as to ensure a wider representation of perspectives. Still, the relatively small number of participants in each participant group may have limited the representation and therefore generalisability. Nonetheless, data richness and saturation has been confirmed by the author and the supervisor of this study.

The second limitation of this present research concerns the timing of study and data collection. Data were collected during the first year of policy

implementation in the research context. This probably means most stakeholders were facing change, and with any change comes a storming period (Cantore & Passmore, 2012). It is unknown and beyond the scope of this research to investigate if some of the powerless feelings about this policy implementation were related to underlying resistance to change, rather than to the factual barriers stipulated by the participants. In addition, as the policy implementation continues, improvement plans are being launched according to sources of information at the university, such as workshops on the fundamental concepts of rubrics. The experiences and perceptions of stakeholders may have been different if the study had been conducted this year, or if there had been a follow-up data collection period on perception change as the policy became more mature. Hence, the results and implications of this research may only apply to those situations where policy implementation is rudimentary.

Third, this present study lacks rubric samples for readers to appreciate the varieties of rubric quality and actualisation pattern. This information would have added to the understanding of the barriers and opportunities with rubric actualisation, since the varieties might have contributed to how difficult or easy it was to adopt rubrics. During the semi-structured interviews, two educator participants briefly showed their rubrics to the researcher. However, due to privacy concerns these samples were not collected. In retrospect, the researcher could have attempted to seek consent to gather and include some rubric samples in this thesis, with the goal to aid understanding of the full actualisation picture.

Lastly, while this present research has delved into the experience of the implementation of a policy, policy analysis was not conducted in-depth for several reasons. First of all, the policy statement (Appendix 1) at the time of the study was very brief with limited information available. In addition, the educator and management participants of this study focused on developing and using rubrics, and whether or not the department met the target number of rubrics utilised. Hence the conversation became utilisation- and barrier-focused. Furthermore, due to the scope of this study and the need to explore perceptions of Assessment for Learning and rubrics, breadth was traded over depth in some topics, for example the deeper factors of welcoming a top-down policy was not explored. Separate research is suggested to tackle policy factors in-depth.

9.4 Recommendations for Future Studies

The findings and implications of the present study provide directions for future research. Within the research context, a follow-up study can be conducted to explore how the participants or various stakeholders perceive and experience the rubric policy after refinement. As expressed previously, more training, especially on the fundamentals of rubrics, has been delivered by the university after the data collection period of this research; it is therefore reasonable to assume that the participants' experience on rubric policy implementation has changed. Since including the fundamentals of Assessment for Learning and rubrics is a suggestion in the proposed holistic assessment policy model, it is intriguing to find out if this knowledge can indeed alleviate some of the barriers with this policy implementation. Findings

will shed light on whether this direction is appropriate, or if other implicit or hidden factors hinder the actualisation of this policy.

Another recommendation for future research centres on the area of assessment policy in higher education in Hong Kong. Since no previous literature has reported on assessment policy in Hong Kong's higher education contexts, an in-depth policy analysis of this initiative can provide local scholars with insights on unique factors leading to successful or failed policy development and implementation. The policy analysis should begin with understanding the policy makers' rationale and the objectives of the policy. An all-rounded evaluation of policy outcomes should be conducted to map with the policy objectives and implementation contents and procedures. This in-depth evaluation allows policy makers of this university and other institutions to learn the nuances and to use them as a reference for future initiatives.

Lastly, it would be interesting to investigate the similarities and differences in the perceptions of Assessment for Learning of teachers/students in applied-disciplines versus non-applied-disciplines by conducting a comparable study in stakeholders of the non-applied-disciplines in other universities in Hong Kong. It is assumed that knowledge on Assessment for Learning among the participants of this study came from their applied-discipline backgrounds. However, this is only a reasonable assumption that needs confirmation.

If differences are indeed identified, the hypothesis can be supported and the strategies used to promote Assessment for Learning in these

disciplines should be different. If no differences are found, it may indicate stakeholders in general are exposed to the Assessment for Learning concept despite the influence of cultural and dominant views. Also, knowing to what extent stakeholders are exposed to the Assessment for Learning concept contributes to shaping the assessment policy and initiatives needed (e.g. training) for assessment change.

9.5 Conclusion

This study contributes to understanding students', teachers' and managers' perceptions and experiences of Assessment for Learning and rubric actualisation under the influence of a top-down rubric policy in one university in Hong Kong. The qualitative findings reveal that although the participants' understanding of Assessment for Learning and rubrics are present, they are implicit and not at the level defined as 'assessment-literate'. This finding links to the huge barriers for actualising the concept and using rubrics under the policy. The participants found that the rubric policy is not holistic with unsophisticated implementation strategies, causing tensions within and between stakeholder groups and resulting in a powerless feeling during the actualisation process.

This seemingly negative finding creates opportunities for Assessment for Learning and rubrics in this institution. However, from what the participants had envisioned, much work is needed to refine the existing policy for a more holistic implementation. In addition, knowledge of assessment practice needs

to be enhanced. As such, a holistic assessment policy model (Figure 9.1) is recommended based on critical analysis of this study's findings.

This model suggests an assessment policy should begin with making various assessment functions explicit. All stakeholders need to be educated through the policy paper and associated training on the conceptual basis of Assessment for Learning and Sustainable Assessment and on the practical ways to execute this concept, including but not limited to rubrics. At the basic level, teachers need to be educated on teaching and giving opportunities to students to engage in the evaluative criteria of rubrics. It is also suggested that all stakeholders be included when drafting or refining a policy. Input from various stakeholder groups can enhance approval and enactment.

While this suggested model applies directly to this research context, it can also provide insights to other universities in Hong Kong and in certain Confucian-influenced countries where similar characteristics are shared, for example those influenced by the Confucian-heritage culture yet marginalised by globalisation.

Scaling up Assessment for Learning is imperative (Carless, 2017). Using policy may be an effective way to make Assessment for Learning explicit and to empower the actualisation process, especially in a culture where examinations are heavily ingrained and valued. The policy needs to be holistic and implementation should be deliberate. Despite good policy intention and implementation, there are still historical and structural issues (e.g. research is valued over teaching) that require attention. Nonetheless,

change takes time and effort. Assessment change is always an uphill battle against tradition. With persistent effort and well-rounded strategies, the change will one day occur.

References

- Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, 53(1), 27-31.
- Andrade, H. G., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, 10(3), 1-11.
- Andrade, H. L., Wang, X., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *Journal of Educational Research*, 102(4), 287-302. doi:10.3200/JOER.102.4.287-302
- Ashwin, P., Boud, D., Coate, K., Hallett, F., Keane, E., Krause, K., . . . Tooher, M. (2015). *Reflective teaching in higher education (Reflective teaching series)*. London: Bloomsbury.
- Ashwin, P., Deem, R., & McAlpine, L. (2016). Newer researchers in higher education: Policy actors or policy subjects? *Studies in Higher Education*, 41(12), 2184-2197. doi:10.1080/03075079.2015.1029902
- Ashwin, P., & Smith, K. (2015). Researcher creations? The positioning of policy texts in higher education research. *Higher Education*, 69(6), 1007-1018. doi:10.1007/s10734-014-9819-9
- Atkins, L., & Wallace, S. (2012). Insider research. In *Qualitative research in education* (pp. 47-63). London: SAGE Publications Ltd.
- Bahous, R., & Nabhani, M. (2015). Faculty views on developing and assessing learning outcomes at the tertiary level. *The Journal of General Education*, 64(4), 294-309. doi:10.5325/jgeneeduc.64.4.0294
- Bearman, M., & Ajjawi, R. (2019). Can a rubric do more than be transparent? Invitation as a new metaphor for assessment criteria. *Studies in Higher Education*. doi:10.1080/03075079.2019.1637842
- Bearman, M., Dawson, P., Bennett, S., Hall, M., Molloy, E., Boud, D., & Joughin, G. (2017). How university teachers design assessments: A cross-disciplinary study. *Higher Education: The International Journal of Higher Education Research*, 74(1), 49-64. doi:10.1007/s10734-016-0027-7

-
-
- Bell, A., Mladenovic, R., & Price, M. (2013). Students' perceptions of the usefulness of marking guides, grade descriptors and annotated exemplars. *Assessment & Evaluation in Higher Education*, 38(7), 769-788. doi:10.1080/02602938.2012.714738
- Bennett, C. (2016). Assessment rubrics: Thinking inside the boxes. *Learning and Teaching: The International Journal of Higher Education in the Social Sciences*, 9(1), 50-72.
- Berry, R. (2011). Assessment trends in Hong Kong: Seeking to establish formative assessment in an examination culture. *Assessment in Education: Principles, Policy & Practice*, 18(2), 199-211. doi:10.1080/0969594X.2010.527701
- Biggs, J. B. (1996). Western misperceptions of the Confucian-heritage learning culture. In D. Watkins & J. Biggs (Eds.), *The Chinese learner: Cultural, psychological and contextual influences* (pp. 45-68). Hong Kong: Comparative Education Research Centre and Australian Council for Educational Research.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. doi:10.1080/0969595980050102
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. doi:10.1080/713695728
- Boud, D. (2007). Reframing assessment as if learning were important. In D. Boud & N. Falchikv (Eds.), *Rethinking assessment in higher education: Learning for the longer term* (pp. 14-25). New York, NY: Routledge.
- Boud, D., & Associates. (2010). Assessment 2020: Seven propositions for assessment reform in higher education. . Retrieved from https://www.uts.edu.au/sites/default/files/Assessment-2020_propositions_final.pdf
- Boud, D., & Falchikov, N. (2007). Introduction: Assessment for the longer term. In D. Boud & N. Falchikv (Eds.), *Rethinking assessment in higher education: Learning for the longer term* (pp. 3-13). New York, NY: Routledge.

-
-
- Broadbent, J., Panadero, E., & Boud, D. (2018). Implementing summative assessment with a formative flavour: A case study in a large class. *Assessment & Evaluation in Higher Education*, 43(2), 307-322. doi:10.1080/02602938.2017.1343455
- Brown, G. T., Hui, S. K., Yu, F. W., & Kennedy, K. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5), 307-320. doi:10.1016/j.ijer.2011.10.003
- Brown, G. T., Kennedy, K., Fok, P. K., Chan, J. K., & Yu, W. M. (2009). Assessment for student improvement: understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347-363. doi:10.1080/09695940903319737
- Brown, G. T., & Wang, Z. (2013). Illustrating assessment: how Hong Kong university students conceive of the purposes of assessment. *Studies in Higher Education*, 38(7), 1037-1057. doi:10.1080/03075079.2011.616955
- Cantore, S., & Passmore, J. (2012). *How do people cope with change?: Phases of change models*. London: Kogan Page.
- Carless, D. (2007). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44(1), 57-66. doi:10.1080/14703290601081332
- Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. New York: Routledge.
- Carless, D. (2015a). *Excellence in university assessment: Learning from award-winning practice*. Abingdon, Oxon: Routledge.
- Carless, D. (2015b). Exploring learning-oriented assessment processes. *Higher Education*, 69(6), 963-976. doi:10.1007/s10734-014-9816-z
- Carless, D. (2017). Scaling up assessment for learning: Progress and prospects. In D. Carless, S. M. Bridges, C. K. Chan, & R. Glofchesky (Eds.), *Scaling up assessment for learning in higher education* (pp. 3-18). Singapore: Springer.

-
-
- Cousin, G. (2008). *Researching learning in higher education: An introduction to contemporary methods and approaches*. In. Retrieved from <<http://www.myilibrary.com?ID=193156>>
- Cox, B., McIntosh, K., Reason, R., & Terenzini, P. (2011). A culture of teaching: Policy, perception, and practice in higher education. *Journal of the Association for Institutional Research*, 52(8), 808-829. doi:10.1007/s11162-011-9223-6
- Cox, B. E., Reason, R. D., Tobolowsky, B. F., Brower, R. L., Patterson, S., Luczyk, S., & Roberts, K. (2017). Lip service or actionable insights? Linking student experiences to institutional assessment and data-driven decision making in higher education. *The Journal of Higher Education*, 88(6), 835-862. doi:10.1080/00221546.2016.1272320
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston: Pearson.
- Deneen, C., & Boud, D. (2014). Patterns of resistance in managing assessment change. *Assessment & Evaluation in Higher Education*, 39(5), 577-591. doi:10.1080/02602938.2013.859654
- Denzin, N. K., & Lincoln, Y. S. (2005). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 1-32). Thousand Oaks, California: SAGE Publications Inc.
- Ebersole, T. E. (2009). Postsecondary assessment: Faculty attitudes and levels of engagement. *Assessment Update*, 21(2), 1-2. doi:10.1002/au.212
- Elliott, J. (2012). Gathering narrative data. In S. Delamont (Ed.), *Handbook of qualitative research in education* (pp. 281-298). Glos, U.K.: Edward Elgar Publishing Limited.
- Ewell, P. (2006). *Applying student learning outcomes concepts and approaches at Hong Kong higher education institutions: Current status and future directions*. Retrieved from Hong Kong: https://www.eduhk.hk/obl/files/OBA_2nd_report.pdf

-
-
- Firestone, W. A. (1998). A tale of two tests: Tensions in assessment policy. *Assessment in Education: Principles, Policy & Practice*, 5(2), 175-191. doi:10.1080/0969594980050203
- Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students conceptions of assessment in higher education. *Higher Education: The International Journal of Higher Education and Educational Planning*, 64(1), 119-133. doi:10.1007/s10734-011-9484-1
- Flint, N., & Johnson, B. (2010). *Towards fairer university assessment: Recognizing the concerns of students*. London: Routledge.
- Forsey, M. (2012). Interviewing individuals. In S. Delamont (Ed.), *Handbook of qualitative research in education* (pp. 364-376). Glos, U.K.: Edward Elgar Publishing Limited.
- Forsyth, R., Cullen, R., Ringan, N., & Stubbs, M. (2015). Supporting the development of assessment literacy of staff through institutional process change. *London Review of Education*, 13(3), 34-41. doi:10.18546/LRE.13.3.05
- Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. *Studies in Educational Evaluation*, 53(C), 69-76. doi:10.1016/j.stueduc.2017.03.003
- Friesen, N. (2012). Experiential evidence: I, we, you. In N. Friesen, C. Henriksson, & T. Saevi (Eds.), *Hermeneutic phenomenology in education: Methods and practice* (pp. 39-54). Rotterdam, The Netherlands: Sense Publishers.
- Giacomini, M. (2013). Theory matters in qualitative health research. In I. Bourgeault, R. Dingwell, & R. de Vries (Eds.), *The SAGE handbook of qualitative methods in health research* (pp. 125-156). London: SAGE Publications Ltd.
- Giorgi, A. (1997). The theory, practice, and evaluation of the phenomenological method as a qualitative research procedure. *Journal of Phenomenological Psychology*, 28(2), 235-260. doi:10.1163/156916297X00103

-
-
- Goulding, C. (2005). Grounded theory, ethnography and phenomenology. *European Journal of Marketing*, 39(3/4), 294-308.
doi:10.1108/03090560510581782
- Gravett, K., Kinchin, I. M., & Winstone, N. E. (2019). 'More than customers': Conceptions of students as partners held by students, staff, and institutional leaders. *Studies in Higher Education*.
doi:10.1080/03075079.2019.1623769
- Greenberg, K. (2015). Rubric use in formative assessment: A detailed behavioral rubric helps students improve their scientific writing skills. *Teaching of Psychology*, 42(3), 211-217.
doi:10.1177/0098628315587618
- Grix, J. (2010). *The foundations of research* (2nd ed.). Basingstoke, New York: Palgrave Macmillan.
- Guba, E. G., & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 191-215). Thousand Oaks, California: SAGE Publications Inc.
- Hammersley, M. (2012). Transcription of speech. In S. Delamont (Ed.), *Handbook of qualitative research in education* (pp. 439-445). Glos, U.K.: Edward Elgar Publishing Limited.
- Harvey, M., & Kosman, B. (2013). A model for higher education policy review: The case study of an assessment policy. *Journal of Higher Education Policy and Management*, 36(1), 1-11.
doi:10.1080/1360080X.2013.861051
- Hawe, E., & Dixon, H. (2017). Assessment for learning: A catalyst for student self-regulation. *Assessment & Evaluation in Higher Education*, 42(8), 1181-1192. doi:10.1080/02602938.2016.1236360
- Ho, R. (2010). Assessment of Chinese students' experience with foreign faculty: A case study from a Chinese university. *Journal of Teaching in International Business*, 21(3), 156-177.
doi:10.1080/08975930.2010.504461
- Jessop, T. (2017). Inspiring transformation through TESTA's programme approach. In D. Carless, S. M. Bridges, C. K. Chan, & R. Glofchesky

(Eds.), *Scaling up Assessment for Learning in higher education* (pp. 49-64). Singapore: Springer

- Jones, G. A. (2014). Building and strengthening policy research capacity: Key issues in Canadian higher education. *Studies in Higher Education*, 39(8), 1-11. doi:10.1080/03075079.2014.949543
- Jones, L., Allen, B., Dunn, P., & Brooker, L. (2017). Demystifying the rubric: A five-step pedagogy to improve student understanding and utilisation of marking criteria. *Higher Education Research and Development*, 36(1), 129-142. doi:10.1080/07294360.2016.1177000
- Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, 39(7), 840-852. doi:10.1080/02602938.2013.875117
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130-144. doi:10.1016/j.edurev.2007.05.002
- Joughin, G., Dawson, P., & Boud, D. (2017). Improving assessment tasks through addressing our unconscious limits to change. *Assessment and Evaluation in Higher Education*, 42(8), 1221-1232.
- Kamberelis, G., & Dimitriadis, G. (2005). Focus groups: Strategic articulations of pedagogy, politics, and inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 887-907). Thousand Oaks, California: SAGE Publications Inc.
- Kennedy, K., Chan, J. K., Fok, P. K., & Yu, W. M. (2008). Forms of assessment and their potential for enhancing learning: conceptual and cultural issues. *Educational Research for Policy and Practice*, 7(3), 197-207. doi:10.1007/s10671-008-9052-3
- Kickert, R., Stegers-Jager, K. M., Meeuwisse, M., Prinzie, P., & Arends, L. R. (2018). The role of the assessment policy in the relation between learning and performance. *Medical Education*, 52(3), 324-335. doi:10.1111/medu.13487
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-

based teaching. *Educational Psychologist*, 41(2), 75-86.
doi:10.1207/s15326985ep4102_1

- Knight, P. (2007). Grading, classifying and future learning. In D. Boud & N. Falchikv (Eds.), *Rethinking assessment in higher education: Learning for the longer term* (pp. 72-86). New York, NY: Routledge.
- Lambrechts, W. (2015). The contribution of sustainability assessment to policy development in higher education. *Assessment & Evaluation in Higher Education*, 40(6), 801-816. doi:10.1080/02602938.2015.1040719
- Macdonald, R., & Joughin, G. (2009). Changing assessment in higher education: A model in support of institution-wide improvement. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education* (pp. 193-214). London: Springer.
- Mack, L. (2010). The philosophical underpinnings of educational research. *Polyglossia*, 19, 5-11.
- Mauri, T., Colomina, R., & de Gispert, I. (2014). Transforming collaborative writing tasks into opportunities for learning: Educational assistance and the use of rubrics in Higher Education. *Culture and Education*, 26(2), 298-348. doi:10.1080/11356405.2014.935111
- McCullough, C. A., & Jones, E. (2015). Creating a culture of faculty participation in assessment: Factors that promote and impede satisfaction. *Journal of Assessment and Institutional Effectiveness*, 4(1), 85-101. doi:10.5325/jasseinsteffe.4.1.0085
- McDowell, L., & Sambell, K. (2014). Assessment for learning environments: A student-centred perspective. In *Advances and Innovations in University Assessment and Feedback* (pp. 56-72): Edinburgh University Press.
- Medland, E. (2015). Examining the assessment literacy of external examiners. *London Review of Education*, 13(3), 21-33. doi:10.18546/LRE.13.3.04
- Medland, E. (2018). 'I'm an assessment illiterate': Towards a shared discourse of assessment literacy for external examiners. *Assessment & Evaluation in Higher Education*, 1-16.
doi:10.1080/02602938.2018.1523363

-
-
- Menéndez-Varela, J., & Gregori-Giralt, E. (2016). The contribution of rubrics to the validity of performance assessment: A study of the conservation-restoration and design undergraduate degrees. *Assessment & Evaluation in Higher Education*, 41(2), 228-244.
doi:10.1080/02602938.2014.998169
- Mercer, J. (2007). The challenges of insider research in educational institutions: Wielding a double-edged sword and resolving delicate dilemmas. *Oxford Review of Education*, 33(1), 1-17.
- Meyer, L. H., Davidson, S., McKenzie, L., Rees, M., Anderson, H., Fletcher, R., & Johnston, P. M. (2010). An investigation of tertiary assessment policy and practice: Alignment and contradictions. *Higher Education Quarterly*, 64, 331-350.
- Miedijensky, S., & Tal, T. (2016). Reflection and assessment for learning in science enrichment courses for the gifted. *Studies in Educational Evaluation*, 50, 1-13.
doi:http://dx.doi.org/10.1016/j.stueduc.2016.05.001
- Murphy, E., Dingwall, R., Greatbatch, D., Parker, S., & Watson, P. (1998). Qualitative research methods in health technology assessment: A review of the literature. *Health Technology Assessment*, 2(16).
- Norton, L. (2004). Using assessment criteria as learning criteria: A case study in psychology. *Assessment & Evaluation in Higher Education*, 29(6), 687-702. doi:10.1080/0260293042000227236
- Norton, L., Norton, B., & Shannon, L. (2013). Revitalising assessment design: What is holding new lecturers back? *Higher Education: The International Journal of Higher Education and Educational Planning*, 66(2), 233-251. doi:10.1007/s10734-012-9601-9
- O'Neill, M., & Chapman, A. (2015). Globalisation, internationalisation and English language: Studies of education in Singapore, Malaysia and Australia. *Education Research and Perspectives*, 42, 1-24.
- Olteanu, C. (2017). Reflection-for-action and the choice or design of examples in the teaching of mathematics. *Mathematics Education Research Journal*, 29(3), 349. doi:10.1007/s13394-017-0211-9

-
-
- Panadero, E., Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation*, 39(3), 125-132. doi:10.1016/j.stueduc.2013.04.001
- Panadero, E., Andrade, H., & Brookhart, S. (2018). Fusing self-regulated learning and formative assessment: A roadmap of where we are, how we got here, and where we are going. *Australian Educational Researcher*, 45(1), 13-31. doi:10.1007/s13384-018-0258-y
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9(1), 129-144. doi:10.1016/j.edurev.2013.01.002
- Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, 21(2), 1-16. doi:10.1080/0969594X.2013.877872
- Panadero, E., Tapia, J. A., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences*, 22(6), 806-813. doi:10.1016/j.lindif.2012.04.007
- Pham, T. T. (2011). Issues to consider when implementing student-centred learning practices at Asian higher education institutions. *Journal of Higher Education Policy and Management*, 33(5), 519-528. doi:10.1080/1360080X.2011.605226
- Pham, T. T., & Renshaw, P. (2014). Formative assessment in Confucian heritage culture classrooms: Activity theory analysis of tensions, contradictions and hybrid practices. *Assessment & Evaluation in Higher Education*, 40(1), 1-15. doi:10.1080/02602938.2014.886325
- Popham, W. (1997). What's wrong--and what's right--with rubrics. *Educational Leadership*, 55(2), 72-75.
- Poth, C., Riedel, A., & Luth, R. (2015). Framing student perspectives into the higher education institutional review policy process. *Canadian Journal of Higher Education*, 45(4), 361-382.

-
-
- Price, M. (2012). *Assessment literacy: The foundation for improving student learning*. Oxford: The Oxford Centre for Staff and Learning Development.
- Price, M., Carroll, J., O'donovan, B., & Rust, C. (2011). If I was going there I wouldn't start from here: A critical commentary on current assessment practice. *Assessment & Evaluation in Higher Education*, 36(4), 479-492. doi:10.1080/02602930903512883
- Raaper, R. (2016). Academic perceptions of higher education assessment processes in neoliberal academia. *Critical Studies in Education*, 57(2), 175-190. doi:10.1080/17508487.2015.1019901
- Ratnam-Lim, C. T., & Tan, K. H. (2015). Large-scale implementation of formative assessment practices in an examination-oriented culture. *Assessment in Education: Principles, Policy & Practice*, 22(1), 61-78. doi:10.1080/0969594X.2014.1001319
- Reddy, Y., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4), 435-448. doi:10.1080/02602930902862859
- Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research & Evaluation*, 15(8).
- Rezaei, A. R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, 15(1), 18-39. doi:https://doi.org/10.1016/j.asw.2010.01.003
- Rezvani Kalajahi, S. A., & Abdullah, A. N. (2016). Assessing assessment literacy and practices among lecturers. *Pedagogika*, 124(4), 232-248. doi:10.15823/p.2016.65
- Rhind, S. M., & Paterson, J. (2015). Assessment literacy: Definition, implementation, and implications. *Journal of Veterinary Medical Education*, 42(1), 28-35. doi:10.3138/jvme.0714-067R1
- Robinson, J. (2012). Using focus groups. In S. Delamont (Ed.), *Handbook of qualitative research in education* (pp. 391-404). Glos, U.K.: Edward Elgar Publishing Limited.

-
-
- Rodgers, C. (2002). Defining reflection: Another look at Joh Dewey and reflective thinking. *Teacher College Record*, 104(842-866).
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), 77-84. doi:10.1080/0969595980050104
- Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30(2), 175-194. doi:10.1080/0260293042000264262
- Sagarra, M., Mar-Molinero, C., & Rodríguez-Regordosa, H. (2015). Evaluating the success of educational policy in Mexican higher education. *The International Journal of Higher Education Research*, 69(3), 449-469. doi:10.1007/s10734-014-9785-2
- Saldana, J. (2011). *Fundamentals of qualitative research: Understanding qualitative research*. New York, New York: Oxford University Press.
- Sambamurthy, M. N., & Cox, M. F. (2016, June 26-29). *Developing a behavior-based rubric for assessing professional skills during hiring*. Paper presented at the American Society for Engineering Education's 123rd Annual Conference, New Orleans, LA.
- Sambell, K., Brown, S., & Graham, L. (2017). *Professionalism in practice: Key directions in higher education learning, teaching and assessment*. Switzerland: Palgrave Macmillan.
- Sambell, K., McDowell, L., & Montgomery, C. (2013). *Assessment for learning in higher education*. New York, NY: Routledge.
- Saunders, M., & Sin, C. (2015). Middle managers' experience of policy implementation and mediation in the context of the Scottish quality enhancement framework. *Assessment & Evaluation in Higher Education*, 40(1), 135-150. doi:10.1080/02602938.2014.892056
- Sayigh, E. A. (2006). Refining lecturers' assessment practices through formal professional development at Rhodes University. *South African Journal of Higher Education*, 20(1), 157-169. doi:10.4314/sajhe.v20i1.25564

-
-
- Schoeb, V., & Chong, D. (2019). Students' learning preferences and experience in a globalised world: Opportunity to optimise internationalisation in physiotherapy education. *OpenPhysio*.
- Schwartz, M. (n.d.). Flexible learning. In. Learning & Teaching Office, Ryerson University.
- Shay, S. (2005). The assessment of complex tasks: A double reading. *Studies in Higher Education*, 30(6), 663-679. doi:10.1080/03075070500339988
- Sindelar, P., & Rosenberg, M. (2000). Serving too many masters: The proliferation of ill-conceived and contradictory policies and practices in teacher education. *Journal of Teacher Education*, 51(3), 188-193.
- Smetherham, D. (1978). Insider research. *British Educational Research Journal*, 4(2), 97-102. doi:10.1080/0141192780040209
- Smith, C. D., Worsfold, K., Davies, L., Fisher, R., & McPhail, R. (2013). Assessment literacy and student learning: The case for explicitly developing students 'assessment literacy'. *Assessment & Evaluation in Higher Education*, 38(1), 44-60. doi:10.1080/02602938.2011.598636
- Starks, H., & Brown Trinidad, S. (2007). Choose your method: A comparison of phenomenology, discourse analysis, and grounded theory. *Qualitative Health Research*, 17(10), 1372-1380. doi:10.1177/1049732307307031
- Tagg, J. (2012). Why does the faculty resist change? *Change: The Magazine of Higher Learning*, 44(1), 6-15. doi:10.1080/00091383.2012.635987
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *The International Journal of Higher Education Research*, 76(3), 467-481. doi:10.1007/s10734-017-0220-3
- Tanner, J. (2017). *Enabling collaboration during written examinations: A pilot study in a biomedical sciences examination*. Paper presented at the 9th AMEA Symposium cum Frontiers in Medical and Health Science Education 2017, Hong Kong.
<http://www.bimhse.hku.hk/frontiers2017/op/op20.pdf>

-
-
- Tay, H. Y. (2015). Setting formative assessments in real-world contexts to facilitate self-regulated learning. *Educational Research for Policy and Practice*, 14(2), 169-187.
- Terhart, E. (2013). Teacher resistance against school reform: Reflecting an inconvenient truth. *School Leadership & Management*, 33(5), 486-500. doi:10.1080/13632434.2013.793494
- Tight, M. (2012). *Researching higher education* (2nd ed.). Berkshire, England: McGraw Hill Society for Research into Higher Education & Open University Press.
- Tijs, R., David, B., Vaes, P., & Kerckhofs, E. (2012). Developing an assessment and evaluation policy in physiotherapy education. *Procedia - Social and Behavioral Sciences*, 69(C), 545-548. doi:10.1016/j.sbspro.2012.11.444
- Tracy, S. J. (2013). *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact* (1st ed.). U.K. : Wiley-Blackwell.
- Trowler, P. (2016). *Doctoral research into higher education: Making theory work*. Great Britain: CreateSpace Independent Publishing Platform.
- Van Helvoort, J. (2010). A scoring rubric for performance assessment of information literacy in Dutch higher education. *Journal of Information Literacy*, 4(1), 22-39. doi:10.11645/4.1.1256
- Wang, Z., & Brown, G. T. (2014). Hong Kong tertiary students' conceptions of assessment of academic ability. *Higher Education Research and Development*, 33(5), 1063-1077. doi:10.1080/07294360.2014.890565
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58(C), 149-162. doi:10.1016/j.tate.2016.05.010
- Yilmaz, K. (2013). Comparison of quantitative and qualitative research traditions: Epistemological, theoretical, and methodological differences. *European Journal of Education*, 48(2), 311-325. doi:10.1111/ejed.12014

Zimmerman, B. J. (2000). Attaining self-regulation. A social cognitive perspective. . In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego: Academic Press.

Appendix 1 - The Rubric Policy Paper

Learning and Teaching Committee

Paper for : Information

Subject : Policy on the use of rubrics in major assessment tasks

1 Background

- 1.1 In 2005 the University adopted a criterion-referenced approach to assessment (CRA) in which students are graded according to pre-determined criteria and standards. To facilitate implementation, guidelines for implementing CRA were developed and disseminated to all staff via publications, web links and staff development activities.
- 1.2 While CRA is generally adopted across departments, there are variations in the rigour with which the principles of CRA are put into practice.
- 1.3 To ensure that the principles of CRA are consistently implemented with adequate rigour, it is proposed that the use of rubrics should be made compulsory for major assessment tasks at the subject level.

2 Use of rubrics in higher education

- 2.1 A rubric is a coherent set of criteria for students' work that includes descriptions of levels of performance quality of the criteria. Rubrics have a natural affinity with the CRA approach because they make explicit the criteria and standards against which assessment are made. They can help students come to grasp the ways experts in the field think, practice, express ideas and appraise each others' work. In a way, they help scaffold the development of students' own expertise in the field as a professional.
- 2.2 Rubrics are important tools for internal moderation of student assessment results. They reduce uncertainty and disagreement over grades by making the bases of judgement transparent. Communicating the criteria to students before the assessment and ensuring that all markers have shared understanding of the criteria and standards before marking commences are both important aspects of CRA.
- 2.3 Internationally, rubrics are being accorded much importance as a tool for benchmarking academic standards as it allows grades to be interpreted in terms of the level of competence that they represent. Increasingly professional accreditation bodies are requiring rubrics to be specified at the subject level. It would be a good practice to subject rubrics to external moderation even if it is not required by accreditation.

3 Policy on the use of rubrics ^[See notes]

- 3.1 Rubrics must be specified for all ‘major’ assessment items at the subject level, made available to students before the assessment, and used for grading the assessment. Departments have the flexibility to determine what is ‘major’. As a rule of thumb:
- For subjects without examinations, rubrics should be required for single assessment items with a weighting of 30% or above of the subject’s overall assessment.
 - For subjects with examinations, rubrics should be required for single assessment items with a weighting of 20% or above of the subject’s overall assessment.
- 3.2 There is no fixed format for rubrics. Any format (e.g., analytic, holistic) is acceptable as long as it clearly defines the main grades (A, B, C, D, Fail) in a way that is understandable to students and is adhered to by teachers in grading.
- 3.3 To ensure that the rubrics reflect a suitable level of academic standards, samples of the rubrics should be periodically reviewed by Departmental Academic Advisors, External Examiners and/or Overseas Academic Advisors, as part of the review process during Departmental Review and other periodic visits by these individuals where appropriate. This being a measure of external benchmarking is not a substitute for internal moderation of assessment processes and results by relevant departmental committees/panels/boards.

4 Notes on implementation

- 4.1 The deadlines for specifying rubrics for ‘major’ continuous assessments and examinations are the end of 2017/18 and 2018/19 respectively. External review of sample rubrics shall be evident in relevant reports by 2018/19.
- 4.2 Subjects that focus primarily on basic mathematics or basic science concepts/topics for which numerical calculation as the only valid and viable means to demonstrate understanding may be given one more year to make the transition, i.e., the deadlines for them will be the end of 2018/19 and 2019/20.
- 4.3 The DLTC of the subject-offering department may decide on whether a subject should be given the additional year based on the criterion given in 4.2. The DLTC should monitor the progress of implementation and give a brief interim report to the LTC near the end of each academic year until the transition is completed, or as otherwise requested by the LTC.

5 Implementation support

- 5.1 For guidelines on developing clear criteria and standards for rubrics, teachers may refer to the “Guidelines for implementation of criterion-referenced assessment” developed by the LTC [REDACTED]

5.2 The Educational Development Centre will provide continuous support to academic staff/departments in the introduction of rubrics via workshops and consultations.

NOTES

The proposal was endorsed at the 59th LTC meeting. The original paper did not mention the format of rubrics within the body of the policy proposal, but the expectation as expressed in 3.2 was nevertheless discussed and agreed upon at the 58th meeting, and was mentioned as a note when the policy was announced on 1 April 2016. The current version merely incorporates this into the body of the policy to make it more complete. This version also includes a correction to 3.1, where the expression ‘over 30%’ was replaced with ‘30% or above’ to give a more accurate representation of guidelines as endorsed by the Committee.

August 2016¹

¹ A subsequent editorial change (removal of a reference to a missing appendix in 2.1) was made in October 2016.

Appendix 2 - Interview Questions

Students

- What does assessment mean to you?
 - What does it comprise?
 - What is its function?
 - Why do you think assessment is needed?
- How has your experience with assessment in university been?
 - Types
 - Do they match with your expectation of functions?
 - Do they serve their purpose?
- Do you think assessment can be used for learning?
 - Can you think of an example of assessment that has helped you learn?
 - Can you think of an example of assessment that's detrimental to your learning?
- What would it take to switch an assessment task to a learning task?
 - What needs to change?
 - What needs to be done to make these changes?
 - Is there a need to do so?
- What does rubrics mean to you?
 - What is it?
 - What is its function?
 - Do you think it is needed?
- Do you think rubrics can help with learning?
 - Can you think of an example of rubrics that has helped you learn? What does it look like?
 - Can you think of an example of rubrics that is detrimental to learning? What does it look like?
- Is it only rubric, or is there something else?

Educators

- What does assessment mean to you?
 - What does it comprise?
 - What is its function?
 - Why do you think assessment is needed?

- How has your experience with assessment in university been?
 - What assessment types have you used?
 - Why these assessments?
 - Do they match with your expectation of functions? E.g. assess the intended learning outcomes

- Do you think assessment can be used for learning?
 - Can you think of an example of assessment that has helped your students learn?
 - Can you think of an example of assessment that's detrimental to your students' learning?

- How do you interpret the rubric policy?
 - Something management needs to comply?
 - Do you know the purpose of it?
 - Do you think it can serve its purpose?

- Have you developed rubrics for your assessments? How was the experience?
 - Yes – development process, understanding of rubrics
 - No – why not? Barriers? Understanding of rubrics

- What does rubrics mean to you?
 - What is it? Is it only a mandated policy?
 - What is its function?
 - Do you think it is needed?
 - Do you think it can help with learning?

- What is the relationship between rubrics and assessment for learning?
 - Have you seen an example that you can share?

Management

- What does assessment mean to you as a manager?
 - What does it comprise?
 - What is its function?
 - Why do all subjects need assessments?

- How does management use assessment results?
 - For quality control? Reassurance?
 - Bench marking?
 - KPI? Policy needs?
 - Do you think the current assessment results truly reflect these?

- Do you think assessment can be used for learning?
 - Have you seen examples of assessment that has helped students learnt?
 - Have you seen examples of assessment that's detrimental to students' learning?

- How do you interpret the rubric policy?
 - Something management needs to comply?
 - Do you know the purpose of it?
 - Do you think it can serve its purpose?

- What does rubrics mean to you?
 - What is it? Is it only a mandated policy?
 - What is its function?
 - Do you think it is needed?

- What is the relationship between rubrics and assessment for learning?
 - Have you seen an example that you can share?