Indonesian junior secondary school teachers' conceptions of assessment: A mixed methods study

Astuti Azis

February 2014

A thesis submitted to the Victoria University of Wellington
in fulfilment of the requirement for the degree of
Doctor of Philosophy in Education

Abstract

Understanding teachers' conceptions of assessment is important because of their influence on teachers' assessment practices. Despite numerous quantitative studies on teachers' conceptions and practices of assessment, little research exists regarding the unique assessment environment of Indonesia. This study uses both quantitative and qualitative data to examine how Indonesian junior high school teachers understand assessment, how their conceptions of assessment relate to their assessment practices, and what factors contribute to their conceptions and practices of assessment.

This mixed methods study adopted a participant selection model in which quantitative data was analysed to select participants for the qualitative phase. A validated measure of teacher conceptions of assessment was adapted for use in the quantitative phase to explore teacher (N=107) conceptions of whether they thought assessment was for improvement, accountability or whether it was irrelevant. These three criteria were used to select twelve Indonesian teachers for semi-structured interviews and to contribute documents for analysis in the qualitative phase.

The Indonesian teachers believed that the purpose of assessment was to improve teaching and learning and also to demonstrate the accountability of students and school. They tended to disagree with the view that assessment is irrelevant. Further analysis of the data revealed that teachers' conceptions of assessment were conflicted. They were keen to use assessment practices to improve their classroom teaching, but felt that the state-wide examination policy requirements constrained their efforts. These distinctive and conflicting conceptions of assessment held by Indonesian teachers appeared to arise from the interplay of socio-ecological factors including culture and the Indonesian education system.

Conceptions of assessment are unique to every setting. These findings highlight that valid measurement of teacher conceptions is likely to require national and regional accommodations based on contextual factors. Furthermore, government, policy makers, and curriculum developers must work to build a strong synergy among themselves in order to share consistent goals with teachers. If cultural expectations of school assessment and government policy were aligned, Indonesian

teachers may be better able to resolve conflict between their beliefs and assessment practices.

Acknowledgements

Alhamdulillah Rabbil Alamien; this thesis is a dream come true, one that is accomplished with the *Ridho* of Allah SWT and the support of many generous people.

I am indebted to my supervisors, Dr. Margaret Gleeson and Professor Luanna H Meyer, for their wisdom, encouragement and continuous support. I cannot thank them enough for everything they have done for me to make this dream a reality.

I also extend my utmost gratitude to Dr. Carolyn Tait, Dr. Flaviu Hodis, and Dr. Lex McDonald for giving me a research assistant role within the School of Educational Psychology and Pedagogy (SEPP). The experience with SEPP contributed immeasurably to my academic and professional growth. I thank the Head of SEPP and staff, Dr. Vanessa Green, Amy de Boer and Sandy Fackney; who were always concerned to create an enabling work environment for me.

I am grateful to Dr. Deb Laurs for always supporting me during the brainstorming and discussion of my drafts, Dennis Dawson and Dr. Dalice Sim who kindly guided me in the quantitative realm, Laura Dimock for her assistance in hard times, and Louise Jennings many thanks for your proof reading.

I am extremely thankful to Dr. Syahrul Yasin Limpo, the Governor of South Sulawesi, for financial assistance through the Overseas Doctorate Scholarship Programme. I would also like to thank the Victoria University Scholarship Committee for funding the last years of my study. It would have been impossible for me to complete my PhD journey without their sponsorship.

My special thanks to the Faculty of Education VUW for the facilities, Elizabeth Morris for granting my conferences, Sheila Law, Alex Klein, and Nicola Atkinson who have been very supportive to me. I would also like to thank my colleagues and friends particularly those at 33 Campbell Street, Karori, for the friendship, laughter, food, encouragement and lunches shared. May Allah SWT bless their activities.

To my late father, my mother, brothers and sisters and my extended family who always believed in me and prayed for me, thank you very much for your love, trust,

and support. To my super wonderful husband and best friend, Muhammad Taufiq Badewi who has been the great guardian and the best father of our kids Hanum Muthia and Hanif Muthahhar while I am far away from home; you are the pillar of my strengths. To Hilmy Mu'tashim, my big boy who always listens to all my stories thanks for providing your waterproof shoulder for me to cry on. Mantang, the greatest nanny ever, I cherish your love and affection for my kids and family. To you I dedicate this thesis.

Finally, I would especially like to thank the teachers who participated in this study, without whom this thesis would not be completed. Thank you very much and may Allah SWT grant you reward, Amen.

Table of Contents

Abstract ... ii
Acknowledgements ... iv
Chapter One Introduction ... 1
Introduction ... 1
Statement of the problem ... 1
Rationale for the study ... 2
Focus of the study ... 3
Purposes of the study ... 3
Research questions ... 4
Key question: ... 4
Sub questions: ... 4
Significance of the study ... 4
Overview of education and assessment systems in Indonesia ... 5
The educational context ... 5
Background to the assessment process in Indonesia ... 6
Assessment system in the region of Gowa ... 8
Definition of key terms ... 10
Organisation of thesis chapters ... 11
Chapter Two Literature Review ... 13
Introduction ... 13
Conceptions in this study ... 13
The importance of studying conceptions ... 13
The ecological theory: framework of the study ... 14
The macrosystem and the exosystem ... 16
A culture of examination, and grading ... 16
The education system, policies and regulations ... 19
Mesosystem ... 21
Microsystem ... 22
Student-teacher relationship and curriculum assignment level ... 23
Views of learning ... 24
Teachers' conceptions of assessment ... 26
Assessment for learning (AfL) ... 29
Assessment of Learning (AoL) ... 31
Summary ... 33
Chapter Three Methodology ... 34
Introduction ... 34
Research approach ... 34
Justification for using mixed methods ... 35
Sampling strategy ... 35
Limitation of mixed methods inquiry ... 36
Phase One: Quantitative ... 37
Mixed Methods Sequential Explanatory Design ... 37
Participants ... 38
Characteristics of the participants ... 40

Instrument 41
The Hong Kong model of TCoA survey 41
Data collection 44
Pilot study 44
Administration of questionnaires 44
Data Analysis 45
Phase two: Qualitative 46
The case study approach 46
Participants 48
The semi-structured interview 48
Documents 50
Data collection process 50
Interview process 50
Qualitative Analysis 51
Interview data 51
Document analysis 52
Validity and reliability of the mixed methods design 53
Quantitative issues 53
Reliability 53
Validity 55
Objectivity 58
Qualitative issues 58
Credibility 58
Transferability 58
Dependability 59
Confirmability 59
Ethics 59
Summary 60
Chapter Four Findings 61
Introduction 61
Quantitative findings 61
Conceptions about assessment to improve teaching and learning 61
Conceptions about the irrelevancy of assessment 63
Conceptions about accountability in assessment 64
Overall result of teachers' conceptions of assessment 66
Links between quantitative and qualitative data 67
Qualitative findings 68
Data from assessment documents 70
Common regional assessment documents 71
Students report book 72
Semester test 72
Answer sheet for the semester test 72
Common school documents 73
Education calendar 73
Annual programme 73

Semester programme 73
Syllabi 74
Learning journal 74
Mark book 75
Classroom documents 75
Teacher test 75
Mid-term test 75
List of grades for final report 76
Student documents 76
Summary of document analysis 76
Data from interviews 78
Improvement group (IM) 78
Conceptions of the purposes of assessment 78
Internal assessment 79
Openness to change 79
Developing values 82
Authentic formative assessment practices 84
Grading to show achievement 87
External assessment 88
Conflicting conceptions 88
The credibility of external assessments 90
IM case summary 94
Irrelevance group (IR) 94
Conceptions of the purposes of assessment 94
Internal assessment 95
Modifying teaching 95
Developing values 97
Authentic assessment practices 98
Grading for accountability 101
External assessment 102
Conflicting perceptions 102
Teachers' autonomy 103
The credibility of summative assessment 104
IR case conclusion 107
Accountability group (AC) 107
Conceptions of purposes of assessment 107
Internal assessment 108
Openness to change 108
Developing values 110
Mixed understanding of assessment 112
Grading practices 115
External assessment 116

Conflicting conceptions 116
Teachers' autonomy 118
AC case summary 121
Chapter Five Discussion of findings and conclusion 122
Introduction 122
The complexity of teachers' conceptions of assessment 122
Cultural factors 125
Contextual factors 126
Personal factors 129
Indonesian teachers' perceptions and practices of assessment 130
Classroom assessment with external accountability focus 131
Preferences for assessment practice: Compliance and defiance 134
Credibility of assessment 136
Limitations 138
Instrument 138
Sampling 139
Implications of this study 139
Direction for future research 140
Concluding Statement 141
References 143
Appendices 152
Appendix A: Key competencies for English subjects 152
Appendix B: TCoA Survey 155
Appendix C: Information Sheet 160
Appendix D: Consent Form 164
Appendix E: Interview Guide 166
Appendix F: Assessment Documents 167
Appendix G: Ethical Approval 201
Appendix H: Permission letter from the Province 202
Appendix I: Permission letter from the Regent of Gowa 203
Appendix J: Permission letter from the Education Department 204
Appendix K: Letter to PD Leader 205
Appendix L: Letter to teacher (phase 1) 206
Appendix M: Letter to teacher (phase 2) 207

List of Tables

Table		Page
1.	Conceptions based on purposes……	28
2.	Participants in quantitative data collection……	39
3.	Basic demographic information of participants: gender, age, education level and teaching experience……..	40
4.	Basic demographic information of participants: curriculum level assignment, certification status and assessment training attended….	40
5.	Dimensions of the Teachers Conceptions of Assessment (TCoA) survey…….	43
6.	Details of questionnaire administration…….	45
7.	Details of interview participants……..	48
8.	Samples of the interview questions……...	50
9.	Details of interview process……	51
10.	Reliability of the Hong Kong TCoA questionnaire……...	54
11.	Factor rotation with 7 components…….	56
12.	Factor rotation with 3 components…….	57
13.	Improvement conception…….	62
14.	Irrelevance conception…….	63
15.	Accountability in assessment……..	64
16.	Agreement level of teachers for components in TCoA, N=107……	66
17.	Demographic information of case study participants……..	70
18.	Overview of participants' assessment documents……..	77

List of Figures

Figure		Page
1.	Model of Indonesia educational standards (adapted from (MoNE, 2008)……..	6
2.	Model of assessment system in Indonesia (Adapted from (BSNP, 2006)……..	8
3.	Model of assessment system in the Region of Gowa…….	9
4.	Procedure of KKM establishment……..	11
5.	The relationship between personal and contextual factors in shaping teachers' understanding of assessment……	15
6.	Continuum of assessment……..	29
7.	Visual diagram of sequential explanatory design: participant selection model (Creswell & Plano Clark, 2007, p. 73)……..	38
8.	The embedded units within the three clusters (adapted from Thomas, 2011; Yin, 2009……...	47
9.	Strategies for analysis and interpretation of documents (adapted from McMillan & Schumacher, 2010, p.362)……	53
10.	Mean score of improvement conception-related items……..	63
11.	Mean score of irrelevance conceptions-related items……	64
12.	Mean score of accountability conception-related items…….	66
13.	Overall results of teachers' conceptions of assessment……	67
14.	Questions to be followed-up……...	68
15.	Profile of teachers who participated in the case study……...	69
16.	Participant assessment document…….	71
17.	Profile of teachers in the improvement conception case……	79
18.	Profile of teachers in the irrelevance conception case……...	95
19.	Profile of teachers in the accountability conception case……..	108
20.	The interrelationship of Indonesian teachers' conceptions and practices of assessment with socio-ecological factors……...	124
21.	Indonesian teachers' perceptions and practices of assessment…..	131

Chapter One

Introduction

Introduction

This thesis reports the conceptions and practices of assessment among junior secondary school English teachers from one region in south Sulawesi province, Indonesia. This is a mixed methods study with quantitative and qualitative phases conducted to examine the extent to which assessment was valued by these teachers and to discover how their conceptions of assessment were operationalised in the classroom.

Statement of the problem

Assessment is considered to play a critical role in education for both policy makers and practitioners. Assessment may be conducted for purposes of accountability which includes establishing how well students have learned, or to inform the design of instruction (how to improve) in educational contexts (Danielson, 2008). These two purposes sometimes support one another, and sometimes compete or conflict with one another, which indicates that assessment is a complex process (Earl, 2003).

For decades, the practice of summative assessment (SA) for accountability purposes has dominated classroom assessment activities. However, recent trends in teachers' classroom practices have reflected a paradigm shift (Hargreaves, Earl, & Schmidt, 2002) towards assessment as a practice to improve learning (Guskey, 2003). This paradigm change is most visible in countries where low-stakes examination policy is implemented. In other contexts where the practices of SA have been widely accepted as part of history and culture (I.-C. Choi, 2008), teachers and the public tend to maintain and value the familiar SA processes and practices (Earl, 2003). This phenomenon suggests that there is a discrepancy in assessment policy and practice particularly between high-stakes and low-stakes assessment environments. Both policies appear to work well in particular contexts indicating that assessment may be connected with cultural beliefs and practices. This signals the importance of investigation into teachers' conceptions of assessment in different contexts and cultures.

The literature on conceptions of assessment has discussed teachers' conceptions, assessment strategies, values of assessment practices, and teachers' views of learning. These studies have mostly been quantitative investigations that describe the uses of assessment, and the types of conceptions held by teachers, but do not incorporate the voices of participants to describe the meaning behind their conceptions.

One issue that arises then is that quantitative results may not fully describe and explain teachers' conceptions of assessment. Furthermore, there is little research about assessment conducted in Asian contexts where examinations are embedded in the culture of teaching and learning. This study illustrates how assessment works in an Indonesian context and shows how teachers' conceptions affect their practices through giving participants the opportunity to articulate their perceptions using qualitative methods.

Rationale for the study

I am interested in investigating teachers' conceptions of assessment based on my background as an English teacher at junior secondary school level for almost twelve years. During that time, there was little professional development on how to assess against the compulsory national requirements. My own view is that assessment of students' work is part of my teaching job and allows me to evaluate my teaching and my students' learning. I also believe that the assessment process functions to control the quality of teachers and schools (Harianti, 2005). I believe that assessment stimulates students to study harder as well as encouraging teachers to perform a better job. That is why I believe that an assessment system which focuses on examinations works well in Indonesia. There is limited research on assessment conceptions in Indonesia, so it interested me to discover what conceptions Indonesian teachers' may have about assessment and how they interpret those conceptions in practice.

In conducting this research I drew upon networks within the English teaching community of my home province in order to solicit participants. Any conflict of interest in working with teachers who may be known to me was addressed by inviting all teachers to participate after receiving permission to conduct the study from the Department of Education of Gowa Region. Moreover, participants'

confidentiality was protected in both phases of the study. In the information sheet I explained that they could withdraw from the research without providing any reasons.

Focus of the study

This study focuses on junior secondary school English teachers from the Gowa region, South Sulawesi, Indonesia. The reason for choosing participants who teach at this level is that English is one of the subjects tested in the national junior secondary school examination. Furthermore, the changes in curriculum appeared to impact the teaching and assessment for English subject. The current curriculum requires teachers to give a composite English score for each student that includes English skills, attendance, character and personality. It is quite different from the previous competency based curriculum where teachers reported their students' score against the four micro skills (reading, listening, writing, and speaking). In deciding the final score teachers consider both students' knowledge and behaviour (McMillan, Myran, & Workman, 2002). One composite English language assessment allows some latitude for English teachers to add a subjective assessment. They may be influenced by a student's previous academic work or opinions about how the student speaks, particularly when borderline marks are being decided (Noor, et al, 2010). This study was conducted in the Gowa Region as the context is typical of South Sulawesi in terms of population; 594,423 people and size; 1.883,32 square kilometres (Pemerintah Kabupaten Gowa, 2011).

Purposes of the study

The purpose of this sequential explanatory mixed methods study is to explore English teachers' conceptions of assessment by obtaining statistical, quantitative results from a sample, then following up with a few individuals to probe or explain those results in more depth. In the first quantitative phase of the study, an internationally validated questionnaire was used to address teachers' conceptions of assessment. In the second phase, qualitative semi-structured interviews and document analysis were utilised to construct cases to investigate how and why teachers believed in particular assessment conceptions.

Research questions

The following questions guide the study:

Key question:

How do Indonesian junior secondary school teachers understand the role of assessment and how do they perceive that assessment impacts upon student learning?

Sub questions:

1. In what ways do teachers of English in the region of Gowa, South Sulawesi Indonesia explain the value of assessment in student learning?
2. What factors do teachers perceive contribute to their understanding and use of assessment in student learning?
3. In what ways do teachers perceive their practice reflects their understanding of the role of assessment to promote student learning?

Significance of the study

The issue of how teachers' conceive of assessment has not fully been studied (Brown, 2008), particularly in an eastern context and at secondary school level. My study contributes to the literature by extending Brown's research on teachers' conceptions of assessment. It investigates Indonesian junior secondary school English teachers' conceptions of assessment in student learning. In addition, the results of this study are valuable because of its unique mixed methods research design. This methodological integration reveals a deeper insight into teachers' conceptions of assessment than previous quantitative designs. Firstly, it identifies the conceptions held by teachers and then it explores why and how participants believe in a particular conception. Methodologically, this study adds to mixed method research by following the procedure of the sequential explanatory design, which connects the quantitative and qualitative data, as well as integrating the results of the two sequential phases of the study. It is also anticipated that the results of this study will contribute to a comprehensive understanding of Indonesian junior secondary school teachers' conceptions of assessment, factors contributing to their conceptions and how teachers perceive these conceptions influence their assessment practices.

Overview of education and assessment systems in Indonesia

The educational context

Currently, Indonesia follows an education system which consists of six years primary (Years 1-6), three years junior secondary (Years 7-9), three years senior secondary (Years 10-12) and four years of tertiary education. National Examinations are conducted in the last year of each level except for tertiary education. Primary and junior secondary schools are categorised as basic education and are compulsory. This policy was declared in 1994 and described as 'nine-year compulsory basic education' (Ministry of National Education, 2005a). The junior secondary school level, which is an intermediate or middle level, is attained on completion of primary school. At the end of Year 9, it is compulsory for students to sit their second National Examination. In other words, students have to sit for two National Examinations up to this point, the first examination conducted in Year 6 and the second one in Year 9.

The National Examination is designed and conducted by the Board of National Standards for Education. Four courses are tested in the examination: Bahasa Indonesia, English, Mathematics and Science. The National Examination is used to evaluate the quality of each school, the region, and the province against the national standards (Ministry of National Education, 2005a). The Ministry of National Education collects and ranks the results of the examination then uses these rankings to map school quality and to provide financial aid for low-achieving schools/provinces.

In 2001, a new regulation was implemented in the country regarding the sharing of power. This policy was granted under Law no. 22/1999 on regional government and Law no. 25/199 on the fiscal balance between the central government and the regions. Previously, the Indonesian system of government relied heavily on central authorities. The new regulation decreed all sectors of government to be decentralised, including the educational sector. This meant that local governments were granted the authority to hire, fire, pay and train civil servants in their area. Educational decentralisation introduced more democratic authority structures and incorporated more people in decision-making processes including at the regional level. Through school based management, teachers were expected to actively design the school level curriculum and experiment with instructional strategies. In

other words the decentralisation system promoted teachers' autonomy and supported them to be more active agents of change in the community.

Background to the assessment process in Indonesia

The quality of education in Indonesia is obtained through the national education benchmark. The national education benchmark determines the minimum acceptable criteria within the education system in the country, based on eight standards. There are standards for content, process, graduate competence, teachers and staff, facilities, management, finance, and assessment. The national standards are developed, monitored and evaluated by the Board of National Standards for Education (BSNP) an independent and professional board working for the Ministry of National Education. The board has responsibilities for managing the National Examination, assessing all textbooks to be used at school and recommending and controlling the quality of education. Standards developed by the board become a requirement for all schools in the country.

The Indonesian education system and its educational assessment model ensure quality assurance (QA), quality control (QC) and quality improvement (QI). While QA is derived through the eight standards, QC is conducted across three levels of assessment. Both requirements monitor the quality of learning outcomes (or the national standards) to ensure QI (Ministry of National Education, 2008). Figure 1 presents the model of Indonesian education standards.

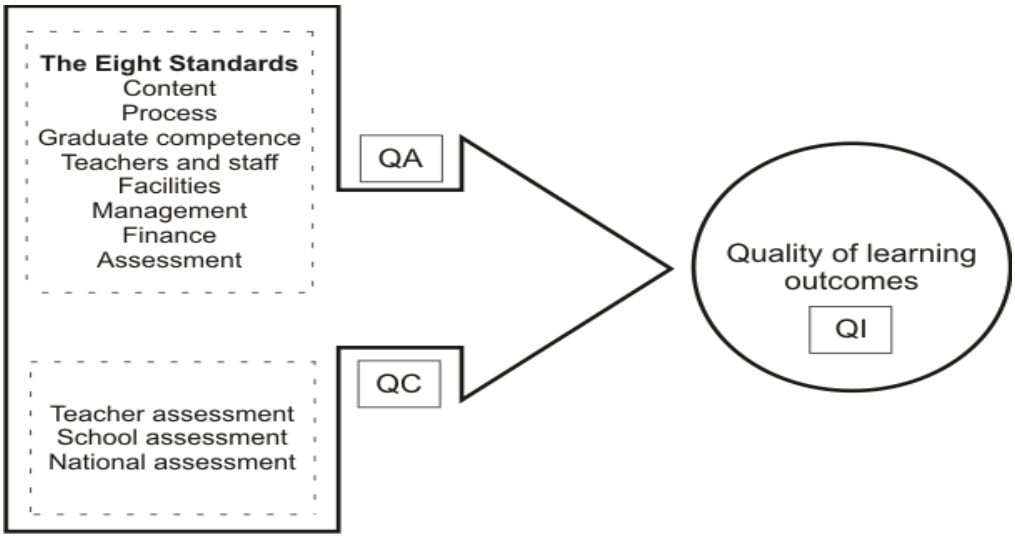


Figure 1: Model of Indonesia educational standards (adapted from (Ministry of National Education, 2008)

The standard of content and the standard of graduate competence are used to develop the school level curriculum (BSNP, 2006). The school level curriculum consists of several standard competencies and basic competencies which were developed based on grade level. Although schools are given the authority to develop their own curriculum, all standard competencies and key competencies must be included in teachers' syllabi. Each syllabus contains about twelve standard competencies and twenty-six basic competencies that teachers need to assess (see appendix A).

To assure quality control, the government regulated assessment through Act no. 20/2007 that requires educational assessment to be conducted at three levels: at teacher, school, and national levels. This ensures that the assessment system in Indonesia involves both internal and external agencies (Harlen, 2007). Classroom assessment is conducted continuously with three basic functions: to monitor the process of learning, to monitor students' progress and to improve learning. These purposes reflect the conceptions of improvement, school accountability and student accountability (Brown, 2008; Webb, 1992). Classroom tests, mid-semester tests, semester tests and class promotion tests are forms of teacher assessment recommended in the policy. All tests are conducted "to check up on what students have learned from a series of lessons over a period of time" (Harlen, 2007, p. 53). In other words, tests are conducted to measure and evaluate students' mastery against the curriculum every two to three weeks, as well as in the middle and at the end of each semester.

School assessment is designed and conducted by each school as a prerequisite for students leaving school. These tests cover all subjects which are not tested in the National Examination at the last grade of junior secondary school level. This school examination is conducted either before or after the National Examination in Year 9 of junior secondary school.

The highest level of assessment adopted in Indonesian schools is the National Examination. Like the school assessment, the National Examination is conducted at Year 9 of junior secondary school level. The examination result is also helpful for the selection of students for senior secondary. The following figure illustrates the assessment system in Indonesia.

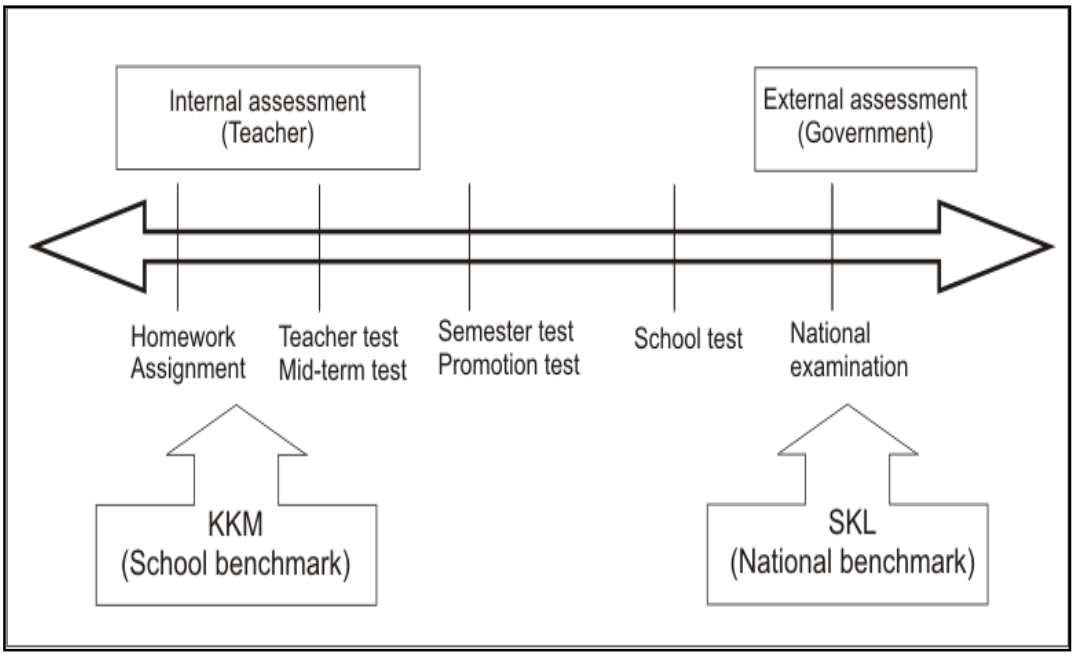


Figure 2: Model of assessment system in Indonesia (Adapted from (BSNP, 2006)

The quality of education in the country is ultimately measured through assessment against the national standard. The school level curriculum requires all students to achieve the minimum criteria set by the school. These criteria are called the minimum criteria of mastery learning (KKM); it is the standard for competencies achievement.

Assessment system in the region of Gowa

Every region in every province in Indonesia is granted local autonomy through the decentralised governmental system, however, all regulations set at the regional level should align to the national regulations. Gowa was the first region in the province to implement subsidised education for students at elementary, intermediate, and high school levels. The programme was implemented to provide free education to all school-age children, with no requirement for tuition or activity fees.

The region of Gowa has made a serious commitment to the national education system and to achieving the national education standards. This is visible in the regional government's decision to play the role of external assessor. The regional government took over the semester test and the class promotion test which were

previously managed by teachers. In managing both tests, the government of Gowa uses similar assessment formats as those used in the National Examination. This might imply that the government wishes to familiarise students with the National Examination. The following diagram illustrates the regional policy in Gowa regarding the assessment process.

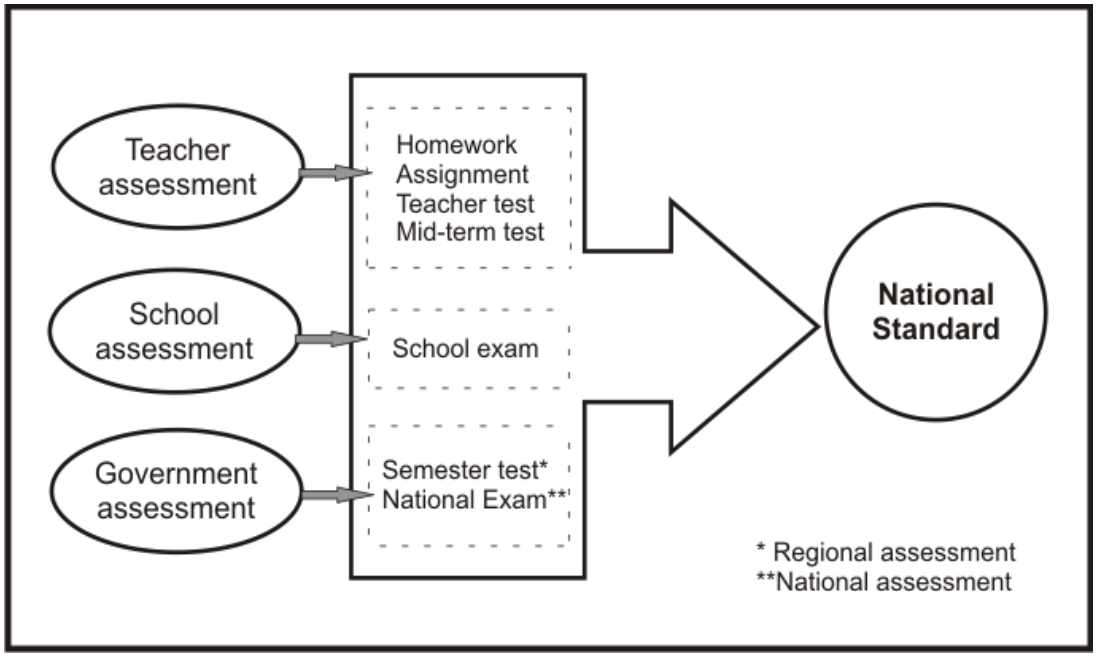


Figure 3: Model of assessment system in the Region of Gowa

The management of the semester and class promotion test has extended the role of local government in assessment. Although it is locally implemented, the region now runs two government assessments which (at the same time) reduce the significance of teacher assessment. The region also publishes a report book for each student using guidelines suggested by the BSNP. The report book generates a single composite number that represents a calculation involving the twelve subjects learnt, the KKM, the student's score and a brief description of the students' position against the KKM.

In addition, the region introduced a new policy called 'automatic promotion' in 2011 and this was implemented officially in the region on 2 May 2012 (Pemerintah Kabupaten Gowa, 2012). This programme encourages teachers to assist students in achieving all key competencies in the curriculum by using particular teaching methods. The government expects that no student will be required to repeat a grade

even if they fail in the National Examination. The philosophy of the programme is that all students can learn effectively as long as they are given sufficient time. Therefore, when a students' school attendance level has been 80% or above, they deserve to proceed to a higher level. The government is convinced that another benefit of the programme is that students can complete their education faster due to the utilisation of a credit system in the programme. Following the official launch of the programme, teachers in the region were invited to attend three days of workshops, followed by training with an individual trainer for each school. The workshops and training involved experts from universities in Indonesia. At the time of this study, the government of the region is preparing new syllabi and twelve books to support and explain the programme in more detail. Yet, the programme is not fully established and although schools are expected to manage teacher and school tests autonomously, the regional government's decision to manage semester tests and the regional policy that no child should repeat a grade, appear to undermine teachers' assessment practices.

Definition of key terms

In order to clarify the key terms used throughout this thesis, I present a concise definition of these terms.

National Examination (UN)

UN is the standardised National Examination conducted at the end of elementary, intermediate and high school. The UN is used to capture the quality of education, and provide a tool for student selection and certification. By ranking the results of the National Examination, the UN can identify schools, regions and provinces that need further guidance for quality improvement.

KKM (school benchmark)

Mastery learning is a students' highest level of competency for a subject; the KKM is the minimum level of mastery learning that a student must obtain. The KKM is established by subject teachers in each school before the school year begins. These subject teachers consider three aspects: complexity (level of difficulty), facilities (schools and teaching materials) and the student intake (students' competence and background knowledge of the subject). The KKM, of a subject at each year is decided by looking at the average score of learning indicators, basic

competencies, and key competencies stated in the curriculum. The KKM, which is adjusted up each year, becomes the reference point when teachers are assessing students. They mark students' work and score it to determine whether the standards set by the school have been achieved or not.

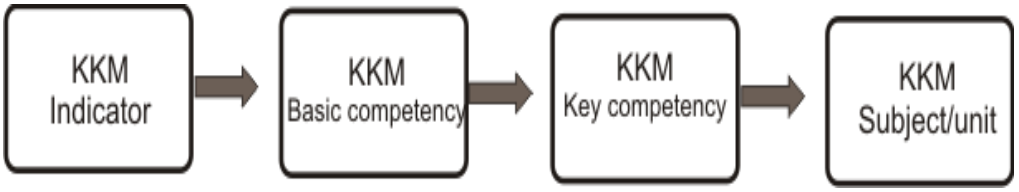


Figure 4: Procedure of KKM establishment

Although all schools utilise the same aspects in setting the KKM, each school has different KKMs to accommodate the different context, school standard (regional and national) and students' background.

Internal assessment

Internal assessment refers to teacher assessment that is conducted to evaluate teaching and learning. In this study, homework assignments, teacher tests of students' mastery of the curriculum and teacher observation of students during instruction are categorised as internal assessment. To some extent, internal assessment signifies classroom assessment, formative assessment or assessment for learning.

External assessment

External assessment covers tests that are conducted by external agencies like the regional department of education and the Ministry of National Education in the country. It denotes summative assessment or assessment of learning. In Indonesia, there is a stronger focus on external assessment compared to internal assessment.

Organisation of thesis chapters

This thesis is made up of five chapters as illustrated in the following order:

Chapter One: Introduction

Chapter Two: Literature review

Chapter Three: Methodology
Chapter Four: Findings
Chapter Five: Discussion and conclusion

Chapter One presents the rationale, focus, and purposes of the study. The research questions are stated along with the expected significance of the findings. An overview of the Indonesian education and assessment system, particularly in the site of the study, was also presented to introduce the official assessment process in the region of Gowa. This chapter also defined key terms and outline structure of the thesis.

Chapter Two reviews literature on assessment including types and purposes of assessment. Several studies of teachers' conceptions of assessment from different settings are presented, along with the gaps that clarify the need for further research in the Indonesian context.

Chapter Three examines the design, the rationale for using mixed methods, and the methods of data collection and analysis. This chapter also addresses the issues of validity, reliability and the trustworthiness of the study. Following this is the theoretical framework used for the study. Chapter Four presents the findings of the study, starting with the quantitative phase and this is followed by the qualitative findings across the three case studies.

Chapter Five gives in-depth insights into the contribution of the study. The major findings are discussed through the lenses of the existing literature and the theoretical framework. This chapter also reviews the implications and limitations of the study, and suggests directions for further research.

Chapter Two

Literature Review

Introduction

This chapter reviews studies in the areas of conceptions of assessment. I start my investigation with a definition of conceptions and an explanation of why the issue is important to study. Following this section I explain how ecological theory frames the contributing factors to teachers' perceptions of assessment. The chapter concludes with a review of the literature on teachers' conceptions of assessment, mostly from elementary and secondary school levels across various disciplines. This review includes a discussion of the two primary purposes of assessment: assessment for learning and assessment of learning.

Conceptions in this study

Some authors differentiate the terms *beliefs* and *conceptions* (Remesal, 2011) while others prefer to use them interchangeably (Calveric, 2010; Vardar, 2010). Pajares (1992) argued that beliefs travel in disguise and often under the aliases of perceptions, values, conceptions, judgments, opinions and the like. Educational researchers have not typically agreed on a working definition of conceptions, However, this study uses Thompson's (1992) understanding of conceptions as a framework of propositions, preferences or general mental structures which are flexible and can change (Green, 1971). I prefer to use the term conceptions because it has been used in relationship to knowledge and facts in comparison with the term beliefs which may rely heavily on evaluative and affective components (Abelson, 1979) and are more likely to involve feelings and emotions (Nespor, 1987). As conceptions relate to practice knowledge, teachers' conceptions could be expected to be influenced by evidence about effectiveness through PD activities.

The importance of studying conceptions

The influence of conceptions in shaping teacher behaviour and action has been investigated in various studies (Brown, 2002; Calveric, 2010; Remesal, 2011). A conception is a lens through which a teacher views, interprets and interacts with

his/her understanding of the world (Pratt, 1992). This means that conceptions influence the way an individual defines his/her work (Nespor, 1987) including teachers' teaching and students' learning (Calderhead, 1996; Pajares, 1992). Conceptions are related to shared social and cultural phenomena (van den Berg, 2002). People's beliefs and the norms of their social environment appear to be crucial in shaping their type of behaviour and practices (Brown, 2008). Griffiths, Gore and Ladwig (2006) report that beliefs affect teaching practices to a greater degree than teaching experience and socioeconomic school context. Thus, any study of teachers' conceptions of assessment should include both teachers' personal beliefs about assessment and the influence of environmental or contextual factors.

The ecological theory: framework of the study

I use sociocultural perspectives to frame my study. The consideration for utilising this theory is that human development processes and outcomes are influenced by environmental factors that consist of several interrelated social systems (Bronfenbrenner, 1977). Bronfenbrenner states that the process of development starts from smaller individual elements known as the microsystem, and then moves to bigger contextual components: the mesosystem, exosystem and macrosystem. Bronfenbrenner contends that the microsystem consists of three patterns that influence a person's development. These are an individual's activities, roles, and interpersonal relations. The mesosystem occurs when two or more settings interact with one another as dyads. The mesosystem extends and develops continuously as people move to new settings, new schools, new offices, or new neighbourhoods. The exosystem highlights hidden factors (like parental work environments) that may not directly relate to an individual but could influence his/her development. The biggest system in ecological theory is the macrosystem which embraces the forms and contents of smaller or lower order systems. A macrosystem 'could exist at the level of the subcultures or the culture as a whole, along with any belief systems or ideology underlying such consistencies' (Bronfenbrenner, 1979, p. 26). Consequently, the macrosystem covers settings in which a person shares the same values, cultures or systems with others. The interrelationship of the systems in this study of teachers' conceptions of assessment can be seen in Figure 5.

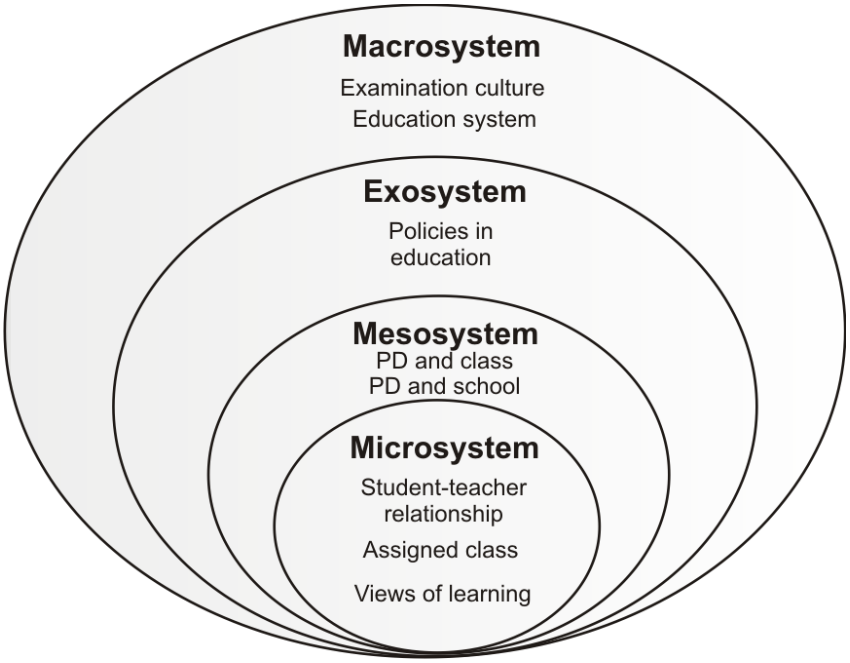


Figure 5: The relationship between personal and contextual factors in shaping teachers' understanding of assessment.

Therefore, in the study of teachers' conceptions of assessment, the influences range from wider environmental factors from the macrosystem where there is a shared culture of valuing examinations (Brown, Hui, & Yu, 2010; Degbey, 2009; Li & Hui, 2007); to an exosystem including the education system (Remesal, 2011; Winterbottom et al., 2008), leading to teacher participation in professional development (PD) (Calveric, 2010; Vardar, 2010), and teaching assignment level (Brown, 2011; Philippou & Christou, 1997) through a microsystem at the level of individual schools to personal factors such as views of learning (Bonner & Chen, 2009; Brown, 2002).

Bronfenbrenner describes the relationship of the systems as 'a set of nested structures, each inside the next, like a set of Russian dolls' (Bronfenbrenner, 1979, p. 3). Bronfenbrenner's framework is mostly used to study child development or parenting (Adamson, O'Brien, & Pasley, 2007; Swick & Williams, 2006). As these studies focus on children, the emphasis centres on the micro level (the child) then moves to wider contextual elements such as home, school, neighbours, community, and so on. My study however, concentrates on the professional life and work of teachers and the way they understand assessment as part of their job. The context of this study suggests that the macrosystem may be the dominant factor that

influences teachers' understanding. This unique interrelationship allows me to present the ecological theory in reverse; starting the discussion at the macro level before examining the micro level.

The macrosystem and the exosystem

The macrosystem encompasses general prototypes that exist in culture and subcultures. It can be seen in formal constructs like regulations, rules, or laws but is mostly informal and implicit like customs, life-styles or bodies of knowledge (Bronfenbrenner, 1977). The macrosystem is the blue print for a particular culture or subculture that ultimately affects the conditions and processes that occur in the microsystem (Bronfenbrenner, 1994). This means that culture strongly influences the participants' ways of understanding. In this study, cultural domains that are embedded in the macrosystem include the cultural expectations of examinations, competition and grading.

The level following the macrosystem is the exosystem. The exosystem includes implementation of educational policies and regulations that are outside the control of teachers but influence their professional decisions. In Indonesia this includes decisions about whether a school will be granted a national or regional standard, and the enforcement of regulations or policies (like ensuring that no child will fail) that override teachers' professional judgements. Most schools have to meet the regional standard and if schools are interested in upgrading to meet the national standard, they must meet additional requirements including a minimum KKM of 7.5 for every subject. This decision is outside the remit of individual teachers.

I now draw upon literature illustrating how the exosytem impacts upon teachers' conceptions of assessment along with studies representing the macrosystem because these circles of influence are closely related.

A culture of examination, and grading

Teacher acculturation into an examination focus is the most evident factor contributing to teachers' conceptions of assessment according to literature (Berry, 2011; Brown, Kennedy, Fok, Chan, & Yu, 2009). While examinations may be unpopular in low-stakes examination contexts and sometimes at elementary school level, at secondary school level examinations are usually a crucial assessment focus for teachers. Interestingly, most Asian studies on assessment reveal that the

examination is a primary consideration for teachers at all levels. A number of these studies use Brown's validated Teacher Conceptions of Assessment survey (TCoA).

In one such study, Brown and his colleagues (2009) used TCoA to survey teachers in Hong Kong. Participants in the study supported the purpose of assessment to improve teaching and learning as well as to make students accountable for their learning, but disagreed with the assumption that assessment has little impact on teaching. However, teachers' assessment practices were in opposition to their beliefs as their teaching focus was on preparing students to pass examinations. In other words, Hong Kong teachers believed that the improvement conception was compatible with the examination process.

The Hong Kong study was very interesting in the sense that although the country was colonised by British, and English is emphasised in the language curriculum, the English education system has had very little influence on teachers' practice (Sweeting & Vickers, 2007). Huge numbers of refugees moving from China to Hong Kong after the civil war in China in 1949 caused Hong Kong to adopt a screening mechanism for schooling (Berry, 2011). Even now teachers, education officials and parents believe that examination success is the best qualification and the main determinant for admission to either secondary or tertiary education in Hong Kong (C.-C. Choi, 1999). As a result, even though the education policies in Hong Kong have tried to promote assessment for learning for more than a decade, teachers' conceptions of the high-stakes social function of assessment and their subsequent practices tend to block the reform agenda (Brown, et al., 2009; Kennedy, Chan, Fok, & Yu, 2008). This does not necessarily stem from teachers' disagreement with the policy, it exists as a result of ingrained practices of testing and competition among people of Hong Kong, Indonesia (Zulfikar, 2009) and other Asian settings like Korea. One reason for this firm belief in the value of examination was revealed in Choi's (2008) study which found that Korean teachers believed that 'tests provide opportunities for the entire population to climb up the ladder of social status' (p.41).

Correspondingly, history and culture were considerations among Chinese teachers (Li & Hui, 2007). Using the first version of the TCoA survey (the one used in New Zealand) the study found that participants differentiated between the functional purpose of assessment and its evaluative function (Li & Hui, 2007). These teachers agreed to the functional purpose of assessment where teachers could use

assessment to improve teaching and learning. However, Chinese participants understood this to mean that assessment assisted in preparing students for better performance in an examination. Participants of this study did not support the evaluative function of assessment which refers to the credibility and validity of assessment. In fact, these teachers doubted that assessment could provide valid information about students' learning. In other words they argued that assessment results were inaccurate.

Similar findings were revealed in studies conducted in Hong Kong and Southern China (Brown, et al., 2010). Teachers participating in these studies understood improvement to mean improving teaching and learning through examinations. For this conception, teachers also believed that assessment motivates students to work harder and become better people (C.-C. Choi, 1999). This literature indicates that Asian teachers shared a culture of assessment by examination which is also reflected in long established teaching practices such as transfer and drilling (Brown, et al., 2009). Rote learning was found to be effective preparation for examinations that in turn were used to evaluate teaching or to improve learning. Developing students' ability to recall facts and information is a relevant strategy given the test format in these countries. Thus mastering more facts in preparation for examinations is believed to improve learning. This phenomenon encouraged teachers to practice traditional assessment strategies like grading and illustrates how the improvement conception is perceived differently in Asian contexts in contrast to New Zealand, and Australian studies.

In contrast to participants in these Asian studies, teachers holding improvement conceptions in New Zealand and other lower-stakes examination contexts perceive that improvement means allowing them to use a range of assessment strategies to improve learning. In these countries, rote learning and tests are less prioritised possibly because a lower-stakes system of evaluation places less emphasis on student scores. Different national policies indicate the different priority given to assessment purposes and types. In countries where low-stakes assessment is implemented, like in New Zealand and Australia, teachers are encouraged to use assessment for learning (formative assessment) whereas in Asian settings, teachers prefer and are accustomed to assessment of learning (summative assessment). As stated earlier, studies conducted in Asian contexts find that teachers' believe good examination results are an effective means of determining

students' future success. This understanding is reflected in the importance of grading practices in examination cultures. The Hong Kong, China, and Korean studies suggest that the role of examinations in assessment conceptions may be culturally-embedded and shared (Berry, 2011; Kennedy, et al., 2008; van den Berg, 2002). However, comprehensive investigation is needed to test this assumption.

The education system, policies and regulations

Ravitch (1995) notes that national assessment protocols are intended to promote equal educational opportunity by providing accurate information to students, parents, teachers and administrators. Nevertheless, the results across regions, schools, and various population groups typically reveal differences that have been attributed to factors such as culture, social class and school composition (Amrien & Berliner, 2002). Internationally, various countries have implemented national assessment protocols, including the No Child Left Behind (NCLB) legislation in the USA, the Standard Attainment Tasks and Tests (SATs) in the UK, the National Standard (NS) in New Zealand, National Testing (NT) in Norway, and Primary School Leaving Examinations (PSLE) in Singapore. These high stakes assessments are implemented at certain levels of education for particular subjects with the stated aim to evaluate students' knowledge and skills enabling them to succeed in the world of work (Cheng, Rogers, & Wang, 2008; I.-C. Choi, 2008), to show whether students have reached the national curriculum targets, and/or to prepare students with educational opportunities (Grant et al., 2002; Gregory & Clarke, 2003). While the assessment standard system in the USA, UK and Singapore have been in use for decades, the systems in New Zealand and Norway are relatively new (Thrupp, 2008). The latter two countries based the introduction of NS and NT on their national results on the internationally administered Progress in International Student Achievement (PISA) (Ozerk & Whitehead, 2012). This standardised testing is seen to ratchet up the accountability of students and teachers as well as raise expectations for students expected to meet the educational standards (Linn, 2000) although the process is highly influenced by the political system (Isaac, 2010). Despite attaining the accountability purposes of the national assessment system, or developing clearer conceptions of performance standards among teachers (Gregory & Clarke, 2003) national standards have been censured for narrowing the teaching focus (Segers & Tillema, 2011) and changing teacher education accordingly. This is particularly evident in New Zealand where

currently literacy and numeracy dominate teacher education at the expense of the arts (Ozerk & Whitehead, 2012).

The education system, policies and regulations affecting education are sub cultures that exist in the macrosystem. Policies, regulations and the education system are among crucial aspects contributing to teachers' conceptions of assessment (Degbey, 2009; Remesal, 2011; Winterbottom, et al., 2008). In most studies cited in this review, participants felt obliged to follow policy, regardless of their own assessment values, particularly in high-stakes educational environments.

The role of educational policy in influencing teachers' conceptions of assessment was evident in the study of Cypriot and Greek teachers (Philippou & Christou, 1997). Their mixed methods study found that participants prioritised the power of policy and regulation to shape their understanding of assessment. This potentially brought teachers into a conflict between their beliefs and practices of assessment. Likewise, in a Cambridge study, Winterbottom, et al. (2008) drew similar conclusions. Although *facilitating better performance* gained the lowest preference among teachers, they nevertheless felt that it had the strongest impact on their assessment practices. The learning environment and high-stakes policy system caused teachers to disregard their preference for using assessment to inform learning. Teachers in this study felt that they complied with educational policy at the expense of their beliefs about good assessment practices.

Similar conflict is also found in a Finnish study (Degbey, 2009). Participants reported their preferences for using assessment strategies and techniques to improve teaching and learning. Yet, they did not feel able to follow through in their own practices due to perceived pressures to prepare students for examinations which they regarded as fulfilling policy requirements. In line with the Finnish study, Remesal (2011) found that Spanish teachers' conceptions of assessment were also determined by the educational system. Secondary school teachers there held societal conceptions of assessment because of the practice of using assessment as a tool for certification or accreditation of student achievement.

The power of policy was even more visible in a Singaporean study (Noor, Muniandy, Krishnan, & Mathai, 2010). These authors revealed that a strong accountability and certification focus led the Singaporean government to ignore issues of

trustworthiness in assessment. Teachers questioned the reliability of tests because different examiners had different interpretations of the descriptors and the test rubrics. Although these teachers perceived the examination to be unreliable, they still argued that it had an important position in describing student achievement. The study suggests that a strong focus on accountability might override concerns about the quality of a test.

Findings from these studies suggest that educational policies and regulations play crucial roles in shaping teachers' conceptions and beliefs about assessment (Barnes, Clarke, & Stephens, 2000; Chan, 2007). This includes priority given to types of assessment (White, 2007). In all these studies, teachers were reported to agree with the function of assessment to improve teaching or to enhance learning but they were given insufficient latitude to implement and develop this understanding. Participants in these studies balanced accountability conceptions resulting from the national educational policy with conflicting conceptions, indicating that their assessment practices were likely to contrast with their beliefs. Moreover, teachers' autonomy tended to be overlooked in the sense that they were not involved in the decision-making process, including negotiating the priority given to certification. Evidence from this literature shows that the exosystem strongly influences teacher perception and practice of assessment.

Mesosystem

The relationship between systems that involve the developing person forms the mesosystem. This includes relations between workplace and other settings such as home and school (Bronfenbrenner, 1994). My study comprises an investigation of a mesosystem in teachers' conceptions of assessment through the interaction between professional development programmmes (PD) and the expectations from the school. Through PD teachers interact with their peers and other experts who could improve their assessment knowledge. In this review, the term PD covers seminars, workshops or training that teachers attend either within or outside school.

The effect of assessment training in shaping teachers' understanding of assessment is evident in Calveric's study (2010) of USA's central Virginian elementary school teachers. Her participants reported that their assessment literacy was improved after participating in PD meetings. In the same way, Turkish teachers in Ankara believed that PD could assist them gaining further information about assessment

(Vardar, 2010). In other words, teachers in both Calveric and Vardar's studies agreed that PD could improve their assessment knowledge which previously highlighted by Borko (2004) or that it could help to reconcile the conflict of belief and practices in assessment. This conclusion is shared by Bumen (2009) who claimed that involvement in PD activities could positively influence teachers' ability to teach effectively. This author argued that whenever teachers saw new strategies modelled and were given opportunities to apply their new knowledge, they could communicate new ways of learning to students or implement new teaching strategies. In this way PD on assessment is able to improve teachers' assessment literacy and may be used to improve teaching and learning. Relevant to this claim, Dole, Nisbet, Warren, and Cooper (1999) reported that PD positively changed teachers' conceptions of assessment. Focusing their study on professional interpersonal relationships, they found that teachers from Queensland, Australia, were more confident in assessing students after participating in critical group sharing in PD. The programme appeared to enrich teachers' ideas and reflections on classroom practices which in turn improved both their teaching and assessment practices. These studies suggest that PD can help teachers to be more literate in and more skilful at assessment as well as enriching the range of assessment strategies they use in the classroom.

However, changing teachers' conceptions of assessment through PD programmes is only one among several ecological factors that could contribute to teachers' beliefs about assessment. Teachers may possess good comprehension of assessment; however, other ecological factors can dominate and block the application of this knowledge. These inhibitors include culture and educational systems, which I highlighted earlier. Furthermore, influences within the inner circle of the system such as student-teacher relationships and teachers' views of learning are worth investigating.

Microsystem

Bronfenbrenner describes the microsystem as:

> "…a pattern of activities, roles, and interpersonal relations experienced by the developing person in a given face-to-face setting with particular physical, social, and symbolic features that invite, permit, or inhibit engagement in sustained, progressively more complex interaction with, and activity in, the immediate environment". (1994, p. 39).

This section looks at the influence of teachers' immediate environment including their relationship with learners, the class assignment level, and the different views of learning held by teachers.

Student-teacher relationship and curriculum assignment level

As this study focuses on the professional life and work of teachers, their immediate microsystem is the workplace: classroom and school. Within the classroom and school environment, teachers develop professional relationships with students that inform their perceptions about students' ability, competency, and proficiency. Numbers of studies have been conducted to address teacher-student relationships and how teachers' perceptions of students influence teachers' assessment of student proficiency (Fowler, Banks, Anhalt, Hinrich Der, & Kalis, 2008; Hamel, 2003).

Teacher-student relationships were a focus of Hamel's (2003) case study of three high school teachers in Tacoma, Washington. Hamel found that teachers' understanding of students included their perceptions of student ability and student social factors as well as the teachers' level of experience. The social factors are the settings or the social context where students live and teachers' experience ranged from teachers' experience as students, their formal teacher education experience (either at college or PD) and teachers' teaching experience. The author concludes that these influencing factors directed teachers to adjust the content of curriculum according to their perceptions of student competencies.

Similarly, Fowler and his colleagues investigated 230 students and twenty teachers in two high-poverty, low-performing schools in the US Midwest to study the quality of teacher-student relationships. These authors suggested that although the relationship between the quality of teacher-student relationship and the way teachers assessed student academic performance was not statistically significant, the use of multiple regression revealed that a relationship between the two exists and needs further investigation.

Another microsystem that influences teacher-student relationships and teacher perception and practice of assessment is the teaching assignment level. Elementary and secondary school teachers (or early and final level teachers) tend to have dissimilar approaches to teaching that affects their classroom interaction. Studying

teachers from two states in Australia (New South Wales and Victoria), Barnes, Clarke, and Stephens (2000) found that teachers at early levels of junior secondary (Years 7-10) were not affected by mandated assessment. This means that teacher-student interaction could focus on assisting student learning and giving less priority to test-taking techniques. In contrast, teachers of students in their final years (Years 11-12) reported that their assessment practices focused on a combination of school-based assessments and end of year examinations. Similarly, in a qualitative study of fifty school teachers in Catalonia, Spain, Remesal (2007) found that primary teachers implemented formative assessment practices whereas secondary teachers maintained and promoted summative assessment practices. These practices reflected the educational reform in Spain that promoted the practice of formative assessment. It eliminated the external standardised tests at primary school level, but tests remained implemented at the senior secondary school level. This study also connects teachers' access to assessment training to the different teaching assignment levels and finds that primary and secondary school teachers have different assessment literacy.

Views of learning

The teaching assignment level influences and is influenced by teachers' views of learning which in turn plays a crucial role in shaping their beliefs about assessment. James (2008) theorised three major views of learning as behaviourist, constructivist and socio-cultural. She explains that those who hold a behaviourist view of learning tend to focus on performance, and students' ability to recall facts and information. These teachers are likely to assess students' responses as correct or incorrect. On the other hand, teachers who hold constructivist views of learning focus on problem solving and understanding. This view allows teachers to expand and vary tasks so that students can demonstrate deeper understanding, and such teachers are likely to assess students' responses to the task against specific criteria. For teachers who hold socio-cultural views, learning is seen as a 'social and collaborative activity in which people build knowledge and develop their thinking together' (James, 2008, p. 30). According to this view, assessment is carried out alongside learning, and involves self-assessment, peer assessment and teacher assessment. The following studies of teachers' conceptions of assessment reveal that different views of learning contribute to teachers' beliefs and practices of assessment, and these views may also be influenced by the students' curriculum level.

Bonner and Chen's (2009) quantitative study of how New York teachers' views of learning shaped teachers' conceptions of assessment found that elementary level teachers endorsed what James (2008) might consider to be a constructivist or socio-cultural approach. Participants in their study believed in alternative assessments such as portfolio and project work. Conversely, secondary school teachers supported a more traditional approach to grading, suggesting a behaviourist view of learning. Likewise, a study conducted in Queensland, Australia revealed similar findings (Brown, Lake, & Matters, 2011). Primary school teachers in the study perceived assessment as improving teaching and learning, indicating that they held constructivist views whereas secondary school teachers viewed assessment as making students accountable for their learning which reflects behaviourist views.

These findings are disputed in a comprehensive study involving New Zealand teachers (Brown, 2002) which found that primary and secondary school teachers shared similar views of assessment. Participants favoured assessment to improve teaching and learning more than for the purpose of accountability. New Zealand teachers at both primary and secondary levels appeared to hold socio-cultural views of learning which led them to believe that assessment for external accountability does not measure deep transformative learning.

These findings suggest that whatever assessment types are implemented in a setting, the policy influences teachers towards holding behaviourist, or constructivist/socio-cultural views of learning (James, 2008). In high-stakes assessment contexts, early and final year teachers apparently hold dissimilar views of learning. In contrast, teachers teaching in low-stakes assessment settings appear to hold similar views regardless of the teaching level they are assigned. In other words, teachers' decisions to use different strategies in assessment may depend on their interpretation of policies affecting the level of schooling in which they were teaching.

These studies reveal that different microsystems may lead to dissimilar conceptions and practices of assessment. Different levels between elementary and secondary teaching may also influence teachers' perspectives. Although teachers frequently hold interconnected conceptions, secondary teachers appear to be more closely affected by the assessment policy determined by their educational system. This

indicates that teachers' ways of understanding assessment are complex and are influenced by several wider contexts like the exosystem, mesosystem and macrosystem. This literature shows that both personal and contextual components play a role in shaping teachers' conceptions and practices of assessment. While personal elements are embedded in each individual's characteristics, contextual elements are wider systems that impact the development and process of understanding. The following section highlights the literature on assessment conceptions and types of assessment that occur due to the interrelationship of socio ecological factors.

Teachers' conceptions of assessment

An early study of teachers' conceptions of assessment was conducted by Wolf, Bixby, Glenn and Gardner (1991) who distinguished between assessment culture and testing culture. These authors believed that teachers' understanding of intelligence, the process of teaching and learning, the nature of tasks and the evaluation criteria influenced teachers' conceptions and practices of assessment. Delanshere and Jones (1999) also proposed three dimensions to identify teachers' conceptions of assessment. Those dimensions are (a) students' placement according to achievement level; (b) teacher's perceptions of curriculum and professional self-efficacy; (c) teachers' beliefs about teaching and learning and about students' as learners. All these authors consider the relationship between assessment and teaching and learning.

Significant studies of teachers' conceptions of assessment have been conducted by Brown and his colleagues since the early 2000s. In his study of New Zealand teachers, Brown' introduced four teacher conceptions of assessment: a focus on improvement or teachers' views of assessment as a tool to improve teaching and learning; assessment as driven by school accountability purposes; assessment for student accountability; and perceptions of assessment as irrelevant or a meaningless practice in daily school life (Brown, 2002). This model has been subsequently validated in several studies (Brown, et al., 2011; Calveric, 2010; Segers & Tillema, 2011). Brown's model has also been adjusted and modified to fit Asian contexts by adding examination as another crucial dimension for high-stakes assessment settings (Brown, et al., 2010; Brown, et al., 2009).

However, Remesal (2009) found that the TCoA model did not suit Spanish teachers. This challenged her to develop a continuum of assessment purposes. At one end she places pedagogical conceptions (assessment for monitoring of teaching and learning) and at the other extreme she identifies a societal-accreditation view (assessment for teachers' accountability and certification of achievement) with some mixed conceptions between the two poles (Remesal, 2011). The continuum is based on four roles of assessment according to how teachers used assessment in learning, in teaching, in the certification of learning and for accountability when students' achievement is used as an indicator of the teachers' professional capacity. Despite the differences in their studies, both Brown and Remesal agree with previous authors that assessment could and should benefit both teachers and learners.

A more recent study on the issue of teachers' conceptions was conducted in Helsinki, Finland (Postareff, Virtanen, Katajavuori, & Lindblom-Ylanne, 2012). These authors identify conceptions as being either reproductive or constructive. The reproductive conception emphasises students' understanding through memorisation of the content of the study module and how it is applied in real life. The constructive conception on the other hand, looks at deeper understanding including reflection and justification for an argument. Unlike former studies, Postareff and her colleagues focus on the impact of assessment on learning.

Interestingly, although these authors identify negative values of assessment, Brown (2002) has been the only researcher to develop these into an independent category, which he called the *irrelevance conception*. This category covers several indicators including teachers' ignorance of assessment results, beliefs that assessment interfered with teaching, beliefs that assessment is an imprecise process, that assessment has little impact on teaching and that assessment could cause contradictions between teachers' beliefs and practices. Remesal approaches the concept of irrelevance in her continuum of pedagogical and accounting conception (Remesal, 2007) where she identifies a category called *mixed undefined conceptions* which stand in the middle of the continuum. This category represents participants' unclear preference for one wing of the continuum which appears to be a neutral conception rather than opposing or negative conception. Furthermore, in her more recent category of pedagogical and societal conceptions (Remesal, 2011), she migrates negative indicators to one or other of the conception types. These

indicators include assessment that may have an unrealistic or bad influence on learning, assessment that has no or a negative impact on teaching and teachers' disagreement with external evaluation when it does not fit the context. Thus, both Brown and Remesal emphasise potentially negative impacts of assessment on teaching and learning.

More importantly, although these researchers use different terms to address assessment conceptions, they appear to refer to the two classic purposes of assessment: summative (Broadfoot, 2007; Wiliam & Black, 1996) or formative purposes (Black & Wiliam, 1998b). Table 1 summarises how these researchers define conceptions of assessment according to its purposes.

Table 1: Conceptions based on purposes

Purposes	Types of conceptions and authors				
	Delanshere & Jones (1999)	Brown (2002)	Black & Wiliam (2007)	Remesal (2007, 2011)	Postareff. at al., (2012)
To improve teaching and to enhance learning	Teaching, learning and learners	Improvement	For learning	Pedagogical	Transformational
To control students	Placement	Student accountability	For certification	Societal	Reproductive
To control teacher and school	Evaluation	School accountability	Accountability		
To have little impact on teaching		Irrelevance			

Based on the literature on assessment purposes I have drawn a continuum of assessment with two different focuses. At one pole is assessment for learning (AfL) while assessment of learning (AoL) is at the other end of the continuum.

Black & Wiliam, 1998a	Assessment for learning				Assessment of Learning
Delanshere & Jones, 1999	Teaching, learning and learners			Placement	Evaluation
Brown, 2002	Improvement			Student accountability	School Accountability
James, 2008	Socio-cultural	Conctructivist			Behaviorist
Remesal, 2007, 2011	Pedagogical	Mixed Pedagogical	Mixed Undefined	Mixed Societal	Societal
Postaerff, et al., 2012	Transformative		Application		Reproductive

Figure 6: Continuum of assessment

Figure 6 shows that maximum differentiation only occurs at its extreme poles. At the AfL pole, authors used various terms such as improvement, pedagogical, transformative conceptions that align with socio-cultural or constructivist views. At the opposite pole, authors introduce terms like evaluation, school accountability, societal and reproductive, that typically represent behaviourist views. Further investigation of the assessment continuum is highlighted in the next section.

Assessment for learning (AfL)

Assessment for learning or formative assessment (FA) focuses both on students' learning and teacher's teaching and these functions are inseparable (Black & Wiliam, 1998b; Harlen, 1998). In other words, the purpose of using FA to assess students' learning and its application to make beneficial changes in instruction creates a tight link with instructional practices (Boston, 2002). These assessment purposes reflect both improvement conceptions (Brown, 2002) and pedagogical conceptions (Remesal, 2011).

Improvement or pedagogical conceptions emphasise the use of information to monitor and produce valid changes in teaching and learning (Black, Harrison, Lee, Marshall, & Wiliam, 2002). This conception requires teachers to make reliable and accurate descriptions of students' performance (Brown, 2002). Various strategies and techniques used in the practice of teachers holding this conception include

informal teacher-based intuitive judgment to complement formal assessment tools. These techniques function to 'identify the content and process of student learning with the explicit goal of improving the quality and accuracy of instruction and/or enabling students to improve their own learning' (Harris & Brown, 2008, p. 2).

AfL requires the involvement of students in the assessment process, either through self-assessment or peer assessment or their participation in determining criteria for evaluation. Student involvement in assessment enables them to view the quality of their own work and modify it to meet the criteria (Sadler, 1998). This process requires the teacher to give feedback on student understandings and areas to be improved, or to offer suggestions about how to improve (Boston, 2002; Hattie & Timperley, 2007). In this way AfL, improvement or pedagogical conceptions assist teachers to achieve a comprehensive and balanced picture of student achievement which is believed to be the key link between assessment and learning (Gipps, McCallum, & Hargreaves, 2000).

AfL also relates to socio-cultural and constructivist views of teaching where a focus is placed on problem-solving and understanding as well as developing thinking (James, 2008). Constructivist pedagogy is "...concerned with the teacher's modelling of how individual pupils are thinking and understanding so that the next challenge, prompt, question or information can lead the learner forward" (Butterfield, Williams, & Marr, 1999, p. 228). This view requires teachers to employ careful listening and observation to understand their pupils, and this activity is integrated in the teaching process (Butterfield, et al., 1999; Shepard, 2000a). In short, AfL, improvement or pedagogical conceptions are likely to enhance teacher teaching as well as student learning.

Literature shows that improvement or pedagogical conceptions are mostly held by primary school teachers (Brown, 2011; Remesal, 2007) or teachers teaching at early levels in secondary schools (Barnes, et al., 2000). Reasons for this preference include different policies at primary level such as less formal or external tests. This allows primary school teachers greater opportunities to improve their literacy in assessment for learning compared to their secondary school teacher counterparts (Remesal, 2011). Further, it indicates that PD may play a role in empowering teachers as suggested by several authors (Calveric, 2010; Dole, et al., 1999; Vardar, 2010).

Assessment of Learning (AoL)

Assessment of learning or summative assessment (SA) is a way of looking back or summing up learning (Broadfoot, 2007; Wiliam & Black, 1996). SA involves marking and grades which require unified procedures that enable comparability among the results of all students (Harlen, 1998). This assessment type focuses on measuring the extent to which students reach or do not reach required standards (Firestone, Mayrowetz, & Fairman, 1998). It is used as a means to record the nature and level of students' achievement throughout their academic careers (Hill, 2000). SA or AoL can be conducted by teachers as well as external agencies. External agencies often use the results for accreditation, accountability and monitoring (Harlen, 2005), selection, placement and certification (Black & Wiliam, 1998a), as well as for public reporting (Guthrie, 2002).

Brown's (2002) student accountability and school accountability conceptions or Remesal's (2011) societal conceptions reflect AoL. This conception type focuses on institutional goals or communal and societal perspectives and interests. In other words, this assessment purpose is used to account for teachers, schools, or systems use of society resources. To this end, teachers who equate assessment with school accountability or societal or conventional conception emphasise two rationales; that assessment is used for demonstrating the quality of school and teacher instruction (Smith & Fey, 2000), and for improving the quality of that instruction (Linn, 2000).

To some extent, this conception considers student learning, however it focuses more on societal interests such as reporting student achievement, attitude and effort against curriculum standards, or comparing students against one another and/or against their prior individual achievements (McMillan, 2001). In order to achieve these purposes, several strategies such as grading, criterion reference tests, and awarding certificates or qualifications based on performance are popular practices (Harris & Brown, 2008). Teachers who hold accountability or societal conceptions support high-stakes tests which they believe to be practical, and aspire to make assessment transparent (Linn, 2000).

Studies of teacher's conceptions of assessment reveal that accountability or pedagogical conceptions are usually held by teachers with behaviouristic views of learning (James, 2008). Teachers with such views are likely to teach at secondary

level (Barnes, et al., 2000; Brookhart, 2011; Remesal, 2011) where high-stakes tests are implemented for accountability and certification purposes (Black & Wiliam, 2007) as well as for measuring student mastery of content (Postareff, et al., 2012). At this level, tests are also conducted to signify whether schools and/or teachers are doing a good job (Butterfield, et al., 1999).

Accountability conception seems to be preferred by teachers in educational settings in which priority is given to summative assessment (Chan, 2007). Educational settings that have practiced screening mechanisms (Berry, 2011) and have historically implemented competitive tests, are the likely contexts for this conception because people are acculturated to accept that this is the most effective judgement (Brown, et al., 2009; Philippou & Christou, 1997). In those settings, scoring good results in a high-stakes test is believed to be the best indicator of achievement and these results are influential in determining a student's future working life (Cheng, 2008; I.-C. Choi, 2008; Kennedy, et al., 2008). When assessment is perceived in such a way, it is usually a value that is shared socially (van den Berg, 2002) or is embedded in a community's identity and culture.

Nevertheless, teachers may hold various combinations of conceptions like those falling between pedagogical and societal (Remesal, 2011), improvement and accountability (Brown, 2002), transformational and reproductive (Postareff, et al., 2012). This condition potentially brings teachers to assessment practices that conflict with teachers' genuine understanding of assessment. Such complex understandings of assessment might cause internal disagreement, refusal to comply, or negative impressions of the purposes of assessment which Brown identifies as irrelevance (Brown, 2002).

Overall, literature on assessment conceptions appears consistent with ecological perspectives of human development (Bronfenbrenner, 1977, 1979). The different policies implemented at curriculum assignment level reveal sub-cultures in the macro system. Teachers' participation in PD and how this is implemented in the classroom and school denotes the mesosystem and teachers' views of learning and interaction with students signify the microsystem.

Summary

In this chapter, I have reviewed and presented literature regarding teachers' conceptions of assessment and types of assessment. The conception types of either accountability, or societal and improvement, or pedagogical closely relate to the types of assessment: summative or formative. In most study settings in this literature, summative assessment was dominant and sometimes conflicted with teachers' desires to use formative assessment. It appears that a teacher's beliefs and practices are influenced by interrelated factors at the levels of the macrosystem, exosystem, mesosystem and microsystem.

The next chapter will discuss the methodology used for this study; the pragmatic paradigm.

Chapter Three

Methodology

Introduction

This chapter presents the research approach of my study. Here I justify using a mixed methods design and explain the sampling strategy. Each of the two methodologies is presented separately. Firstly, I consider the quantitative design component, selection of participants and the instrument used, followed by procedures for data collection and analysis. Next, the qualitative component is presented, including the case study design, purposive sampling of participants and procedures for data collection and analysis. The last part of the chapter comprises of the validity, reliability and trustworthiness of the mixed methods design and outlines ethical considerations.

Research approach

This study applies a pragmatist philosophical position so that the research design could be planned and conducted to address the research questions (Johnson & Christensen, 2008; Tashakkori & Tedddlie, 2003). A major tenet of pragmatism is that quantitative and qualitative methods are compatible in the sense that qualitative and quantitative data sources can be mixed (Leech, Dellinger, Brannagan, & Tanaka, 2009). Both approaches are combined and integrated in this study to complement one another (Hewson, 2006). In other words, the combination aims to "provide strengths that offset[1] the weaknesses of both quantitative and qualitative research" (Creswell & Plano Clark, 2011, p. 12).

This pragmatist paradigm is also called multi methods research or mixed methods research (Gray, 2009) where the researcher collects and analyses both quantitative and qualitative data and integrates the two forms of data concurrently or sequentially and gives priority to one or both forms of data in a single study or in multiple phases of research (Creswell, Plano Clark, Gutmann, & Hanson, 2003).

[1] Offset refers to the suggestion that research involving both quantitative and qualitative data have their own weaknesses and strengths and the combination allows the researcher to offset their weaknesses to draw on the strengths of both (Bryman, 2008).

The premise of the combination is to provide a 'better understanding of research problems than either approach alone' (Creswell & Plano Clark, 2007, p. 5). Thus the combination seeks the best of both methods (Bergman, 2008).

Justification for using mixed methods

There are several reasons for using mixed methods research as the paradigm for this study. The complementary function of a mixed method approach allows the researcher to see "elaboration, enhancement, illustration, and clarification of the result of one method with results from the other method" (Johnson & Christensen, 2008, p. 451). It also enables the researcher to use the quantitative data and results to identify those individuals who may expand the results through qualitative data (Mertens, 2003). Another reason is that previous studies on the issue of teachers' conceptions and practices of assessment were conducted using either a quantitative survey design or a qualitative design. While large scale quantitative studies allow generalisation of the findings, the method is not designed for in-depth understanding of a phenomenon (Johnson & Christensen, 2008), such as the understanding behind teachers' conceptions of assessment. In contrast, although qualitative methods provide information relevant to an in-depth understanding of the phenomenon under investigation (Degbey, 2009; Noor, et al., 2010; Remesal, 2011), they cannot be generalised to other people or settings (Johnson & Christensen, 2008). Therefore, in this study, quantitative analysis preceded the collection of qualitative data to probe patterns emerging from survey findings.

Sampling strategy

The sampling strategy in mixed methods involves constructing a sample scheme and determining a sample size in both the quantitative and qualitative components of a study (Johnson & Christensen, 2008). A sampling scheme is defined as 'special strategies used to select units, for instance people, groups, events, settings' (Onwuegbuzie & Collins, 2007, p. 283). Considering the purpose and to maximise understanding about the underlying phenomenon, I used purposive sampling or non-random sampling for both the quantitative and qualitative phases. Purposive sampling refers to the selection of participants by the researcher based on his/her consideration that participants involved in the study have experience of the central phenomenon (Creswell, 2007).

In the first quantitative phase of the study, purposive sampling was used specifying characteristics of the population relevant to the study and locating those individuals matching those set characteristics (Johnson & Christensen, 2008). This convenience sampling approach yielded two characteristics of teachers as identified in the database: (1) those teaching English at junior secondary school in Indonesia; and (2) those who are actively involved in attending meetings as part of a structured PD. In the second qualitative phase, a purposive sampling model was utilised to select participants who were likely to best answer the research questions. This study used a *nested sequential* mixed sampling design (Johnson & Christensen, 2008) meaning that the participants selected for the second phase were selected from the first phase of the study. Findings from the first phase provided the basis for selection of the second, qualitative phase of the study.

Limitation of mixed methods inquiry

One of the challenges in using this model is the extensive time needed to gather data (Creswell & Plano Clark, 2007; Johnson & Christensen, 2008). However the focus placed on the qualitative phase enabled the researcher to limit the number of participants in the study. There were 107 teachers who completed the questionnaire and only twelve of those were interviewed. These two phases helped me gather data within a limited time (Creswell, Plano Clark, & Garret, 2008). In addition, the sample sizes also justified the purpose of collecting enough qualitative information in order to develop meaningful themes (Creswell & Plano Clark, 2011).

Other challenges were ensuring the practicability of instruments and analysing both types of data (Ivankova, 2004). An internationally validated survey was used for the quantitative phase of the study. This questionnaire was translated into Bahasa, with the translation checked using back translation. Pilot testing was also conducted in order to overcome any ambiguity that might prevent participants from understanding the questionnaire. Similar procedures for the second phase involved developing relevant interview questions and also piloting these with teachers who were not participating in the real study.

The interpretative nature of the qualitative phase carries a risk of investigator's bias when analysing the findings. I mitigated against this by triangulating different data

sources, using member-checking of transcript data and using self-reflection through memos and a reflective journal (Creswell, 2003a).

Phase One: Quantitative

Mixed Methods Sequential Explanatory Design

There are three issues to be considered in designing a mixed methods study: priority, implementation, and integration (Creswell, et al., 2003). Priority refers to which approach, either quantitative or qualitative or both, is given more attention throughout the data collection and analysis process in the study (Creswell, 2003). Decisions about the priority of the approach to be used in a study can be made before data collection, during data collection or later in the analysis process (Ivankova, Creswell, & Stick, 2006). This study placed more emphasis on the qualitative strand because I wanted both to investigate reasons behind teachers' conceptions of assessment, and how their conceptions are implemented in practice. To do this, case studies seemed an appropriate means of capturing the contextual nature of the phenomenon under investigation (Yin, 2009).

Implementation refers to whether the quantitative and qualitative data collection and analysis are conducted in sequence, one following another, or concurrently, at the same time (Creswell, et al., 2003). This study adopts a *sequential explanatory design* where a quantitative survey precedes the qualitative phase (Creswell & Plano Clark, 2007).

Integration refers to the stages in the research process where the mixing of quantitative and qualitative methods occur (Creswell, et al., 2003; Tashakkori & Tedddlie, 2003). The integration can occur in the formulation of research purposes (Tashakkori & Tedddlie, 2003) or in the intermediate stage where the results of the first data analysis are used to guide data collection for the second phase of the study (Hanson, Creswell, Plano Clark, Petska, & Creswell, 2005). The mixing can also occur at the interpretation stage of the study (Onwuegbuzie & Tedddlie, 2003). I connected the quantitative and qualitative phases during stage two by selecting participants for the qualitative phase from those who responded to the survey in the quantitative phase. Secondly, I connected the phases by using the quantitative survey findings to inform the development of relevant interview questions to explore in greater depth the participants' beliefs. The third integration is conducted during

the discussion of findings. A visual model of how the study was integrated is presented in the following figure.

Figure 7: Visual diagram of sequential explanatory design: participant selection model (Creswell & Plano Clark, 2007, p. 73)

To summarise, the first phase of the study was undertaken to categorise teachers' conceptions of assessment employing quantitative methods to cast a "wider net" which then allowed for purposive selection of participants to "target a specific population of interest" for the second qualitative phase the research (Hesse-Biber, 2010, p. 465). The second phase employed a qualitative case study approach involving semi-structured interviews and documents analysis to examine more closely teachers' assessment conceptions, to understand reasons for their preferences and to capture how teachers perceive their conceptions are reflected in practice. In accordance with mixed-methods research protocols, the qualitative builds upon initial quantitative results (Creswell, et al., 2008) towards better understandings of the phenomenon (Mertens, 2003).

Participants

Participants were teachers of English at junior secondary schools from the Gowa region South Sulawesi, Indonesia. They were recruited by accessing the database held by the Department of Education to identify teachers of English who were actively participating in a PD programme currently being conducted in the Gowa Region, South Sulawesi. This yielded 152 potential respondents.

Teachers working in the Gowa region were specifically targeted as Gowa is fairly typical in size and population of a region in the South Sulawesi province. There are nineteen districts in this region, and each district has a minimum of two junior secondary schools. Recruitment of participants for the study was through their PD groups which were organised according to district. This recruitment system would potentially involve all English teachers in the region. There were eight PD facilitators in the region and all were willing to assist in the recruitment process.

Two days after the regional Department of Education office issued its permission letter, I was invited to attend a meeting involving the majority of PD leaders for all education levels. The leaders were planning PD sessions for English teachers, with a focus on using action research to improve teaching. This provided an opportunity to give all eight leaders invitations to participate in my study to distribute to their PD groups along with the letter of explanation about the purposes of the investigation. The PD leaders shared their meeting schedules with me and allowed me to contact them for confirmation. Schedules for the visits were tabulated including information about place, time, number of active members and contact numbers. Based on this information I visited PD groups during their next formal meeting in order to invite individual teachers to participate. I could then distribute the questionnaire for completion by those teachers.

The teachers from the eight PD groups who completed the survey were drawn from both urban and rural areas. This convenience sample of only teachers who attended PD meetings resulted in 107 English teachers completing the survey. Teachers varied in terms of grade level teaching assignment, years of teaching experience, qualification, and gender. For the purpose of confidentiality all PD members involved in the research were assigned a number reference, for example, G.1.1 referred to participant from Group 1 Number 1. Table 2 provides details regarding the numbers and percentages of teachers participating from each of the districts.

Table 2: Participants in quantitative data collection

PD group/location	Ideal number of participants	Survey participants	Involvement rate (%)
G.1/Urban	18	16	88
G.2/Urban	22	14	64
G.3/Urban	17	15	88
G.4/Urban	18	14	77
G.5/Urban	19	7	37
G.6/Rural	21	15	71
G.7/Rural	19	14	73
G.8/Rural	18	11	61
Total	**152**	**107**	**70**

Characteristics of the participants

Table 3 summarises the demographic characteristics of participants including age, gender, qualification level and teaching experience.

Table 3: Basic demographic information of participants: gender, age, education level and teaching experience

Characteristics	Category	Number	% Total N=107
Gender	Female	78	72.9
	Male	29	27.1
Age	Under 23 years	1	0.9
	23-30 years	32	29.9
	31-40 years	30	28.0
	41-50 years	40	37.4
	Over 50 years	4	3.7
Education level	Diploma	2	1.9
	Bachelor	100	93.5
	Master	5	4.7
	Doctor	0	0
Teaching experience	Less than 3 years	15	14.0
	Between 3-10 years	46	43.0
	Between 11-20 years	27	25.2
	Over 20 years	19	17.8

It was notable that the largest group of participants were female teachers. Participants mostly belonged to the 41-51 year age group and the majority of them had attained a Bachelor's degree. Almost half of the participants had been teaching between 3-10 years and only 14% of the total respondents reported having less than three years teaching experience. Table 4 summarises teachers' curriculum level assignment, certification status and assessment training attended.

Table 4: Basic demographic information of participants: curriculum level assignment, certification status and assessment training attended

Characteristics	Category	Sum	% Total N = 107
Curriculum level assignment	Year 7	24	22.4
	Year 8	18	16.8
	Year 9	19	17.8
	All levels	19	17.8
	other	27	25.2
Certification	Certified	32	29.9
	Not certified	75	70.1

Assessment training	Never	6	5.6
	Training at bachelor level	28	26.2
	Training at school	6	5.6
	Training beyond school	4	3.7
	More than one training session	63	58.9

It was notable that almost a quarter of the participants were teaching more than one assignment level. This means that some teachers were teaching at Year 7 level as well as Year 8, or were teaching at Year 7 as well as Year 9, and some were assigned to teach Year 8 and Year 9. The second largest group was teachers teaching at Year 7, while Year 8 was comprised of the least number of participants. The table also shows that the majority of the participants were not certified and had attended several training sessions in assessment.

Instrument

For this phase of the research, a questionnaire was identified and modified to identify teachers' attitudes about assessment. A questionnaire can provide a useful measure of perceptions, values and behavioural intentions (Johnson & Christensen, 2008) and identify different kinds of characteristics within large samples (Stake, 2010) while ensuring anonymity for participants (McMillan & Schumacker, 2010).

The Hong Kong model of TCoA survey

The Teachers Conceptions of Assessment (TCoA) survey developed by Brown (2002, 2003, 2008) to examine New Zealand teacher's conceptions of assessment was adapted for use in this study (Appendix B). Previous research using the survey resulted in the identification of three major teacher conceptions of assessment: accountability, improvement and irrelevance. The original survey has been used in assessment research in other locations and cultures, including Hong Kong (Brown, et al., 2009); Barcelona, Spain (Remesal, 2009) Virginia, USA (Calveric, 2010); Ankara, Turkey (Vardar, 2010); Queensland, Australia (Brown, et al., 2011), and The Netherlands (Segers & Tillema, 2011). The original scale is available in both long and short versions in English. However, participants in Hong Kong and Mainland China revealed different interpretations of TCoA compared to interpretations from elsewhere. This led Brown, Hui, and Yu (2010) to develop and expand the TCoA inventory, adding *examination* as a new dimension. The author's revised version of the TCoA comprised of thirty-one items and was called the Hong Kong model of TCoA. I chose to use the validated Hong Kong model of TCoA

(Brown, et al., 2010) because I anticipated that Indonesian teachers might hold similar conceptions to teachers in Hong Kong and China due to the high-stakes examination setting.

Dr. Gavin Brown granted permission for me to use the Hong Kong TCoA survey in this study, with minor adaptations. Although the original 27-item survey been translated previously into Bahasa, certain new factors such as *examination* were added to the Hong Kong English model that had not been covered in the first Bahasa version. Moreover, although the translated first version of the TCoA was piloted with a number of teachers in Indonesia, the results were not analysed nor published (T. Suratna, personal communication, September 21, 2011). In other words, The Hong Kong model of TCoA comprised more suitable dimensions for assessment in Indonesia compared to the original model. The previous translation also seemed unsuitable for junior secondary school teachers due to the use of unfamiliar terms, which made the translation sound awkward. I communicated with the first translator (Suratna) as I prepared a new translation of the Hong Kong model of the TCoA survey, which was then approved by Brown, and piloted with individuals similar to the intended participants in Indonesia. Feedback from the trial was used to revise the questionnaire into its final form and this was checked again with Brown (G. Brown, personal communication, November 03, 2011).

The Bahasa version of the TCoA consists of three sections in line with the Hong Kong version.

1. The first part asks demographic information about participants' backgrounds (gender, age, years of teaching experience, grade level teaching assignment, qualification, whether they are certified or not), and information about participation in assessment training;

2. The second section is comprised of 31 Likert-type items scored on a six point scale (*1= strongly disagree 2= mostly disagree, 3- slightly agree, 4 = moderately agree, = mostly agree and 6=strongly agree*) which address conceptions of assessment (assessment for improvement, assessment for accountability, and assessment as irrelevant).

3. The third section explains my intention to interview some teachers, and asks them to leave a contact number and name if they are interested in participating in the second qualitative phase of the study.

The second section of the questionnaire consists of several subscales. Improvement subscales comprise of eleven items that evaluate teacher agreement against three dimensions. Five items address assessment and student development and three items measure teachers' responses to the function of assessment to improve students' learning. The remaining three items belong to the dimension of accuracy in assessment. There is only one dimension asking about irrelevance. The third subscale (accountability) contains the most items. There are eighteen items which investigate three dimensions: eight items belong to the examination category, two items address error in assessment and the five remaining items measure the accountability purposes of assessment in terms of illustrating the quality of teacher and school. Table 5 summarises the meta-factors, factors, and examples of items in the TCoA survey.

Table 5: Dimensions of the Teachers Conceptions of Assessment (TCoA) survey

Conceptions of assessment	Dimensions	Items
Improvement	Students development (5 items)	▪ Assessment is used to provoke students to be interested in learning ▪ Assessment cultivates in students a positive attitude towards life
	Help learning (3 items)	▪ Assessment helps students improve their learning ▪ Assessment determines if students meet qualification standards
	Accuracy (3 items)	▪ Assessment results are trustworthy ▪ Assessment results can be depended on
Irrelevance	(5 items)	▪ Assessment has little impact on teaching ▪ Assessment forces teachers to teach in a way that is contrary to their beliefs
Accountability	Examination (8 items)	▪ Assessment prepares students for examination

		▪ Assessment selects students for future education or employment opportunities
	Error (2 items)	▪ Teachers should take into account error and imprecision in all assessment ▪ Assessment results should be treated cautiously because of measurement error
	teacher and school control (5 items)	▪ Assessment results contribute to teachers' appraisal ▪ Assessment is an indicator of a school quality

Data collection

Pilot study

A pilot trial was carried out with an independent sample (not otherwise involved in this study) to identify any issues that might require modification prior to distribution of the survey to the research participants. (D. K. Cohen & Hill, 2000; L. Cohen, Manion, & Morrison, 2007). Twelve teachers with similar characteristics to the intended participants participated in the pilot. All were teachers of subjects other than English from one junior secondary school in Gowa region, South Sulawesi province, Indonesia. The pilot survey was conducted to measure the clarity of questionnaire items and to determine the approximate time required to complete the survey. Respondents were also asked to provide feedback on the content, format and lay-out of the questionnaire. I emailed the draft questionnaire to a close colleague who had appropriate expertise in research as she was in the last year of her master's degree. This colleague sent me back the pilot results with the feedback. Respondents of the survey reported that the questionnaire was comprehensive but two items were unclear or ambiguous. Some others suggested increasing the font size used in the questionnaire. In response to their comments and feedback, I made a few minor changes to the wording of items, increased the font size and refined the layout. Thus, the pilot testing process enabled me to revise and refine the questionnaire (McMillan & Schumacker, 2010) and provided reassurance that the questionnaire was understandable and suitable for the context.

Administration of questionnaires

I personally distributed the questionnaire to teachers attending eight English PD groups, mostly at the end of the session. A second visit was required to two of the

PD groups in order to distribute surveys to several PD members who were not present at the meeting when I had arranged to administer the survey. The distribution of the questionnaire took place between the fourth week of October 2011 and the first week of February 2012 (see Table 6).

My presence at the PD venue while respondents completed the questionnaire was useful as this enabled me to answer questions raised by participants and I was able to check that surveys were completed so as to avoid missing data wherever possible. Table 6 presents details of the questionnaire distribution to all PD groups both in urban and rural areas.

Table 6: Details of questionnaire administration

Date	PD group/location	Note
24 Oct 2011	G.1/Urban	
25 Oct 2011	G.2/Urban	
21 Nov 2011	G.3/Urban	
29 Nov 2011	G.4/Urban	First visit
02 Jan 2012		Second visit
02 Dec 2011	G.5/Urban	
03 Dec 2011	G.6/Rural	First visit
04 Jan 2011		Second visit
24 Dec 2011	G.7/Rural	
04 Feb 2012	G.8/Rural	

Data Analysis

Data obtained from the questionnaire were entered into an Excel spread sheet and analysed using SPSS version 19. The Excel analysis provided a plotting of each participant's conceptions as measured by the survey, thereby allowing me to identify participants with unique response patterns for the second stage interviews. The selection for interview participants was based on predominant preferences or patterns aligned with each of the particular conceptions of assessment as these were the case units for the study. Examining participant patterns also allowed me to identify those with a jagged or inconsistent profile of responses.

Using SPSS version 19, descriptive statistics were employed to analyse respondents' demographic information. Percentages and frequencies were calculated to ascertain the extent to which respondents believed in the function of assessment to improve learning, the extent of their confidence in assessment being

irrelevant, and the extent that they believed in the accountability purposes of assessment. Cronbach's Alpha was conducted to determine the internal consistency and reliability of the questionnaire. In addition, the validity of the questionnaire was also tested by executing exploratory factor analysis.

Phase two: Qualitative

The case study approach

Yin (2003b) defines case study as an empirical enquiry that "investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" p.13). This study explores the overarching phenomenon of assessment conceptions held by Indonesian teachers of English. As case study particularly focused on the unique characteristics of each case (Stake, 2000) this design provided a strategy for examining how or why teachers understand assessment and how it was influenced by their particular teaching context.

In order to explore this phenomenon thoroughly, I decided to "have sub-groups of cases covering each type" (Yin, 2009, p. 59). The responses from participants in this study clearly identified two groups of patterns that were predominantly either improvement or accountability. The responses from the third group were distinctive in that they revealed a jagged response. Like the participants in the other groups, some favoured both improvement and accountability, but unlike the other participants, teachers in this group rated conceptions of irrelevance over one or more of the other conceptions. Therefore teachers with these mixed conceptions can be categorised as seeing assessment as largely irrelevant despite acknowledging the potential for improvement and accountability. Thus, there are three different cases in the study: improvement, irrelevance and accountability. Each case comprises a subgroup or a cluster of four individuals that hold one of these three main conceptions of assessment. Altogether the three cases are subunits of the larger phenomenon of conceptions of assessment (Yin, 2009). This multiple-embedded design (Yin, 2003a), enabled me to understand the complex phenomenon of assessment conceptions from the real-life situation of the participants and also allowed identification of any ambiguous boundaries that might exist among the three types of assessment conceptions (Merriam, 1988).

Yin (2009) argues that this strategy enables clearer identification of case and is robust and compelling. In addition, it is considered to ensure high authenticity and transferability of the findings (Grunbaum, 2007). The twelve key cases (four within each case) are believed to produce replication and best explanation of the phenomena under investigation, which is the main purpose of multiple case studies. By looking at three different groups or the three assessment elements separately, I was able to compare each type of conception in order to gain the integrity, the wholeness of assessment conceptions (Thomas, 2011).

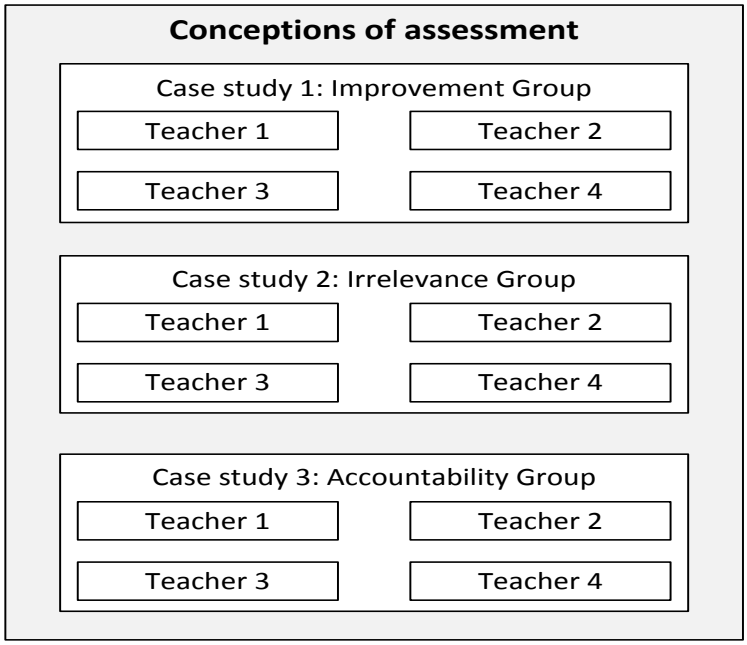


Figure 8: The embedded units within the three clusters (adapted from Thomas, 2011; Yin, 2009)

This multiple embedded case study did not aim to generalise the findings or seeking a desirable outcome. This study is designed for transferability through thick description that respects the particularity of each case (Mertens, 2005). This case study gathered substantial data from different sources including semi-structured interviews and document analysis to enable data triangulation, which is important in case study analysis (Creswell, 2005). Triangulation refers to "checking the validity of an interpretation based on a single source of data by recourse to at least one further source that is of a strategically different type" (Hammersley, 2008, p. 23). The combination of two or more viewpoints is described as the right incorporation of jigsaw puzzle pieces in order to provide the full image of a certain object (Erzberger & Kelle, 2003). Therefore, the use of interview and documents as data collection

instruments aimed to reduce the possibility of drawing false conclusions (Hammersley, 2008).

Participants

The participants' willingness to share their contact details on the last page of the questionnaire enabled me to contact them via phone. The participants in the qualitative component consisted of twelve Indonesian teachers of English purposively selected based on the survey results. There were three groups of teachers representing each conception, and each cluster included four teachers. This is a sufficient sample size in a qualitative research as the focus is not upon generalisation of findings but elaboration of ideas and explanations for those findings (Creswell, 2002; Creswell, et al., 2008). Sample size in qualitative research is dictated therefore by saturation, or the point at which participants are no longer introducing new ideas or explanations but seem to be recycling information already revealed in the data. Participants were provided with an interview invitation and information sheet (Appendix C) as well as a consent form to sign provided in Bahasa Indonesia. A copy of the consent to participate in the interview is included in Appendix D. Table 7 provides a list of participants and their cluster membership.

Table 7: Details of interview participants

No	Group	Pseudonyms	Clusters
1.	G.4.6	Intan	Improvement
2.	G.6.9	Emma	Improvement
3.	G.5.6	Andin	Improvement
4.	G.1.10	Lisa	Improvement
5.	G.3.1	Eva	Irrelevance
6.	G.2.8	Akbar	Irrelevance
7.	G.4.12	Ira	Irrelevance
8.	G.8.11	Rahmat	Irrelevance
9.	G.2.4	Santi	Accountability
10.	G.2.6	Putri	Accountability
11.	G.1.1	Naya	Accountability
12.	G.7.12	Angga	Accountability

The semi-structured interview

An interview is 'a data-collection method in which an interviewer asks questions of an interviewee' (Johnson & Christensen, 2008, p. 203). The main purpose of using

an interview is ‘to verify and extend information obtained from other sources’ (McMillan & Schumacker, 2010, p. 355) such as questionnaires and/or documents. Accordingly the interview process assisted me to gain a deeper understanding of teachers’ conceptions of assessment and to locate factors contributing to their conceptions and practices of assessment as well as their beliefs in the role of assessment in student learning.

The in-depth semi-structured interviews used in this study enabled me to set topics and issues to guide questioning in outline form but did not compel me to use exactly the same words or sequence for each one (Appendix E). This interview feature aims to increase the comprehensiveness of data and ensures that the data collection is systematic (Johnson & Christensen, 2008). The interview protocol included fifteen open-ended questions that had been pilot-tested. Debriefing with the pilot participants who did not participate in the formal interviews that are part of the study provided information on the clarity and their relevance of the interview questions and enabled me to avoid unnecessary redundancy.

The interview questions were prepared in both English and Bahasa and participants chose which language they wished to use in the interview. All interviews were conducted in Bahasa, the language preferred by participants. The questions were designed to elicit the participants’ values of assessment, factors influencing their conceptions, their assessment practices, and their conceptions about accuracy of assessment. I used prompts to probe for clarity and in-depth information. The interviews were conducted face-to-face and lasted between 45 to 60 minutes for each of the twelve participants. Table 8 provides values of assessment measured in the interview along with some examples.

Table 8: Samples of the interview questions

Aspects to measure	Examples
Values and influence	In your current teaching, what is the purpose of conducting assessment?
	What in your experience led you to this understanding of assessment?
Practice of assessment	How do you assess your students? /What methods do you use in assessing your students?
Impact of assessment	What is the impact of assessment on your student learning/on your teaching/ on your school accountability?
Accuracy in assessment	To what extent do you perceive that assessment results provide an accurate measure of students' performance?

Documents

Documents or artefacts describe people's experience, knowledge, actions and values (McMillan & Schumacker, 2010). Document analysis was undertaken to supplement the information obtained during and after the interviews. After capturing teachers' conceptions of assessment using the questionnaires, I also sought teachers' permission to view their assessment documents. These documents covered teachers' lesson plans, assessment files, student tasks, tests set by the teacher, teacher assessment records and copies of student report books. These documents illustrated some of the functions and values of assessment in the participants' schools relating to the participants' assessment practice. Moreover, these documents also helped me to clarify items from the questionnaire (McMillan & Schumacker, 2010). As a result these materials were used not only as prompts for the interview to delve more deeply into teachers' conceptions of assessment but also as information on how teachers' conceptions were reflected in their teaching and learning. Samples of all documents are included in Appendix F.

Data collection process

Interview process

The interviews were conducted between January and March 2012 (see Table 9). Some interviews took place while I was waiting to conduct the last survey with a PD group located in rural area. All interviews were carried out at times and a venues convenient to the participants and each interview was conducted on a different day. No other person was present during the interviews and all interviews were audio-

recorded. Field notes were also taken during the interviews to highlight some important points made by participants. Once an interview was transcribed, the transcript was returned to each participant for member checking which allowed them an opportunity to read through their transcriptions before and after I translated them into English. Member-checking was conducted to avoid any misunderstanding or misinterpretation of the participants' words (Mertens, 2005). All participants accepted and agreed to the transcription and did not request any changes. Member checking was also conducted with a fluent Bahasa-English speaker to check the meaning of the quotes used in findings. A summary of the interview process is listed in Table 9.

Table 9: Details of interview process

No	Pseudonyms	Date	Date of member check
1.	Intan	30 Jan 2012	29 Feb 2012
2.	Ira	31 Jan 2-12	18 Feb 2012
3.	Naya	06 Feb 2012	13 Feb 2012
4.	Andin	07 Feb 2012	26 Feb 2012
5.	Angga	15 Feb 2012	24 Feb 2012
6.	Eva	20 Feb 2012	08 Mar 2012
7.	Emma	21 Feb 2012	01 Mar 2012
8.	Akbar	22 Feb 2012	28 Feb 2012
9.	Putri	25 Feb 2012	28 Mar 2012
10.	Lisa	01 Mar 2012	10 Mar 2012
11.	Rahmat	08 Mar 2012	17 Mar 2012
12.	Santi	28 Mar 2012	30 Mar 2012

Qualitative Analysis

Interview data

The semi-structured interviews were transcribed verbatim by the researcher using Word processing software. Qualitative analysis involved both manual thematic analysis and NVivo 8. Manual analysis was dominant due to the use of Bahasa in the interview. In addition, analysing transcripts in the original language was more authentic because it enabled me to capture the key ideas expressed by the participants in their own words.

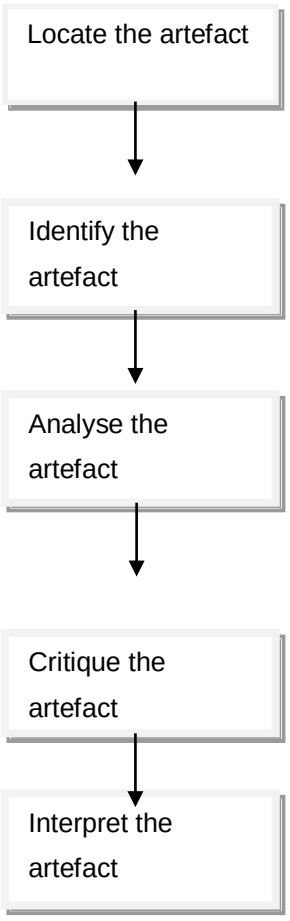
The steps in the qualitative analysis included: (1) Listening to the audio-file, (2) Reading the transcript and checking it with the participants, (3) Coding the data by

segmenting and labelling the text according to the main ideas, (3) Merging the codes together by sorting, cutting and pasting, (4) Naming themes, (5) Checking the themes and sub-themes, (6) Renaming themes, (7) Writing the report, and (8) Renaming themes (Braun & Clarke, 2006; Creswell, 2003). Reflections on my findings led me to repeat the steps of this analysis in order to find categories underlying each theme so that I was able to rename my themes more accurately. NVivo 8 was used particularly to track the page numbers of each quote used in findings.

Analysing my study within the multiple-embedded case design allowed me to elaborate findings within a cluster or family of phenomena and offered me the opportunity to observe similarities and differences in each case study in order to produce new knowledge (Khan & VanWynsberghe, 2008). I was interested to ascertain whether these three nested units or the sub-classes of the major phenomenon (conceptions of assessment) did in fact "share certain patterns or configurations" (2008, p. 5). Essential elements or themes were bracketed case by case (Denzin, 1989). This process allowed me to understand how and why conceptions arose in a particular context and at the same time allowed me to make comparisons across cases. In other words, no second layer analysis was needed for the study because it would have stated the same information without the construction of any new knowledge. This means that a discussion of the larger phenomenon immediately follows the case findings.

Document analysis

Document analysis involved five strategies suggested by McMillan and Schumacher (2010). This strategy allowed me to study the characteristics of the persons who prepared the documents, as well as to note the teachers' judgements of their students. Figure 9 illustrates the steps for analysis.



- Inform and ask participants to share their assessment documents and assessment records

- Note the category of the artefacts
- Give brief descriptions of the artefacts

- Describe the purposes of the artefacts
- Who uses it
- How, where and what is the purpose of its use

- Determine the authenticity and accuracy of the artefacts in relation to identifying the meaning in the social settings

- Produce subtle meanings about the artefacts based on the context and other data

Figure 9: Strategies for analysis and interpretation of documents (adapted from McMillan & Schumacher, 2010, p.362).

Validity and reliability of the mixed methods design

There are two important ways to evaluate quality in qualitative research; these include credibility and trustworthiness (Mertens, 2005), whereas in quantitative research the factors that are relevant are generally referred to as reliability, validity, and objectivity. As this study is a sequential mixed methods study, issues connected with reliability, validity and objectivity of the quantitative data collection will be examined before examining parallel issues in qualitative design.

Quantitative issues

Reliability

Reliability refers to the consistency of the instrument to present a test score free from measurement error (Muijs, 2011). The issue of measurement error can arise from the participants (such as their motivation and alertness to participate), from the

administration process (such as flexibility in time for questionnaire completion), or from the instrument used for the study, for example changes in the items of the questionnaire (Mertens, 2005). In order to address this issue, I conducted descriptive statistics of means and standard deviations for the instrument. Cronbach's coefficient Alpha statistic was also carried out. Muijs (2011) and as Singh (2007) suggests that an alpha of 0.70 or more is an accepted reliability level, overall the TCoA items scored within a reliable range. Table 10 illustrates the internal consistency of items in the questionnaire.

Table 10: Reliability of the Hong Kong TCoA questionnaire

Scale	Item	Total Correlation	If Item Deleted
Improvement	9	.590	.832
α = .85	17	.696	.822
	10	.665	.829
	13	.637	.829
	11	.646	.827
	1	.437	.843
	2	.316	.852
	3	.390	.846
	29	.502	.839
	8	.530	.838
	4	.522	.837
Irrelevance	12	.535	.434
α = .60	18	.319	.546
	15	.480	.453
	7	.153	.626
	27	.299	.573
Accountability	23	.628	.833
α = .85	31	.355	.849
	19	.479	.842
	22	.419	.845
	5	.489	.841
	26	.633	.834
	14	.439	.844
	24	.478	.842
	21	.297	.851
	28	.338	.849
	25	.502	.841
	20	.546	.838
	30	.651	.832
	16	.624	.833
	6	.367	.849

The eleven survey items that focused on teachers' beliefs that assessment is effective in improving student learning had good internal consistency, α = .85. Similarly robust results were also recorded for the subscale measuring the accountability purposes of assessment: the fifteen items in that scale scoring optimal internal consistency with α=.85 where all items alpha values were above

0.3. Thus, the two subscales measuring teachers' conceptions about improvement and accountability purposes of assessment had overall high reliability.

In contrast, the irrelevance section only received α=.60 meaning that this section had low reliability (Singh, 2007). However this does not necessarily mean that these items were not reliable (Muijs, 2011). The low alpha value for irrelevance sections might be influenced by the small number of items used in the section. However, low reliability may also signal that these items were not as well understood by participants which would indicate a measurement problem or even signal a validity problem. It is also possible that misunderstandings occurred because this concept was not originally formulated in the participants' language (Bahasa Indonesia) nor was it adjusted for an Indonesian setting. This means that several important components like the culture and education system of the setting are not captured effectively by the survey.

Validity

Although the Hong Kong version of TCoA has been internationally validated, it was necessary to check the validity of conducting the questionnaire for use in a different country and culture other than those where the survey had been used previously. The aim of this validity test was to look at the internal structure of the TCoA and the extent to which it measured what it was intended to measure (Muijs, 2011). To determine validity, I conducted a factor analysis to reduce data according to its variables, detected the relationship between variables (Singh, 2007), and sought evidence for construct validity (Mertens, 2005). This process involved principal component analysis particularly exploratory factor analysis (EFA). Testing the validity of a translated questionnaire through factor analysis is commonly used to check for consistency with the original version (Isaksson, 2013; Ma, Hwang, & Chen-Sea, 2005) particularly when the translated version is first used. There were two criteria to extract factors for this analysis, namely the Eigen value and the Kaiser criterion.

Using the Kaiser Criterion, I retained the factors with Eigenvalues greater than 1 (Singh, 2007). The initial factor statistics of my data revealed that that there were nine factors having values higher than 1. These nine final factors explain 67.3% of total variance. The result of this variance was dissimilar to the original validity test in which there were only seven factors extracted from the questionnaire with Chinese

participants. To confirm the factors and look closely at the observed variables and the underlying latent variables that might exist, I extracted the same seven factors as those used in the Hong Kong version of TCoA. In so doing, I did factor rotation to improve the interpretability of the solution using direct oblimin to easily group and interpret the factors (Tabachnick & Fidell, 2007). The results of the analysis are shown in Table 11.

Table 11: Factor rotation with 7 components

Items	Component						
	1	2	3	4	5	6	7
16 : measures school quality	**.687**						.345
25 : teacher appraisal	**.679**						
29 : trustworthy	**.653**						
8 : dependent	**.651**						
30 : indicate school quality	**.609**					-.222	
23 : good score in exams	**.572**			.296		-.212	
20 : indicate good teacher	**.552**	.248				-.230	
22 : sets class schedule	**.540**	-.312	.207		-.218		.274
17 : foster characters	**.429**		-.208		.225	-.281	
1 : improve learning	**.241**	-.204				-.206	
12 : filed and ignored		**.768**					
18 : interfered with teaching		**.700**				-.262	
15 : imprecise process		**.632**				.285	
7 : against belief	.260	**.520**	.269				
21 : error and imprecision			**.809**				
28 : treated cautiously			**.733**				
24 : select future education			**.396**	.291			-.222
19 : teaches exam technique				**.830**			
5 : prepare for exam				**.736**			.278
31 : familiarise exam format				**.628**			-.361
9 : succeed in real world					**.516**		
26 : avoid failure in exam	.404			**.496**			
2 :meet qualification standard			.235		**.726**		
4 : sufficiently accurate					**.690**		-.261
3 : modifies teaching					**.447**		
14 : assign grade or level	**.363**					-.278	
27 : little impact			-.338			**.817**	
11 : cultivate positive attitudes				**.249**		-.671	
13 : stimulate thinking	.210				**.262**	-.584	
10 : provoke students' learning	**.281**		-.272	.210	.221	-.442	
31 : police teachers			.246		.220		**.680**

The original Chinese version of TCoA developed by Brown and colleagues (2010) had both first and second order factors, whereas the Indonesian version of the TCoA scale revealed an even more complicated structure. Several items overlapped with two or more components suggesting that they measured more than one construct. As the original version of the questionnaire used three major factors,

I conducted another extraction. The results of this extraction are simpler but still illustrate some overlapping constructs (see Table 12).

Table 12: Factor rotation with 3 components

Items	Component		
	1	2	3
10 : provoke students' learning	.830		
17 : foster characters	.802		
11 : cultivate positive attitudes	.755		
14 : assign grade or level	.699		
23 : good score in exams	.699	.235	
13 : stimulate thinking	.679		
8 : dependent	.638		
30 : indicate school quality	.632		
25 : teacher appraisal	.606		
9 : succeed in real world	.587		
29 : trustworthy	.585		
4 : sufficiently accurate	.542		
5 : prepare for exam	.533	.220	
20 : indicate good teacher	.522	.237	
16 : measures school quality	.485		.349
1 : improve learning	.484	-.206	
26 : avoid failure in exam	.469		.353
19 : teaches exam technique	.454		
3 : modifies teaching	.431		
31 : familiarise exam format	.400		
12 : filed and ignored		.721	
15 : imprecise process		.636	
18 : interfered with teaching		.599	
7 : against belief		.569	
27 : little impact		.328	-.259
2 : meet qualification standard	.226	-.229	
21 : error and imprecision			.825
28 : treated cautiously			.744
31 : police teachers			.474
22 : sets class schedule			.468
24 : select future education	.359		.373

This analysis suggests that different constructs should be used to measure Indonesian teachers' conceptions of assessment. Indeed, the Chinese version of the questionnaire was developed for a Chinese context that might signify different relationships among factors in the questionnaires. Cultural differences between Hong Kong and Indonesia might actually result in more or less factors. Thus although Indonesia, Hong Kong, and China share a similar examination culture, different languages are spoken and different policies implemented that might influence their teachers' conceptions of assessment. Moreover, the use of the six-point rating scale on the questionnaire might have confused Indonesian teachers because they were more accustomed to a five point scale. Furthermore some of the

descriptors of frequency used on the Likert scale did not have equivalent words in Bahasa.

Objectivity

Objectivity is determined by the disinterest of the person who administers, scores and interprets a test (Mertens, 2005). In this case I administered the survey which allowed me to explain any unclear or ambiguous items in the questionnaire. Soon after the data collection was completed with one group, I stored and managed it using an Excel spread sheet which was later uploaded into SPSS. In short as the sole researcher, I was responsible for the distribution, management and analysis of all data.

Qualitative issues

To guarantee the trustworthiness of the qualitative phase, the factors to consider are credibility, transferability, dependability and confirmability of findings (Guba, 1981; Guba & Lincoln, 1989). In this study, seven primary verification procedures were used to determine trustworthiness.

Credibility

Credibility is the ability of the study to measure what it is intended to measure (Shenton, 2004). This is similar to validity in quantitative studies. To address this issue, two main strategies were used. The first was triangulation, in which I collected data to investigate a phenomenon from different sources to see if the findings would be aligned across sources (Mertens, 2005), through interviews, field notes and documents. Another strategy I used was member-checking which gave participants the opportunity to confirm or challenge the transcriptions of what they had said.

Transferability

Transferability means the extent to which the findings can be transferred to another context (Guba, 1981; Merriam, 2001). Two main strategies were conducted to address this. The first was to provide thick descriptions that enabled me to present my participants' demographic information to allow readers to develop a proper understanding of the phenomenon (Shenton, 2004). Thick description enables readers to make comparison to other situations or contexts with similar characteristics (Guba, 1981). I also compared my findings to previous, related

studies carried out in different contexts (reviewed in chapter two and discussed with reference to my findings in the discussion chapter). Another strategy applied was using purposive sampling. This means that participants were selected based on their responses in phase 1; these participants were believed to have experience of the central phenomenon (Creswell, 2007).

Dependability

Dependability indicates the stability of results over time (Guba & Lincoln, 1989). The issue is addressed through a clear explanation of the methods used. In this chapter, I have presented a thorough justification of the methods and their effectiveness. I have fully discussed the research design, the data gathering process and the process of analysis, also reflective appraisal of the project (Shenton, 2004). In addition, I also conducted stepwise replication by presenting parts of this qualitative study in international conferences in Semarang, Jakarta, Indonesia and in New Delhi, India. A part of this study was also published in an international journal (Azis, 2012).

Confirmability

Confirmability means the consistency of data and its interpretation (Guba, 1981). This can addressed through triangulation and by conducting a confirmability audit (Guba & Lincoln, 1989). In the latter strategy, I asked peers to review field notes and interview transcripts. I also was careful to keep research journals.

Ethics

In order to obtain permission to conduct research In Indonesia, I applied to the Victoria University Human Ethics Committee. The ethics approval /2011/70: RM 18851 was received on 12th October 2011 (Appendix G).

I used the letter granting ethical approval from the university to apply for another permit from the research site. This process involved three sets of permission. Firstly permission was sought from the Research and Development Department of South Sulawesi province. Their letter of permission and recommendation was issued on 14th October 11 (Appendix H). As the research was conducted in one particular region in the province, I took the recommendation letter from the South Sulawesi province to Gowa region. The Unity Section of the Gowa region published

another recommendation letter addressed to the Education Department of the region dated 17th October 2011 (Appendix I). The following day, the Education Department of Gowa region issued a letter of permission, which allowed me to conduct research in the region (Appendix J).

In addition, I adhered to the four guidelines suggested by Christians (2005) for conducting research. These are gaining informed consent, avoiding deception, protecting participants' privacy and the confidentiality of the data, and ensuring accuracy of the data. An information sheet explained the purposes of the research project, the nature and consequences of the research, the duration of the study and their rights as participants including the information that they could withdraw from the research without having to give any reasons. Participants were assured that confidentiality would be maintained in any report of findings. They were also assured that the research process and its findings would have no negative impact on them. The information letter and consent form was prepared both in English and Bahasa.

Summary

This chapter presented the methodology of the study. The study used a mixed methods sequential explanatory design where the quantitative phase proceeded the qualitative phase. The participant selection model utilised for the study was weighted towards the qualitative strand. Purposive sampling was implemented in both phases of the research. The Hong Kong model of TCoA survey, a semi-structured interview and document analysis were the main instruments used for data collection. Data from the qualitative phase were analysed using SPSS version 18 and thematic analysis was used to analyse the qualitative data. Issues relating to reliability, validity and ethics were considered.

Chapter Four

Findings

Introduction

This chapter presents findings from the two phases of the study. Findings from the quantitative questionnaire are described first, in accordance with a sequential mixed methods design. The quantitative data is analysed with descriptive statistics. In the qualitative phase, data from interviews and documents are analysed using thematic analysis (Braun & Clarke, 2006). The integration between quantitative data and the qualitative data is presented at the end of the quantitative findings. This integration technique identifies the link between quantitative and qualitative data.

Quantitative findings

Conceptions about assessment to improve teaching and learning

The participants were asked their views about assessment to improve teaching and learning in the TCoA survey. There were three subscales under this conception: assessment which is conducted for student development, assessment to help learning and accuracy in assessment. All items were rated on a six point Likert scale according to participants' degree of agreement. The scale used two negative (strongly disagree and mostly disagree) and four positive trends (slightly agree, moderately agree, mostly agree and strongly agree). The first two degrees are categorised as disagreement and the remaining four are grouped as agreement. The results are presented in Table 13 and Figure 10.

Table 13: Improvement conception

No.	Items	N	Degree of agreement %						Mean	SD
			1	2	3	4	5	6		
9.	Assessment helps students succeed in authentic/real world experience	107	5.6	3.7	14.0	23.4	20.6	32.7	4.48	1.45
17.	Assessment fosters students' character	107	2.8	1.9	7.5	18.7	29.9	39.3	4.89	1.22
10.	Assessment is used to provoke students to be interested in learning	107	.9	0	1.9	13.1	11.2	72.9	5.52	.90
13.	Assessment stimulates students to think	107	0	1.9	5.6	20.6	29.9	42.1	5.05	1.01
11.	Assessment cultivates students positive attitudes towards life	107	.9	1.9	4.7	25.2	16.8	50.5	5.07	1.13
1.	Assessment helps students improve their learning	107	.9	0	3.7	12.1	7.5	75.7	5.52	.95
2.	Assessment determines if students meet qualification standards	107	.9	1.9	1.9	19.6	18.7	57.0	5.24	1.05
3.	Assessment information helps modify on-going teaching of students	107	.9	0	0	5.6	16.8	76.6	5.56	.72
29.	Assessment results are trustworthy	107	1.9	5.6	11.2	30.8	26.2	24.3	4.47	1.23
8.	Assessment results can be depended on	107	2.8	9.3	12.1	29.0	15.0	31.8	4.39	1.42
4.	Assessment results are sufficiently accurate	107	.9	4.7	8.4	29.9	30.8	25.2	4.61	1.14

Table 13 shows that item 3 received the highest agreement followed by item 1 and 10 all with mean scores of above 5.50 suggesting that teachers believed in assessment as a tool to improve student learning. Lower agreement was addressed to all items asking about accuracy in assessment (item 29, 8 and 4). Hence, these items received mean scores above 4.30. Overall, the mean scores ranged from 4.39 – 5.56 as shown in Figure 10 and Table 13, which signifies that generally participants considered the use of assessment is to improve learning. They particularly believed that it is an accurate way to help develop teachers' teaching and improve students' learning. This suggests that participants supported the improvement conception of assessment.

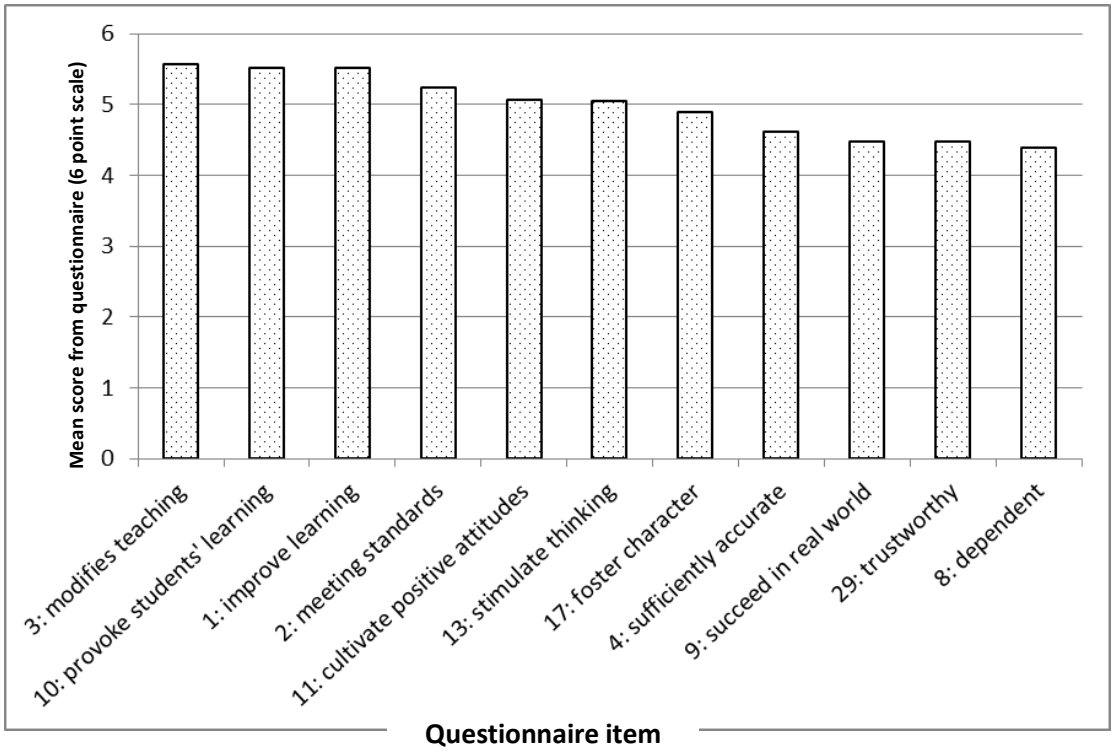


Figure 10: Mean score of improvement conception-related items

Conceptions about the irrelevancy of assessment

Findings from the survey on the irrelevance conceptions of assessment indicated a disagreement toward statements which counted assessment as irrelevant. Table 14 and Figure 11 present teachers' level of agreement toward this type of conception.

Table 14: Irrelevance conception

No.	Items	N	Degree of agreement (%)						Mean	SD
			1	2	3	4	5	6		
12.	Assessment results are filed and ignored	107	65.4	18.7	4.7	7.5	2.8	.9	1.66	1.14
18.	Assessments interfere with teaching	107	79.4	12.1	1.9	.9	4.7	.9	1.42	1.04
15.	Assessment is an imprecise process	107	47.7	29.0	11.2	5.6	3.7	2.8	1.97	1.27
7.	Assessment forces teachers to teach in ways against their belief	107	31.8	22.4	13.1	16.8	9.3	6.5	2.69	1.59
27.	Assessment has little impact on teaching	107	47.7	28.0	12.1	3.7	8.4	0	1.97	1.23

According to Table 14 and Figure 11 item 7 received the highest mean score of 2.69. Four remaining items received an average agreement of 1.41 which according

to the Likert scale used in this study fell between strongly disagree and mostly disagree. This indicates that participants in this study do not feel strongly that assessment is irrelevant and are likely to consider assessment as an important part of their teaching role, particularly to inform the quality of teaching and learning.

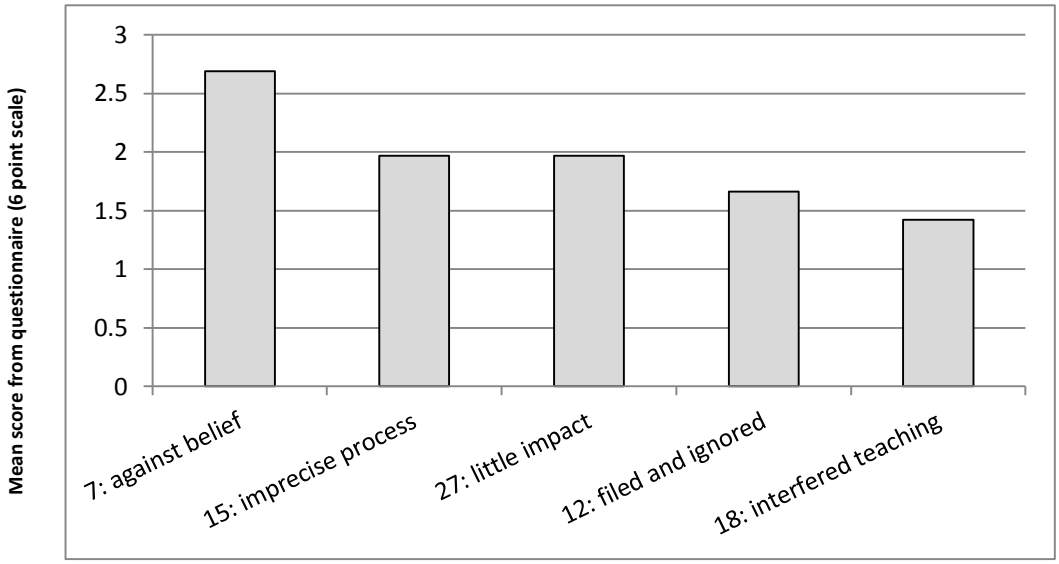


Figure 11: Mean score of irrelevance conceptions-related items

Conceptions about accountability in assessment

In the questionnaire, the participants were also asked to rate their level of agreement toward assessment for accountability purposes. Three subscales were used: examination, error, and teacher and school control. Participants' confidence toward accountability conception related items are listed in Table 15 and Figure 12.

Table 15: Accountability in assessment

No.	Items	N	Degree of agreement (%)						Mean	SD
			1	2	3	4	5	6		
23.	Assessment helps students gain good scores in examinations	107	3.7	2.8	7.5	22.4	22.4	41.1	4.80	1.32
31.	Assessment familiarises students with examination formats	107	4.7	0	9.3	30.8	25.2	29.9	4.62	1.26
19.	Assessment teaches examination-taking techniques	107	3.7	7.5	18.7	30.8	21.5	17.8	4.12	1.32
22.	Assessment sets the schedule or timetable for classes	107	.9	1.9	2.8	21.5	36.4	36.4	5.00	1.01

5.	Assessment prepares students for examinations	107	1.9	7.5	9.3	21.5	26.2	33.6	4.64	1.33
26.	Assessment helps students avoid failures in examinations	107	2.8	.9	7.5	36.4	25.2	27.1	4.62	1.15
14.	Assessment is assigning a grade or level to students work	107	.9	1.9	2.8	13.1	22.4	58.9	5.31	1.03
24.	Assessment selects students for future education or employment opportunities	107	.9	4.7	7.5	19.6	26.2	41.1	4.89	1.21
21.	Teachers should take into account error and imprecision in all assessment	107	0	2.8	7.5	26.2	19.6	43.9	4.94	1.12
28.	Assessment results should be treated cautiously because of measurement error	107	.9	4.7	5.6	29.9	23.4	35.5	4.77	1.18
25.	Assessment results contribute to teachers' appraisals	107	1.9	.9	2.8	19.6	34.6	40.2	5.05	1.05
20.	Assessment indicates a good teacher is	107	4.7	10.3	6.5	27.1	21.5	29.9	4.40	1.47
30.	Assessment is an accurate indicator of a school quality	107	3.7	5.6	10.3	28.0	26.2	26.2	4.46	1.33
16.	Assessment measures the worth or quality of schools	107	5.6	17.8	15.0	25.2	17.8	18.7	3.88	1.51
6.	Assessment is used by school leaders to police what teachers do	107	4.7	6.5	15.0	22.4	20.6	30.8	4.40	1.45

Table 15 and Figure 12 show that participants believed in the value of accountability in assessment. The strongest agreement was given to items 14, 25 and 22, all with mean scores of above 5.00. The lowest agreement was given to item 16 with the mean score lower than 4.00 but this score is still regarded as a moderate agreement. The remaining 11 items received strong agreement with the mean score of above 5.00. More specifically, the error in assessment sub scales (items 21 & 28) received the highest agreement with the average mean score of 4.85. This score is slightly (.1) higher than the agreement given to examination sub scales. The lowest group, still rating as receiving good agreement, was the teacher and school control sub scale (items 25, 20, 30, 16 and 6). This group received a mean score of 4.43. It can be therefore concluded that English teachers in the region of Gowa confidently supported the purposes of assessment to describe the accountability of teacher, students, and school. They also seemed to recognise the importance of examinations in the teaching and learning process.

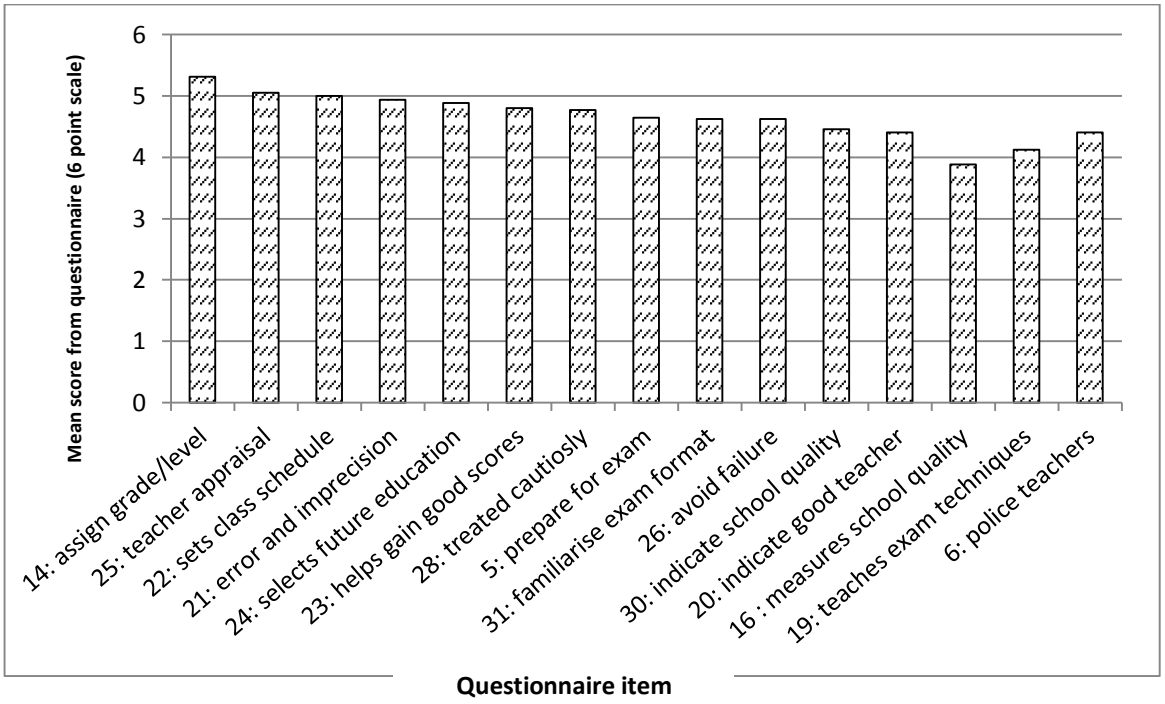


Figure 12: Mean score of accountability conception-related items

Overall result of teachers' conceptions of assessment

Findings regarding overall results of participant responses to the TCoA questionnaire can be seen in the following table.

Table 16: Agreement level of teachers for components in TCoA, N=107

Components	Mean	SD
Improvement	4.99	1.09
Irrelevance	1.94	1.27
Accountability	4.66	1.25

Table 16 and Figure 13 reveal that assessment for improvement received the highest agreement (M=4.99) followed by the accountability conception (M=4.66). The two conceptions were not widely different and participants tended to answer *mostly agree*. On the other hand, the irrelevance conception gained the lowest response (M=1.94) and participants generally chose *mostly disagree*. All subgroups indicated similar standard deviation above 1 %. The irrelevance subgroup received the highest score (SD=1.27), while accountability received 1.25 and improvement 1.09 respectively. This suggests a minimal variation in comparison to the other two subgroups.

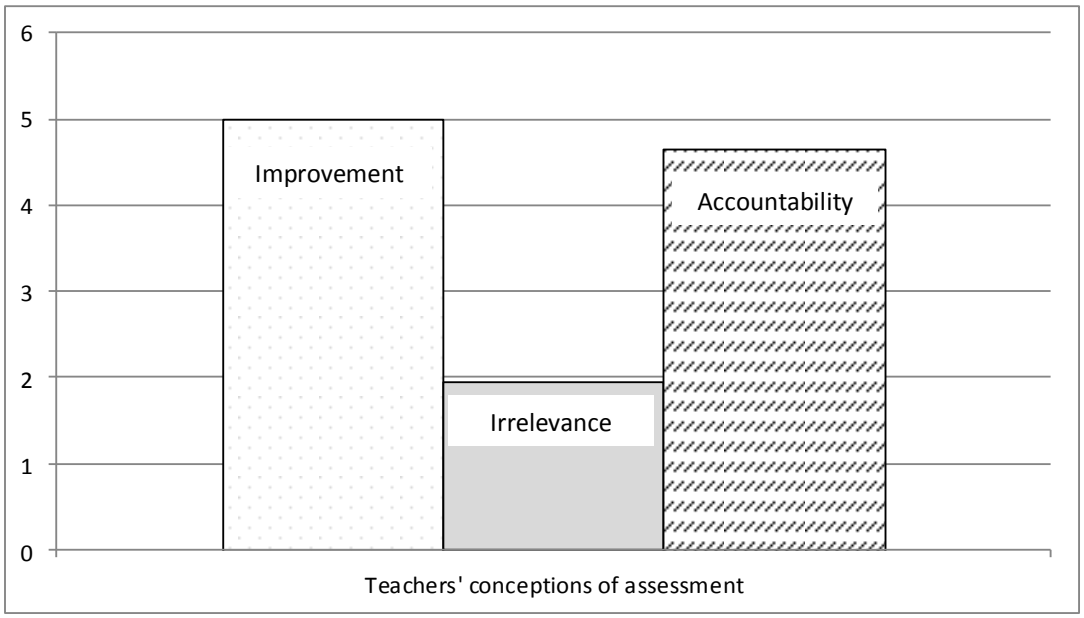


Figure 13: Overall results of teachers' conceptions of assessment

Table 16 and Figure 13 also show that participants in the study believed in conducting assessment for improvement purposes as well as to describe accountability of students and school. Teachers also revealed that assessment was important by giving a low rating to items testing for irrelevance. This response pattern is consistent with previous studies using the same instrument such as those conducted in New Zealand (Brown, 2002), Virginia, USA (Calveric, 2010) Ankara, Turkey, (Vardar, 2010) and the Netherlands (Segers & Tillema, 2011).

Links between quantitative and qualitative data

Findings from the quantitative phase illustrate that participants agreed with the improvement function of assessment particularly when it provides accountability for students and school. They also tended to disagree with the items suggesting that assessment is irrelevant. However, in this study I wished to explore teachers' understanding of assessment in depth. English teachers in the region of Gowa might indicate that they are professional teachers who conduct assessment for learning which is similar to NZ teachers. However, it was important to conduct a deeper qualitative exploration on how these teachers understand their beliefs, why they have such an understanding as well as how they apply their understanding in practice. Some of the teachers' demographic characteristics, such as length of time

in teaching or the assignment level of their classes seemed likely to influence their beliefs and teaching practices. However, data from the quantitative survey could not provide an explanation for these associations. This data posed a question that required further study, which was pursued in the second, qualitative phase of the research.

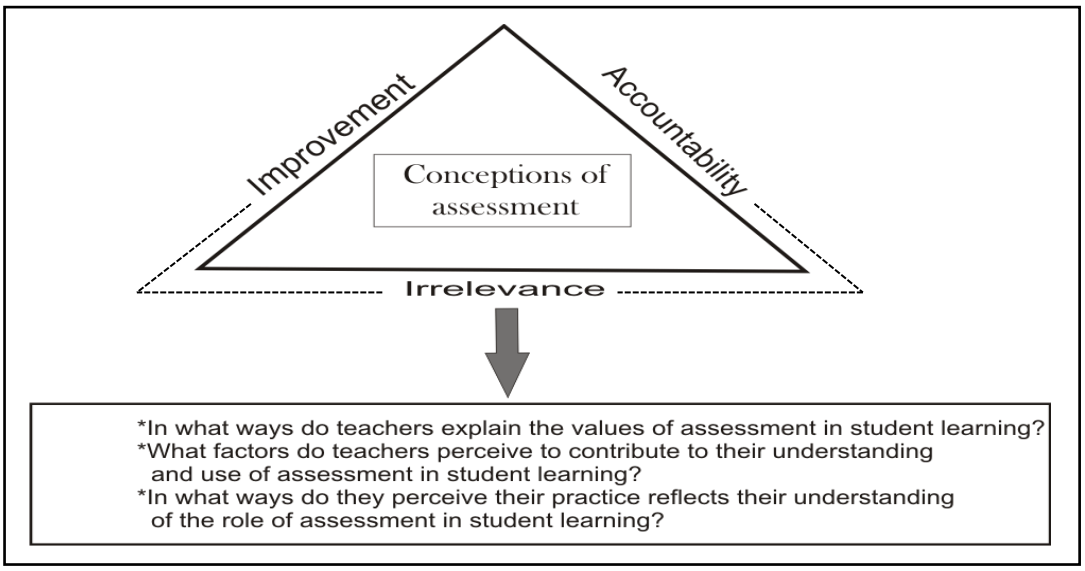


Figure 14: Questions to be followed-up

Qualitative findings

The qualitative phase of this study aimed at obtaining qualitative data using a case study approach that could explain how and why teachers with particular responses to the survey understood assessment. This second phase of the study was also carried out to comprehend teachers' assessment practices and factors contributing to their understanding of assessment. The next section presents participants' assessment practices evident in their assessment documents. This section is followed by descriptions of the qualitative interview results illustrating teachers' understanding of assessment and factors contributing to their understanding. Profiles of case study participants can be seen in the following Figure.

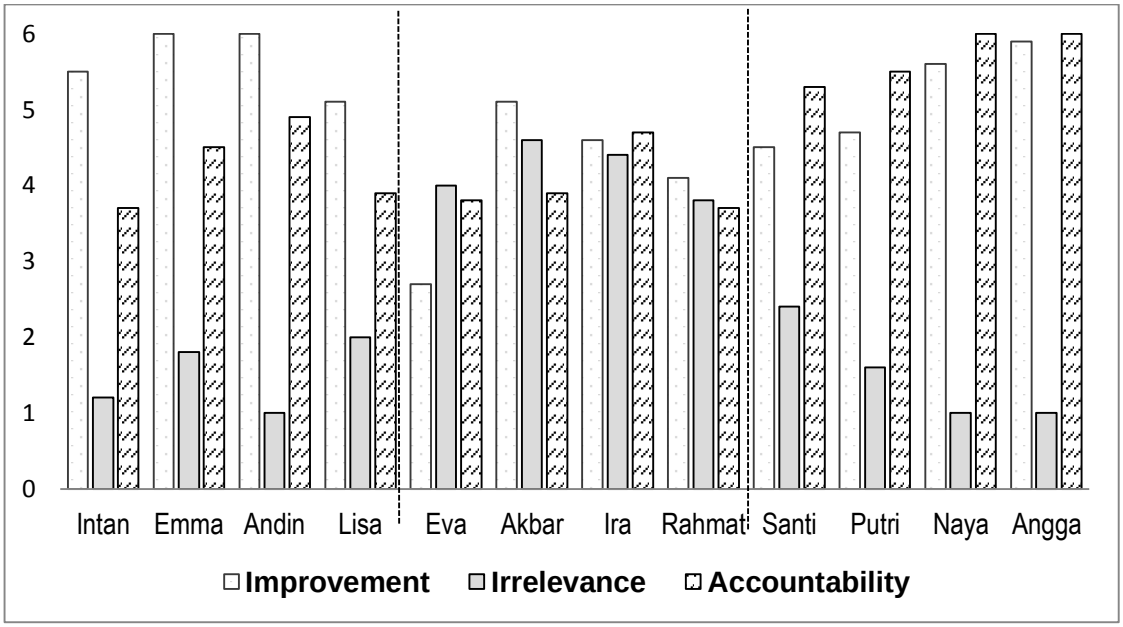


Figure 15: Profile of teachers who participated in the case study

Figure 15 shows that the first four teachers strongly agreed with items asking about improvement conceptions. This signifies an understanding that assessment is used to improve teachers' teaching and students' learning. The second four teachers in the graph were categorised as holding irrelevance conceptions. These teachers appeared to have complicated beliefs about assessment. However their responses indicated that they held strong beliefs about the irrelevance of assessment which was unlike the responses of participants from the other groups. This means that the second group viewed assessment as leading to improvement in teaching and learning, as well as providing accountability, but they also contended that assessment was irrelevant. The last group of teachers in the graph were strong supporters of accountability conceptions of assessment. These four teachers also held improvement conceptions and tended to disagree with items that indicate irrelevant functions of assessment. Thus two groups seemed to provide clear exemplars of teachers who viewed assessment as either predominantly improvement or accountability oriented, whereas the third group had mixed conceptions that were complicated by the view that assessment was also irrelevant. These profiles distinguished between three units of analysis or cases. The first four teachers represent the improvement (IM) conception case, the second four characterise the irrelevance (IR) conception case and the last four participants signify the accountability (AC) conception case. Details about these groups are presented in the following table.

Table 17 shows the purposive sample I constructed which includes 12 teachers all from different schools with various years of teaching experience, teaching locations (rural, suburban, and urban), and assignment levels (grade 7, 8 and 9).

Table 17: Demographic information of case study participants

Cluster	Gender	Teaching experience	Assignment level	School/PD
Improvement				
Intan	F	11-20	9	Urban
Emma	F	11-20	7	Suburban
Andin	F	3-10	7,8 & 9	Rural
Lisa	F	3-10	7 & 8	Suburban
Irrelevance				
Eva	F	3-10	7	Suburban
Akbar	M	Over 20	7	Urban
Ira	F	3-10	7	Urban
Rahmat	M	11-20	7	Rural
Accountability				
Santi	F	3-10	7, 8 & 9	Urban
Putri	F	11-20	9	Urban
Naya	F	11-20	7	Suburban
Angga	M	3-10	7,8, & 9	Rural

The improvement cluster comprises all female teachers teaching at different levels. The irrelevance cluster comprises a balance of male and female teachers teaching at the same level, and the accountability cluster is mostly comprised of female teachers from various assignment levels. All clusters included participants from three areas: urban, suburban and rural. The spread of participants' teaching assignment levels and school geographical locations across the three conceptions groups suggests that these demographic factors were not related strongly to teachers' conceptions.

Data from assessment documents

This section examines information about participants' assessment practices derived mostly from teachers' assessment documents. All teachers appeared to use

assessment documents that were developed on three levels: regional, school or PD, and the classroom. I divided these documents into three categories covering common regional documents, common school documents, and classroom documents. Common regional documents are the uniform documents used by all teachers in the region and are prepared by the Department of Education. Common school documents are developed by teachers at school or at a PD meeting, in other words, teachers at the same schools or those who participate in PD usually use the same documents. Other documents covered personal files developed by teachers for classroom use.

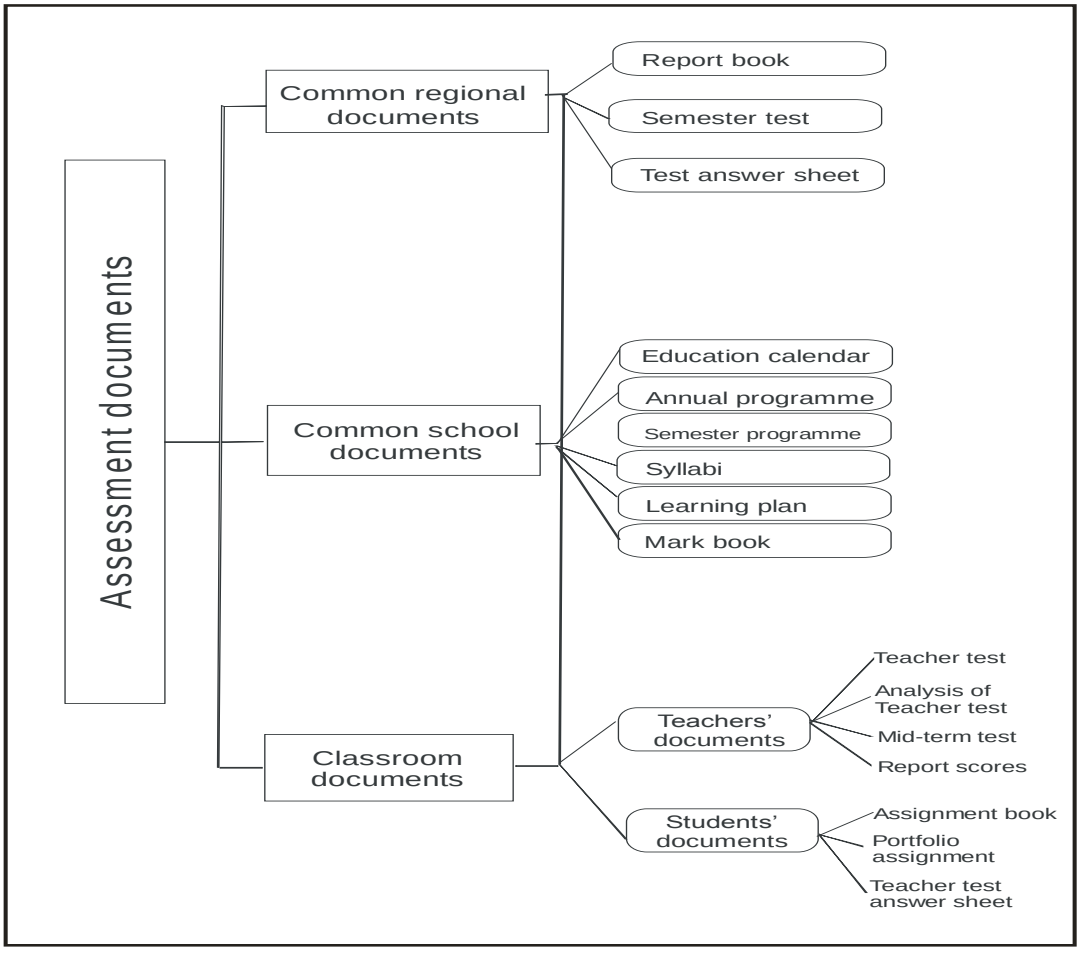


Figure 16: Participant assessment document

Common regional assessment documents

The Education Department of the region of Gowa manages all educational issues at the regional level including assessment documents like the students' report book, the semester test and the test answer sheet.

Students report book

Students' progress within one six-month semester is reported in this document. It is similar to a report card in which teacher record students' learning and report results to parents. It records the achievement of the competencies stated in the school level curriculum. Schools may design their own report book by referring to the format recommended by the Ministry of Education (Harianti, 2005). However, in the region of Gowa, the Education Department of the region designed and published the book. The report book also records non-academic data such as student behaviour and attendance. Teachers incorporate these aspects into their students' composite final score:

> "…their behaviour is also counted. These affective skills are described as A,B or C so we put our affective assessment in certain columns in the students' report book". (Andin, 1, 2012)

Semester test

The semester test measures students' comprehension against key competencies at the end of each semester. The key competencies cover four micro skills including speaking, reading, writing and listening, which are established at the national level. The test covers all basic competencies taught within the time frame of six months (Sudibyo, 2007). This means a semester test is conducted twice a year. In the region of Gowa the Department of Education schedules and organises the test by engaging selected teachers from each subject to develop the test which is then published under the department name. All schools in the region share the same test and carry it out using the same schedule.

The tests for Year 7 and Year 8 each contain fifteen items and use various formats such as short answer, completion, fill in the gap and jumbled sentences. For Year 9 students, the test consists of fifty items, all in a multiple-choice format with ninety minutes time allocation. The administration of the semester test follows that of the National Examination in the sense that it is a common test for the region administered using the same test, schedule, formats and time allocation.

Answer sheet for the semester test

The Department of Education in the region also provides an answer sheet for students to complete. For Year 7 and Year 8 students, the sheet is quite simple; it requires them to write their answers using pen. The answer sheet for Year 9

students is quite different: students need to blacken the options provided (A, B, C, D). The sheet requires students to present their personal data such as name, signature, identification, test number, school name, the subject and the date of test. All information on the sheet must be written using 2B pencils but teachers check it manually. This answer sheet is an exact model of that used in the National Examination, confirming that the government wishes to familiarise students with the final examination format.

Common school documents

Common school documents are documents used by teachers in their teaching and assessment. Teachers prepare these documents according to subject and curriculum level with colleagues at school or at a PD meeting. Shared documents are prepared for one semester or two. These teaching devices contain several sections like the education calendar, annual programme, semester programme, syllabi, learning plans and a mark book.

Education calendar

This document schedules students' learning over one academic year. Among the components in this calendar are the learning time for each of the key competencies per semester, the weeks of classroom teaching, the hours of classroom teaching and a schedule for holidays. In one year there are usually 34-38 teaching weeks and in each week 32-36 hours are allocated for teaching each subject (Suhardi, 2009).

Annual programme

This programme usually covers two semesters and contains the key competences and basic competences to be taught in each semester as well as number of classes teachers should spend teaching for each basic competency (Suhardi, 2009). A detailed list and descriptions of key competencies and basic competencies can be seen in Appendix A.

Semester programme

The semester programme is a plan for one semester. The document outlines the annual programme by scheduling the indicators for each basic competency. Teachers indicate the exact time for the delivery of teaching content (Suhardi,

2009). Both the annual and semester programmes are designed based on the education calendar.

Syllabi

The syllabus connects the key competencies and basic competencies[2] with suitable teaching materials, learning activities and indicators for competency achievement (Suhardi, 2009). Among the components are: the name of subject, the key competencies, basic competencies, teaching materials, learning activities, indicators of competency achievement, and assessments (technique, format, and example of tasks). The school has the right to develop most of these components except the key competences and the basic competencies. The current syllabus model encourages teachers to develop students' personal values while teaching key competencies. This covers moral standards shared by the community which are developed through education. For example when a teacher discusses a topic about making an appointment, he/she does not only explain and share the common utterances used for the topic, he/she might wish students to develop punctuality, honesty and mutual respect. This programme called character education, aimes to encourage students to develop both intelligence and good character (Suyanto, 2009). Through character education teachers are encouraged to educate their students to know the good, love the good and do the good (Ryan & Bohlin, 1999), and also to reason the good (Lickona, 1991).

Learning journal

The learning journal explains the syllabus in more detail. The learning journal is developed and agreed at PD and school meetings. This document describes the teaching procedure and strategies teachers use to organise their teaching in order to achieve key competencies (Suhardi, 2009). The learning journal is the guide for teachers to use in class, the laboratory or field. Components of this document include the teaching goal, teaching materials, teaching methods, teaching resources and assessment. For the assessment component, a rating scale and scoring grid are added to support the technique, format and example of each test presented in the syllabus. The guide also includes information about positive values or characteristics teachers wish to develop along with the teaching of particular contents.

[2] There are six key competencies and thirteen basic competencies in one semester for each grade

Mark book

Some participants use a commercially produced mark book available in book stores. This is an A4 size book presented in columns for recording information about school and class identity, students' names, their attendance per class and students' scores. Others redesigned the assessment book but maintained all elements presented in the printed book. There was no significant difference among IM, IR and AC teachers regarding these common documents. This suggests that all participants in this study were actively involved in PD meetings which were run by the Education Department of the region in conjunction with the Education Quality Assurance Agency in the province.

Classroom documents

In addition to common and shared documents, participants were also asked to share their personal assessment documents. These documents were used by teachers and students.

Teacher test

The teacher test is a periodical test designed and conducted by teachers to measure the achievement of one or two basic competencies (Harianti, 2005). In line with this goal, participants used their tests to assess manageable and discrete areas of learning. They scheduled assessments after teaching no more than two basic competencies. Teachers' tests mainly contained tasks familiar to students from class activities. The difference was that students did not have access to the textbook during this test. A teacher test might consist of five to ten items to be completed within eighty minutes (one class period). The formats used in this test were similar to the semester test particularly for Year 7 and Year 8. This suggests that teachers intended to prepare students by using a test, which is similar to the semester test.

Mid-term test

Teachers are expected to conduct this test to measure students' comprehension of several basic competencies taught within an eight to nine week period (Harianti, 2005). Schools do not set a special schedule for the mid-term test, as they do for the semester test. The mid-term test is similar to the teacher test in terms of the number of test items and the format but teachers use regular class time to conduct the test.

List of grades for final report

Subject teachers prepare this document for the school principal to sign off before each homeroom teacher writes the grades into the students' report books. The sheet contains a compilation of all students' scores covering four main sections: these are the average score from adding classroom assessments, the scores from teacher tests, the mid-term test and the semester test. In order to reduce these into one final score, schools utilised a particular averaging formula which might differ from one school to another in terms of priority. Some schools might accentuate teacher test and others focus more on the semester test.

Student documents

Student assignment books showed that all participants mainly assigned tasks taken from a textbook. Teachers marked and scored these tasks and returned their assignment books to students.

Students' work was also collected in their portfolio assignment. These documents were a useful means for students to demonstrate creativity to their teachers. Although creativity is not one of the key competencies, teachers seemed to be interested in developing this quality. All teachers assigned students a portfolio task. Another document from students was their answer sheets for teacher tests. These contained teachers' responses to students' answers for questions in the test.

Summary of document analysis

Table 18 shows the documents provided by participants according to the frequency of each document type and the users of the document. The table also illustrates that not all participants were willing to share their assessment documents. In addition, it was necessary to collect only one document representing regional and national assessment documents, as this document is common to all teachers.

Table 18: Overview of participants' assessment documents

Frequency and type of document used		User of document	Contributors	Total documents
Per unit	Quiz	Composed by teacher	All IM teacher 3/4 of IR teacher All AC teachers	11
	Quiz answer sheet	Marked by teacher	All IM teacher ¾ of IR teacher ¾ of AC teacher	11
	Analysis of quiz result	Composed by teacher	¼ of IM teacher None from IR cluster ¼ of AC teacher	2
	Assignment	Solved by student	All IM teacher All IR teacher All AC teachers	12
	Portfolio assignment	Solved by student and marked by teacher	¾ of IM teacher ¾ of IM teacher ¾ of AC teacher	9
Per three months	Mid-term test	Composed by teacher	All IM teacher 2/4 of IR teacher ¾ of AC teacher	9
	Mid-term answer sheet	Solved by student and marked by teacher	¾ of IM teacher 2/4 of IR teacher ¾ of AC teacher	8
Per single term (6-months)	Semester test	composed by regional teacher representative	Common regional document	1
	Test answer sheet	Marked by teacher	¾ of IM teacher 2/4 of IR teacher ¾ of AC teacher	8
	Assessment records	Filled by teacher	All IM teacher All IR teacher All AC teacher	12
per school year	Lesson plan	Composed by teacher	All IM teacher All IR teacher All AC teacher	12
	Student report book	Filled by teacher	Common regional document	1
	National Examination test	composed by national teacher representative	Common national document	1

Participants assessment documents were analysed by following the strategies suggested by McMillan and Schumacher (2010). The analysis mostly focused on the purposes of using the documents and how it was used. In other words, teachers' assessment documents were collected to understand if teachers' beliefs in assessment were reflected in practice. Data from these documents revealed that

teachers used common regional documents, common school documents and classroom documents in their assessment practices. IM practices resembled AC practices in terms of the way the teachers developed the test, the range of students' tasks and the marking system. This practice suggests consistency between beliefs and practices in that these teachers conducted assessment both to improve teaching and learning and to signify the accountability of students, teacher, and school.

In contrast, IR teachers reported quite different practices reflecting their beliefs that assessment is irrelevant to learning. These teachers did not believe that assessment had a role to play in improving teaching and learning, and tended to conduct assessment because it was compulsory. Thus, all three case study groups show consistent beliefs and practices in regard to assessment. Further details of teachers' and students' assessment documents are presented in the next section along with interview data.

Data from interviews

The analysis of key ideas involved coding, merging codes, and naming and renaming themes. The interview data revealed that teachers' understanding of assessment purposes fell into two groups. They differentiated between internal assessment and external assessment. Teachers' understanding of these assessment types are illustrated along with supporting quotes in the following section. I start by presenting the case study from the improvement case followed by the irrelevance case and finishing with the accountability case.

Improvement group (IM)

Conceptions of the purposes of assessment

Intan, Emma, Andin and Lisa are examples of teachers holding a conception of assessment which valued improving teaching and learning. Their responses to the questionnaire reflected their strong advocacy for using assessment for students' development, to help learning and develop accuracy. In addition, they agreed with accountability purposes in assessment and were unsure whether assessment could be considered irrelevant. The trends shown by these participants indicated a similar pattern.

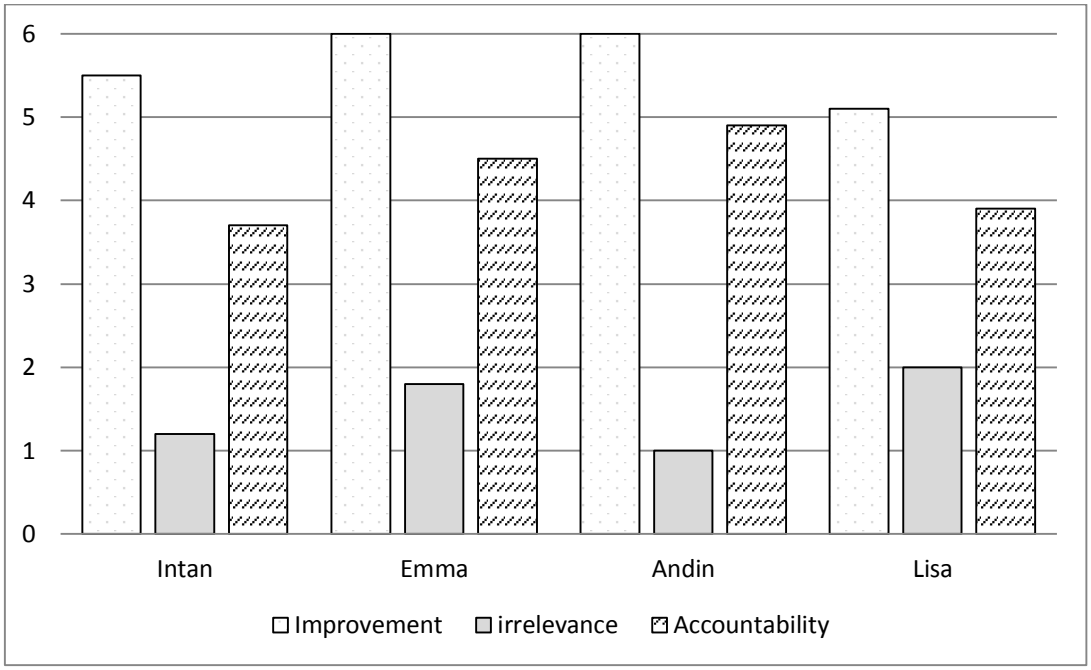


Figure 17: Profile of teachers in the improvement conception case

IM teachers' values of assessment fall into the two main categories of internal and external assessment. These categories are also used to organise the themes in the irrelevance and accountability cases. Internal assessment in the improvement case consists of the themes of openness to change, developing values, authentic formative assessment practices, grading to show achievement and teachers' autonomy.

Internal assessment

IM teachers favoured internal assessment where they could adjust teaching, develop values and practise with numerous assessment strategies. Their statements about assessment revealed that they were open to change, believed that assessment was an effective way to teach students good values and indicated an agreement to score students.

Openness to change

IM teachers believed that the main purpose of assessment was to inform teaching. Through internal assessment teachers could measure the quality of teachers' teaching. IM teachers reported that it was crucial for them to reflect upon students' assessment results.

> "…[low scores] sometimes make me disappointed; I thought I'd taught them well". (Intan, p.10, 2012)

> "I question myself, is my teaching effective? Do my students understand the lesson?" (Lisa, p.9, 2012)

These teachers worried about the efficacy of their teaching and signalled a willingness to modify their teaching continuously. IM teachers tended to articulate a belief in formative assessment because they used their classroom assessments to inform their teaching. In their consideration of students' results, IM teachers involved their students in the reflection process. This two-way communication tracked the efficacy of teachers' teaching:

> "Every semester I ask my students to comment on my teaching…do they like the strategies I use, how they want the learning process run". (Lisa, p.12-13, 2012)

To IM teachers, internal assessment could also bridge a better relationship with students. IM teachers seemed to hold themselves accountable for students' motivation in addition to their academic success. They took responsibility for encouraging low-achieving students through personal conversations:

> "I try to motivate them, I give them feedback, I ask what causes such unsatisfactory results, what is the problem, which part is hard and so on…So I assist them to realise reasons behind their failure". (Intan p.18, 2012)

IM teachers appreciated that a low score could affect students' confidence negatively so these teachers encouraged students not to be pessimistic:

> "I convince them that they have many opportunities to improve and I give them those chances". (Emma, p.3, 2012)

Teachers' reflections and communication with students resulted in changes to their teaching approaches:

> "I realised that I talk fast; this might be one cause". (Intan, p.10, 2012)

And

> "This strategy might work well in class A or B but does not necessarily work in Class D. Sometimes I almost give up, but at the same time, it motivates me to continuously search for the right strategies". (Emma, p.11, 2012)

This is an example of how IM teachers adjusted their teaching in order to meet student needs. IM teachers responded to assessment results positively in the sense that they were encouraged '*to learn more in order to be a good teacher'. (Intan, p.13, 2012)*. This might explain why they valued classroom assessment.

> "My assessment is 99.9% reliable…I assess my students based on their proficiency and I am confident that it is accurate; I am opposed to those who manipulate students' scores". (Emma, p.15, 2012)

Thus, IM teachers favoured internal assessments as an evaluation tool in order to improve the quality of their teaching. They trusted that internal assessment generated trustworthy results.

Another way that IM teachers showed their openness to change was in supplementing the set textbook. IM teachers relied on the textbook as the main source of teaching materials. IM teachers were confident in the validity of the textbook as a teaching tool. The textbook published by the Ministry of Education contains all key competencies and tasks that are accessible for both students and teachers. However, unlike teachers in the other cases, the IM teachers were willing to include additional materials.

> "I use the textbook as the main resource of my classroom activities but also combine with other resources especially for listening activities". (Intan, p.14, 2012)

This extract signals that IM teachers were open to supplementing the textbook with extra resources:

> "I use various [teaching] resources…and I create some of my own". (Emma, p.10, 2012)

> "I love going to book stores to find additional resources. I use all [kinds of] relevant materials in my teaching including CDs and cassettes". (Intan, p.13, 2012)

Teachers from this group also showed me some teaching materials they used in the classroom. Few of these were taken from the textbook. Most were colourful, laminated and taken from a newspaper, magazine and other authentic sources. These teaching materials and the portfolio assignment suggest that IM teachers are interested in the use of authentic materials to assess students' learning.

IM teachers demonstrated an understanding of a semi-constructivist view of learning (James, 2008) particularly in their concern for students' engagement in learning. This was evident in their attempts to create an enjoyable learning experience in internally-assessed classroom activities.

Developing values

IM teachers believed that assessment was meaningful both to illustrate student academic performance and to develop positive values. They described assessment tasks as an effective way to support the government's new programme of building character through education. These teachers appeared to believe that assessment could be used to develop positive characteristics that would benefit students' future lives. IM teachers mentioned several values that students could develop through assessment.

> "…a good score helps [a student] to improve in confidence… discipline can be witnessed through their intention to do homework which also indicates good responsibility". (Andin, p.11, 2012)

Another important characteristic that teachers wished their students to cultivate was honesty. IM teachers encouraged students to be fair in approaching their tests.

> "…no cheating…if we assist them doing this, it becomes a habit and that is how they learn about honesty". (Emma, p.12, 2012)

IM teachers reported that not only could students acquire positive values from assessment, but practicing these values might also have a role in reshaping their own understanding and practices of assessment. These teachers placed fairness at the top of the qualities they themselves developed because they believed in the role of assessment in presenting a *'true score'. (Emma, p.13, 2012).* Some other values they perceived to develop through assessment practices included confidence and responsibility. This teacher reflected on the importance of presenting the authentic score and her sense of responsibility for conducting fair assessment. In return, students appeared to welcome this consistency and developed close relationships with their teachers.

> "...you know, students usually meet me at my office and we walk down to class together". (Emma, p.14, 2012)

Thus, IM teachers believed that assessment has the power to influence students and teachers positively. In addition to these values, these teachers felt internal assessment could be used to stimulate students to learn:

> "Assessment makes students enthusiastic to learn". (Emma, p.11, 2012)

These teachers perceived that assessment was a crucial component of studying. Teachers emphasised their students' investment in assessment and the positive influence this made on their learning. They felt that students willingly prepared themselves for tests:

> "They know that I'll assess and score them so they will study for them". (Andin, p.2, 2012)

This extract suggests that testing is an effective motivational strategy that could positively affect students' persistence as students were eager to obtain a good score.

> "They compete with their friends, compete to get high scores, it makes them study harder". (Intan, p. 9, 2012)

Some IM teachers even used rewards to maximize students' interest in learning.

> "…once I challenged them with a group task and I told them that the best group would get special presents…they truly competed to present the best". (Intan, p.2, 2012)

IM teachers found tests and assignments effective in motivating students to learn. Some IM teachers reported that grades and rewards performed the same function as feedback which was conducted mainly to make learning interesting rather than focusing on students' position against the learning goals. IM teachers found that assessment promoted positive values in their students and they also expressed a strong belief in the value of formative assessment practices to develop learning in accordance with their constructivist approach to teaching. These practices were only visible in classrooms where teachers were given authority to control their teaching. Yet, alongside the IM teachers reported teaching practices that showed their belief in authentic formative assessment they balanced equally strong views about the value of grading to show achievement. The final theme illustrates the toll this places on their autonomy as teachers.

Authentic formative assessment practices

IM teachers based their teaching and assessment on the four micro skills; reading, writing, listening and speaking. They recorded these scores in their modified scoring grid document. They reported that authentic English teaching requires all four modes of the language to be taught. In the current curriculum, called the school level curriculum, English language scores are composite. This means that scores for listening, speaking, reading and writing are collapsed into a single overall score and some teachers prefer to concentrate their teaching on reading and writing at the expense of skills that are not examined. Nonetheless, teachers in the improvement group reported teaching and assessing students according to all four language skills.

> "…I measure speaking, listening, reading and writing. So students are scored based on these skills". (Lisa, p.4, 2012)

IM teachers reported that they specifically taught each of these skills and tried to involve their students in assessments that would inform their learning. These teachers reported using oral tests in speaking, listening and some reading classes, and employed written tests in writing and reading comprehension classes.

> "…for speaking, students perform a dialogue, they make their own dialogue based on the text and materials we discussed and they demonstrate it in front of the class; sometimes they also tell stories when we discuss narrative text…Reading activities cover reading aloud, reading comprehension…For writing, students compose paragraph and design an advertisement or announcement, or write a letter". (Intan p.5, 2012)

IM teachers wished to use assessment that matched each skill which would ensure the validity of the test. These teachers insisted that it was necessary to use several criteria to determine students' scores such as looking at students' responses in listening, as well as whether or not they followed the commands given by teachers. Speaking skills were challenging in that there were multiple aspects to evaluate.

> "I use some criteria such as content, fluency, intonation, and pronunciation". (Emma, p.8, 2012)

Reading aloud is another activity that required teachers to use an oral test where they utilised similar criteria to those in a speaking assessment. Teachers used the following criteria for writing assessments:

> "We look at punctuation, content, vocabulary, grammar and order". (Lisa, p.8, 2012)

These models of assessment reflect IM teachers' belief in the importance of authenticity in internal assessment. IM teachers also reported developing particular criteria and rubrics in assessment to assist student learning. These teachers insisted that sharing the assessment criteria with students was important in improving students' learning because it helped them to:

> "…know which aspects will be measured so that they will think of them, and they will work on them". (Intan, p.8, 2012)

This extract implies that teachers tried to communicate their expectations in order to guide students towards performing work that would meet the set criteria (Martins, 2008). It also suggests that the IM teachers tried to share responsibility for the learning and assessment process with their students.

Using several assessment formats outside the examination schedule such as *'role-play, making a movie review and presentation'.(Intan, p.6, 2012)* signified that these teachers utilised authentic assessment through tasks which represented real-world problems (Frey & Schmitt, 2007) to observe a fuller picture of students' achievement (Horn, 2006). In other words, these teachers chose criteria like those above to represent the real skills students need to obtain in order to be a proficient English user.

> "I want them to be able to speak English". (Intan, 4, 2012)

IM teachers reported another understanding of authenticity by considering students' prior knowledge when planning their teaching. These teachers made assessment a continuous process; one which required them to observe students' learning carefully.

> "Is there any progress from not knowing to knowing, do they make an improvement?" (Andin, p.1, 2012)

This teacher tried to diagnose students' learning and regarded meeting the learning outcomes of English teaching as the major goal in teaching the subject. To meet this purpose, IM teachers reported carefully selecting appropriate pedagogy to support authentic teaching tasks.

> "I use a CTL (contextual teaching and learning) approach: the lesson I present should reflect real experience relevant to our context". (Intan, p.6, 2012)

This teacher demonstrated an understanding that teaching pedagogy could enable students to connect the content to the real world (Berns & Erickson, 2001). The Ministry of Education in the country has encouraged teachers to use this approach since the implementation of this competency-based curriculum unfortunately this has not been popular with Indonesian teachers (Satriani, Emilia, & Gunawan, 2012). Consequently, it was interesting to discover that IM teachers considered this approach important.

> "I ask myself can they [students] practise it in their daily life. Is it useful?" (Lisa, p.1, 2012)

To some extent, IM teachers' support for authenticity is a challenge because English is learnt as a foreign language in Indonesia. This is why they adjusted teaching materials to meet students' level of English.

> "They do not study English at primary school so I start with ABC in our first lessons. I know this is not stated in the curriculum, but I have to [teach the material]". (Lisa, 9, 2012)

IM teachers illustrated authenticity in assessment by assessing students per skill. They also utilised what they believed as the appropriate pedagogy in teaching English. These teachers considered their students' background knowledge crucial to inform their English teaching. They felt that contextualising teaching materials and creating interesting learning experiences could support student achievement.

In addition, IM teachers involved students in aspects of the assessment process by asking students to exchange or swap their books and mark one another's work. This practice indicated that IM teachers had an interest in peer assessment and wanted their students to play an authentic role in the assessment process. IM teachers gave evaluative feedback at the bottom of each task like 'excellent', 'good' and 'well done'. Teachers appeared to consider that these comments motivated students to learn, which suggests a commitment to formative assessment practices. Moreover, in preparing the final grade for students, IM teachers combined the previous and current curriculum final grid format. They included four aspects of classroom assessment like speaking, listening, reading and writing along with students' scores

for the semester test, the mid-term test and the semester test in the scoring grid although these had to be combined into a single score for English. Assessing students' skills separately indicated that they supported a competency-based curriculum but the fusion of all scores into one suggested that they followed the current school level curriculum format. They also added some notes on whether the KKM (standard score) was achieved or not. This practice indicates consistency between their beliefs, teaching and assessment.

Grading to show achievement

IM teachers believed in authentic classroom-based assessment that would inform teaching and learning. However, they also indicated a strong interest in grading practices. IM teachers believed that assessment should be useful in revealing student learning. They argued that testing was a preliminary instrument to measure the quality of students' work. One confidently stated that it helped

> "…to measure students' proficiency, how well they comprehend the lesson". (Intan, p.1, 2012)

IM teachers appeared to believe that students have different levels of competency; assessment assisted them to recognise students' position among their fellows. IM teachers assumed that the higher the score, the more learning a student had achieved. IM teachers seemed to be happy in using assessment results to map student achievement levels. These teachers appeared to believe that assessment and scoring were inseparable.

> "…it is a sign, whether students master the teaching materials or not, if they get 10 (ten), it means they understand the lesson well". (Andin, p.2, 2012)

In other words, IM teachers viewed a score as a symbol of achievement. This understanding appeared to be strong and was reflected in these teachers' teaching and assessment practices. IM teachers marked students' sheets by checking correct answers and crossing the wrong ones and allocating points for each test section (usually per skill) to be calculated later as the score for the test. Teachers justified this marking style both because ticking a correct answer pleased and motivated the students and because crossing an incorrect answer encouraged students to think further about or work out what was wrong. Both aimed to provoke student learning. Teachers also contended that a good score was important for

students' self-actualisation and provided an indication of their further education level.

> "…when students have a good score or when they have good achievement, they will be recognized in our society". (Intan, p.4, 2012)

It seemed that both teachers and students believed that scores reflected students' achievement, suggesting that scoring and grading were a valued parts of learning. Students were '*happy when their work is scored' (Andin, p.9, 2012)* and requested every task to be rated by asking *'will you mark this or not?' (Lisa, p.10, 2012).* Students appeared to '*study only to get a good score' (Lisa, p.2, 2012)*. IM teachers believed that assessment could be used to identify students' position against the standards, acknowledge students' positions among their fellows and determine positive impressions among in Indonesian society.

> "They feel like the score is a reward for their hard work, they feel much appreciated". (Intan, p.9, 2012)

This phenomenon indicates a communal assumption that grading is crucial for communicating the quality of learning. This shared perception of the value of grading suggests that scoring and grading is a significant part of the participants' culture. The IM teachers did not find an emphasis on grade was incompatible with a focus on formative assessment.

External assessment

IM teachers believed that exam-based external assessments are an important means of evaluation to certify students' learning, teachers' teaching, and the accountability of a school. However, they also felt that external examinations had a negative impact on a teacher's autonomy, equity among students and they doubted the credibility of external tests. This ambivalence was reflected in the themes that arose which related to conflicting conceptions, teacher autonomy and the credibility of external assessment practices.

Conflicting conceptions

IM teachers believed that external assessment could assure teachers' accountability. These teachers believed that the collection of students' assessment results was an indication of the quality of teachers' teaching.

> "It tells me whether I am able to transfer knowledge to my students". (Andin, p.9, 2012)

This understanding suggests that IM teachers supported conventional or behaviourist views of learning. These teachers assumed that students' success in learning was visible when they could reproduce teachers' knowledge. In short, IM teachers considered that assessment was a powerful way to communicate the standard or quality of students' learning and teachers' teaching.

> "...they [parents] look at the result. When it is good, it means the teachers are qualified". (Emma, p.4, 2012)

In addition, IM teachers agreed with the government's use of external assessment to evaluate the quality of a school. IM teachers understood the main function of this high-stakes external assessment was to assign each school a rank in the region, province and country.

> "To determine the quality of students, the school, the region, the province, through assessment we can measure the percentage of quality improvement". (Andin, p.11, 2012)

And,

> "The students' proficiency indicates the quality of a school" (Lisa, p.13, 2012)

This extract illustrates that IM teachers acknowledged the role of society in assessment. The strong confidence placed in examination scores as a measure of students' proficiency by students, teachers and the community appeared to be underpinned by the National Examination. While IM teachers supported the use of high-stakes assessment to fulfil the above purpose, they appeared to have mixed feelings about it. In contrast to their support for assessment as a means to hold schools accountable, IM teachers also argued that using a score as the indicator of quality placed schools in a difficult situation.

> "Each school is ashamed [to have low scores] particularly because this is [the overall score] ranked nationally…" (Emma, p.4, 2012)

IM teachers reported that some schools felt obliged to engineer acceptable scores to fulfil the demands of society, or to encourage new enrolments.

> "We are required to report scores which are acceptable or above the standards, otherwise teachers' and the school's reputation is contested". (Emma, p.4, 2012)

The effect of student achievement in external examinations on the wider community was obvious to this teacher:

> "Once, a school announced that 100% [of its] students had passed the examination. Following this the local community really appreciated and trusted it. I mean more parents sent their kids to that school. However, a few years ago when students' scores dropped, the numbers of applicants also declined". (Emma, p.12, 2012).

Students' examination results determined the reputation of a school, influencing prospective parents seeking to identify and select prestigious schools for their children. IM teachers seemed to be in doubt about the validity of using external assessment to hold a school accountable in this way. They suggested that this also led to several unintended consequences for teachers' autonomy and the credibility of tests.

The credibility of external assessments

Teachers interpreted policies and regulations set at regional and national level including the semester test and the National Examination as 'must do' activities. Teachers adhered to policies implemented at both levels.

> "…this is our condition, we are ruled by that regulation [the National Examination and the semester test], both regionally and nationally" (Lisa, 3, 2012)

Regional involvement in creating the semester test is quite recent because up until two years ago, teachers of each subject created and tested their own students. At first, teachers responded positively to the new policy because it allowed both regional and national assessment practices to become more familiar to students. Yet, it soon became apparent that the new policy ignored the rural teachers' input and gave the regional government a dominant role in assessment. Teachers also felt that this policy showed that the government distrusted and disregarded their competence in evaluating students.

> "…in developing the semester test... the department [of education] did not invite teachers from this [rural] area". (Lisa, 3, 2012)

IM teachers believed that taking an active role in constructing the semester test should remain an important part of their job. IM teachers believed that they were the most familiar and appropriate evaluators in understanding students' genuine competence. The prominent role played by the Department of Education in assessment practices worried teachers. They explained that this policy had resulted in a drop in students' scores.

> "…we know the best test for students; their test [regional test] is too hard for my students". (Lisa, p.19, 2012)

This extract seems to indicate two things. Firstly, it appears to show that teachers were anxious about students' inability to meet the standard, when they felt accountable for students' results. Secondly, teachers were certain that the test *'does not measure what my students have learnt' (Lisa, p.19, 2012).* The semester test may have evaluated aspects of the curriculum attainment target which teachers were unable to meet. In other words, teachers might have left some key competencies untaught in their efforts to adjust teaching to suit students' competency level. Also, teachers had more confidence in their own ability to produce a reliable test than that of regional officials. Hence, the regional administration of the semester test and the teachers' exclusion from the assessment process led teachers to judge government assessment to be less than reliable. These teachers argued that the local department should trust individual teachers to design and administer the semester test.

> "I think it would be better if the department returned the making of tests to teachers" (Intan, p.17, 2013).

These teachers objected that the National Examination had become the main evaluation tool for students' achievement or school quality. They contended it was unfair because

> "The National Examination only looks at our students' proficiency from a general viewpoint…they do not look at students' proficiency in each school, they standardise them" [sounding emotional]. (Lisa, p.2, 2012)

This teacher objected to the use of one standard applied to all students. IM teachers reflected that the dominant role played by external agencies in constructing an examination could be unfair to students with dissimilar competences and different school facilities. There were differences between rural and urban schools, in

different teacher qualifications and in what was addressed in national or local standards. IM teachers also believed that external assessment (particularly the National Examination) was unbalanced and did not capture a holistic picture of the students' English proficiency.

> "They do not know what happens in the class, they only look at the final grade". (Andin, p.3, 2012)

Thus it is clear that IM teachers contested the validity of external assessment. They insisted that more localised assessment conducted by familiar evaluators (teachers of the subject) might be fairer and more reliable. In addition, IM teachers doubted the credibility of external assessment.

> "I have witnessed suspicious practices; sure I did not make up this story, two students had exactly identical answers with the key, including the words and commas. I know one student well; his competence is not at that level". (Intan, p.16, 2012)

IM teachers had negative impressions of the credibility of external assessment practices because they suspected answers were leaked to students. These teachers expressed strong doubts in the system.

> "I am sure, it's impossible [for students to answer all questions correctly] …there must be a conspiracy". (Emma, p.3. 2012)

This perception undermined IM teachers' confidence in external assessment. They believed that the emphasis placed on one external test was encouraging schools to present inauthentic assessment results.

> "...the government demands that students must have a good score…." (Emma, p.4, 2012)

And

> "You know we can negotiate it [the score] and students may get assistance [in doing the test]". (Lisa, p.17, 2012)

IM teachers believed that assessment was an important tool to measure student, teacher and school accountability. However, they reported that external examinations had a negative impact on teachers' autonomy and led to untrustworthy external assessment practices. Teachers also argued that the high

value placed on examinations by the wider community to evaluate schools potentially influenced schools to compromise the credibility of external assessment.

IM teachers reported assessment practices that accorded with their understanding of assessment. Being unable to mitigate against the dominant role of external assessment in their context, they conformed to practices that fit AoL. In informal classroom assessment situations where their autonomy was acknowledged, IM teachers confidently implemented constructivist assessment practices. However, these appeared to only receive a very small portion of the teaching time because teachers used the opportunities presented by other internal assessments like teacher tests or mid-term semester tests, to familiarise students with the format used in external assessment.

Teachers are responsible for administering classroom tests and mid-term semester tests. For the classroom test, IM teachers consistently covered the two basic competencies using familiar formats like completion, jumbled sentences, and short answers- even translating passages. The mid-term semester tests covered more key competencies where IM teachers utilised a wider range of tests formats including multiple-choice, writing memo and matching in addition to formats they used in classroom tests. IM teachers made a test inventory identifying the competencies and learning indicators they wished to measure in the test. This level of preparation indicated that IM teachers considered their tests to be valid assessments of learning.

However, their focus on external examinations seemed to conflict with their other, formative practices. These teachers reported that assessments were an exercise to *'familiarise students with the final examination' (Andin, p.7, 2012).* For this reason, the semester test was intended to shadow the final examination and teachers created internal tests that resembled models used in external assessment.

> "I use matching, multiple choice, particularly for the ninth grade students". (Intan, p.7, 2012)

IM teachers' decision to familiarise students with external assessment formats indicated their focus on external assessment. In extreme cases, teachers engaged in intense examination preparation providing extra classes and exercises similar to the examination formats.

> "In the second semester of the third year, we give extra lessons to students which are conducted after school. Moreover, four or five weeks before the examination, the curriculum section [of the school] will design a new lesson schedule so that students will learn only the subjects tested in the examination". (Intan, p.3, 2012)

Teachers seemed to be unaware that such practices limited the subject matter content as well as potentially limiting students' engagement with the content. It seems that the National Examination influenced both these teachers and their schools to maintain a focus on summative assessment practices.

IM case summary

IM teachers reported divergent understandings of assessment purposes that could be categorised as both formative and summative. They favoured teacher-driven assessment (formative) where they had room to develop and implement strategies to improve teaching, learning, and felt that their own assessments were likely to be effective and reliable. In contrast, they reported less faith in exam-based assessment conducted by external agencies. Although they appreciated that exam-based assessment (summative) was necessary to certify students or to make teachers and schools accountable to the community, they argued that external assessment could impact negatively upon teachers, students and the credibility of tests. Their understanding and practices of assessment illustrated a conflict between internal and external assessment. IM teachers demonstrated formative assessment practices to reflect their understanding of internal assessment and more summative types of assessment to reflect their focus on external accountability. The emphasis placed on summative types of assessment despite their strong belief in assessment for learning revealed their conflicting conceptions.

Irrelevance group (IR)

Conceptions of the purposes of assessment

Eva, Akbar, Irma and Rahmat are categorised as holding the irrelevance conceptions of assessment. These teachers indicated inconsistent responses against the three types of assessment conceptions. Their inconsistencies reveal a conception that assessment is irrelevant and highlight these participants concerns about the inaccuracy of assessment, demonstrate ignorance of the purposes for assessment results, and reflect their assumptions that assessment has little or negative impact on teaching.

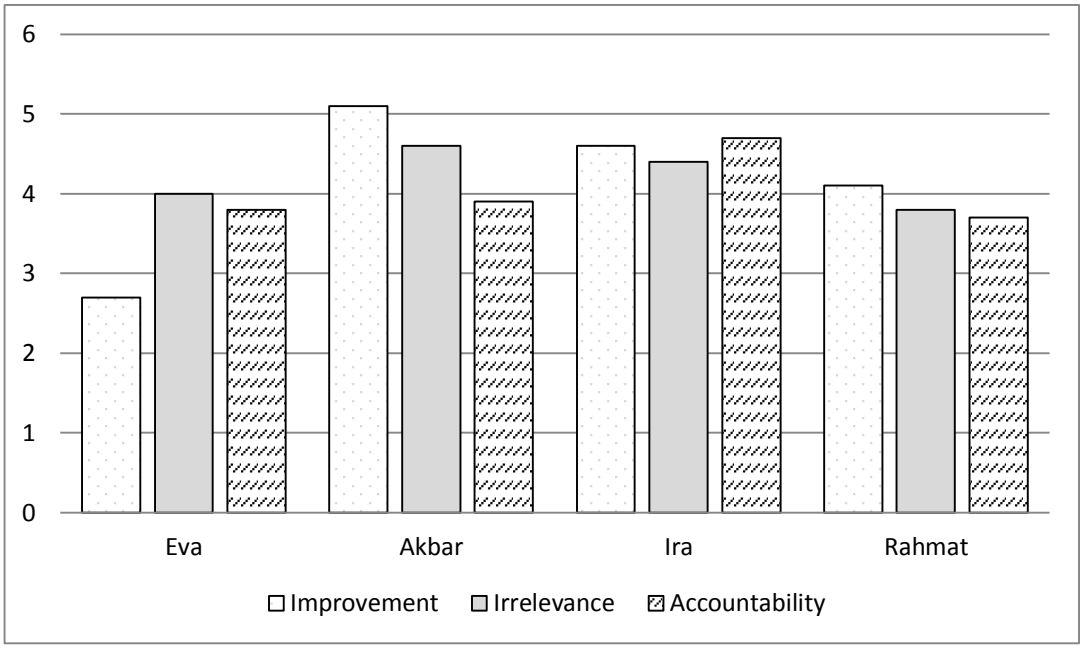


Figure 18: Profile of teachers in the irrelevance conception case

Internal assessment

The category of internal assessment covers IR teachers' beliefs about modifying teaching, developing values, authentic summative assessment practices and using grading as a motivational strategy.

Modifying teaching

IR teachers reported that assessment helped them to modify their teaching. They adjusted teaching in two stages; during class interaction and after conducting internal tests. IR teachers believed that internal assessment was a tool to '*evaluate the quality of teachers' teaching' (Irma, p.1, 2012).* This was measurable in student performance when undertaking tasks as well as in their responses to instructions.

> "When I see students only give limited participation, it means the teaching materials are hard so I have to shift to other materials which are easier and friendlier". (Rahmat, p.10, 2012)

IR teachers argued that another source of information about quality teaching could be obtained through testing what students' learnt within a certain time, for example over a period of two or three weeks. Considering the information from regular class tests, IR teachers argued that internal assessment:

> "Informs me whether I should re-teach the same lesson or conduct remedial classes". (Irma, p.10, 2012)

IR teachers identified that re-teaching the same materials or running remedial classes is a form of teaching adjustment. However these activities tended to be superficial because IR teachers did not diagnose students' low scores or reflect on the methods they used. Teachers repeated the same lesson in the same way, in other words, these teachers did not modify their teaching during any intervention.

> "I re-teach the same materials to students…using the same strategies". (Eva, p.13, 2012)

This comment might imply that IR teachers were not interested in challenging themselves and their students through more effective teaching strategies or more interesting class activities. The implementation of the same strategies for the review programme suggested that they did not conduct the intervention for improvement purposes. This teacher innocently stated *'at least, I repeat my teaching' (Eva, p.13, 2012).* This statement implies that assessment minimally impacts on teachers' teaching. It may also reflect IR teachers' pessimistic expectations for student improvement in learning. These teachers tended to claim that students' competence was unchanged with or without intervention or remedial classes.

> "Although we repeat discussing one theme and give students a second chance to sit a test [teacher test] their scores remain the same". (Eva, p.3, 2012)

> "Well, we know our students' competence level. It is useless to retest or re-teach certain materials, there will be no improvement". (Rahmat, p.5, 2012)

It seemed that internal assessment did not inspire these teachers to make changes in order to help students gain the standard required. When asked about criteria repeating the same teaching approach and using the same teaching materials in intervention programmes, teachers argued

> "…at least there is a slight increase in students' scores". (Irma, p.7, 2012)

This suggests that teachers might stop the remedial classes before students achieved the KKM. To some extent, IR teachers demonstrated indifference to student development in learning and continually complained about low student

competence. It appeared that IR teachers' conceptions of using assessment results to modify their teaching involved; switching materials that were presumed to be hard to easier teaching materials, or re-teaching the same lessons without further modification. Their unwillingness to stretch their learners was also evident in the way that IR teachers used textbooks as the only teaching resource. These teachers concentrated on what they perceived to be easier content from the textbook and concentrated on using tasks from the textbook such *'essay, matching, dialogue, and fill-in-the-gap'. (Eva, p.10, 2012)* that prepared students for external examinations.

Developing values

IR teachers believed that internal assessment encouraged students to develop several positive learning attitudes like *'discipline, and confidence' (Irma, p.8, 2012)*. They attempted to build positive characteristics such as fairness through internal assessment by warning students that '*those who cheat or give assistance [to fellow students] during the test will get no score' (Akbar, p.8, 2012).* These teachers favoured internal assessment as they believed it to be more genuine and more motivating because students' can be directly involved in activities.

> "Students' can confidently participate or raise their hands to answer my questions, regardless of whether they give correct or wrong answers but to me it is a good sign that they want to learn". (Rahmat, p.3, 2012)

In other words, IR teachers believed that internal assessment could *'motivate them [students] to study harder, (Akbar, p.1, 2012).* However, they signalled a hesitation over whether assessments were an effective means of achieving the values.

> "Assessment should teach students about fairness…but it does not ". (Eva, p.14, 2012)

IR teachers explained why they held contradictory understandings particularly of external assessment:

> "Assistance they get during the exam blows the value away". (Eva, p.15, 2012)

Conspiracy among markers and teachers during external examinations like manipulating students' marks or giving assistance undermined these teachers' views regarding assessment and motivation for study.

> "They only study for the first day of the exam…when they are given the answer key, they won't study anymore". (Rahmat, p.15, 2012)

This theme might suggest that IR teachers doubt the value of assessment to support the character education programme run by the Ministry of Education. According to IR teachers, these practices had the potential to:

> "Teach students about unfairness or dishonesty…this could even result in more serious consequences like creating a corrupt generation". (Rahmat, p.15, 2012)

IR teachers contended that assessment was a good strategy to motivate student learning as well as to encourage the development of some values. However, these teachers argued that such values were only achievable in classroom activities. In an external assessment which focuses more on results, IR teachers questioned the concept of teaching and modelling these values.

Authentic assessment practices

IR teachers' classroom assessment followed traditional practices with a focus on preparing for external examinations. These practices included observation, paper-pencil tests and other traditional approaches like translating sentences and working on sentence structures (parsing). These teachers believed that traditional assessment practices were authentic and credible.

IR teachers stated that teaching should be authentic in terms of the relevance to students' real lives; however, they placed greater emphasis on preparing students for external tests. IR teachers reported teaching and assessment practices that were not based on key competencies as suggested in the national curriculum.

> "I've never had listening activities in my class, it is very hard for my students". (Eva, p.10, 2012)

Although some teachers in the group decided to include listening activities, they restricted these to simple exercises.

> "After they make two sentences about a chair, I usually ask them to read it aloud to the class while other students listen". (Rahmat, 11, 2012)

This implied that IR teachers' understanding of listening activities was limited and they did not value teaching authentic listening skills such as responding to

information. They might consider that listening skills are less crucial because these are neither reported in the students' report books nor tested in the final examination.

> "Why should we assess those skills? [listening and speaking] We are not encouraged to report them". (Akbar, p.10, 2012)

IR teachers appeared to consider authenticity in assessment meant focusing their teaching on particular skills that would be tested like reading and writing. Authenticity for IR teachers also related to their reliance on observation during instruction. These teachers conceived that observation could provide genuine insights into student competence.

> "Without any official test, we know our students' quality [level of competence]…it is not hard because we teach them". (Akbar, p.1, 2012)

Direct observation seemed to be the mainstay of IR teachers and they felt it was unnecessary to record this process.

> "I just keep them in my head, I know some students work hard and some others only cheat. So although in one task/activity they might have the same score, my memory records them differently". (Irma, p.3, 2012)

IR teachers appeared to believe that observation is the best way to capture their students' real proficiency. Another reason for IR teachers' reliance on observation was a perception of their students' low competence. They argued that some students were *'less motivated to learn' (Rahmat, p.5, 2012)*. So when they indicated an interest to participate in a lesson, teachers' observation was needed to record the event. IR teachers were confident that their students' low interest in learning was due to the *'minimum support students get from family' (Rahmat, p.11, 2012),* rather than a reflection of their skills as teacher. This reported phenomenon led them to utilise observation more often than any other formal assessment tool. In addition, their suspicions of students' academic records encouraged teachers to depend on observation as a reliable judgement:

> "I have a student who sits at the third grade but still unable to read...where do the scores in her elementary school report come from? They must be fictitious". (Irma, p.5, 2012)

This teacher felt that it was necessary to gather more authentic assessment data through observation. Consideration about institutional circumstances and students' low competence led IR teachers' to implement traditional teaching approaches.

> "I enjoy teaching structure…I think it is very important". (Akbar, p.3-4, 2012)

And,

> "I use a simple approach… like translating sentences". (Rahmat, p.6, 2012)

Their interest in traditional approaches might explain their preference for using essays or short answer formats which they believed was '*suggested in curriculum particularly in the first and second grade' (Eva, p.10, 2012).* Moreover, IR teachers had limited conceptions of essay formats, choosing to:

> "Request students to change sentences from positive to negative or interrogative…change the subject which requires students to change the verb". (Akbar, p.3-4, 2012)

Examples of traditional assessment practices during classroom activities were recorded in student's assignment books. One IR teacher who shared this document with me required students to translate the procedure of making instant noodles into English. She also asked students to write down the procedure for making a cheese omelette by listing the ingredients, utensils needed and steps to make it along with its translation into Bahasa Indonesia. The teacher marked students' work, printed the date, but did not give any verbal comments. This practice signifies that teachers might consider that scoring is sufficient to record students' competency.

Regarding teacher testing, some IR teachers tested students frequently for example after finishing one basic competency. They felt that having more than one basic competency in a test was quite challenging for students, so they designed an easy or student-friendly test.

> "The programme suggests that we test students after finishing two basic competencies…but it is too hard for students". (Eva, p.9, 2012)

IR teachers used similar formats in their tests as they used for classroom activity. This might explain why some IR teachers did not conduct teacher tests and picked particular classroom tasks to be reported as a teacher test.

> "I do not accumulate several key competencies to be measured in one test". (Rahmat, p.4, 2013)

IR teachers also reported comparable practices in mid-term semester test. It appeared that some IR teachers did not differentiate between teacher tests and mid-term tests. Some of them used class activities as the mid-term test. Other teachers who tested several key competencies were only copying some sections from past semester tests. This suggests that IR teachers might not consider teacher tests to be very important forms of assessment. Moreover, in their student report books, IR teachers presented this official document in a simple way. One teacher recorded only the students' semester score in the report book. She did not present other scores, or the source of the final calculation, or details about the status of KKM achievement. IR teachers appeared to demonstrate indifference towards reporting their assessment results. This practice invites speculation about the authenticity of the score.

Grading for accountability

In their assessment IR teachers utilised scores as they believed that *'students always want to be scored' (Rahmat, p.1, 2012).* Scoring seemed to be a symbol of achievement that could lead to recognition, for example among classmates.

> "Students like it…they are proud to show good score to others". (Rahmat, p.12, 2012)

IR teachers believed that scoring worked well for both high achievers and low achievers. IR teachers agreed that grading in assessment is crucial because the practice informed the position of students against the standards as well as their position among their fellows.

> "When students get a good score, they are motivated to maintain or get more…those who get lower scores will try harder to equal their counterparts". (Akbar, 2, 2012)

These teachers contended that the measurement of students' accountability through scoring illustrated both student achievement and teacher quality:

> "I am proud of myself when I find my students scored high, I feel like I taught them well" (Akbar, 7, 2012)

IR teachers demonstrated a unique technique for grading. They only circled wrong answers with no additional notes and did not check students' correct answers. This practice indicates that IR teachers focused on students' wrong answers. These teachers might expect students to prepare for the remedial programme in which students are given a second chance to do the test but by working only on the wrong responses they gave in the first test. To IR teachers grading is a crucial practice for students and teachers, not only to provide accountability, but also as a way of stimulating students to learn. This emphasis on grading appeared to be an expectation shared among students and teachers.

External assessment

IR teachers' views about external assessments were contradictory. Despite valuing external examinations for assessment purposes, they also argued that external assessment lacked credibility and could be intimidating. This category comprises the themes of conflicting conceptions, teachers' autonomy and the credibility of external assessment.

Conflicting perceptions

Like the IM group, IR teachers regarded external assessment as a measurement tool to capture students' proficiency as well as to signify the school and teachers' quality.

> "National Examination is conducted to measure the quality of students, teachers and school management. When the results are low, it might mean students have low competence or teachers are unqualified". (Rahmat, p.2, 2012)

IR teachers appeared to acknowledge the function of external assessment to signify accountability and to evaluate the quality of education. However, they complained about the process and practice of external assessment. They argued that *'the process is misleading' (Eva, p.6, 2012).* This teacher asserted that students' assessments were not accurate in representing a school

> "It is counterfeit…they [staff of Department of Education] say assessment presents the quality of education in the region but everyone knows that the scores are going to be high…all students must pass the test". (Eva, 5, 2012)

IR teachers reported that to achieve the aim of showing students' competence or the quality of education in the region, schools were stuck in a complicated situation and that sometimes led them to tweak the assessment results to:

> "Preserve their good image…let's say the school is famous with a good reputation, if students' scores are low, the school will receive public attention". (Irma, 2, 2012)

Teachers contended that although they disagreed with the policy, they inevitably complied with the system.

> "The system demands us to do that…the school principal instructs us [to assist students], the school principal is instructed by his/her boss [Head of Education Department] and maybe there are higher level instructions". (Rahmat, 14, 2012)

IR teachers judged that the Department of Education concentrated too much on the region's reputation and '*always wanted to be recognised as successful' (Eva, p.5, 2012).* This phenomenon suggests that each school community, particularly those with a high ranking, wished to maintain their reputation and each school principal was determined to maintain their position even if it meant directing staff to manipulate students' examination results. At the same time, complying with such inconsistencies in the system left teachers with a complicated understanding of assessment.

Teachers' autonomy

The inconsistencies in the process of external assessment diminished its credibility in the eyes of IR teachers and also undermined the teachers' autonomy. These teachers reported systemic interference with the processes for regional and national tests. Unhappily, they testified that

> "Before the examination is co*n*ducted, teachers will be gathered [by the school principal] and instructed to assist students". (Eva, p.5, 2012)

IR teachers felt powerless and desperate. They reported spending time marking students' tasks but

> "All we have done is ignored. We have to follow what the boss [school principal] tells us" (Irma, p.2, 2012)

These extracts imply that school principals use their power to insist teachers perform actions that might contradict teachers' beliefs. Another form of intimidation was also visible in decision-making processes where teachers' voices tended to be ignored:

> "Although in a formal teacher meeting we agreed that one student should stay at the same level, he [the school principal] could change it". (Irma, p.3, 2012)

These teachers perceived that external assessment affected teachers' autonomy negatively; as a result they contended that assessment was meaningless

> "...it makes our assessment useless, students get 4, 5, 3 but at the end we can only present 6.5 and above because it is the KKM. We have to base the results on the KKM!" (Rahmat, p.15, 2012)

IR teachers stated that students whose original score was below the KKM had their grade lifted to meet the standard and those who exceeded the minimum standard would receive an even higher score. To some individuals, this circumstance not only disregarded teachers' autonomy, it also created professional uncertainty.

> "I am not confident to be a teacher; I think I am not a good one". (Eva, p.2, 2012)

These teachers believed that students were aware of their own actual competence but when teachers were encouraged to manipulate final scores, teachers felt that they lost respect from students.

> "I feel like students laugh at me, they think that they will always get assistance". (Eva, p.5, 2012)

The credibility of summative assessment

IR teachers expressed a number of negative views of the examination system such as their requirement to follow orders and their involvement in a conspiracy to manipulate scores. They disagreed with but felt powerless to criticise the system, and felt that they received inadequate information regarding new policies implemented in the region. IR teachers also complained about the school level management system. According to these teachers, the school level curriculum gave each school the right to determine the KKM (passing grade of subject) and control students' achievement. Hopelessly this teacher revealed:

> "The school has the authority to determine the KKM … so when they [school] decide on a certain standard [KKM] they will make it so". (Rahmat, p.12, 2012)

This extract signals that IR teachers retain little faith in external assessment. Teachers reported that they were unable to avoid unfairness in examinations and explained the irony of increasing students' scores so they appeared to meet the KKM so as to assure the school principal and the Education Department of the region that '*the teaching process is on the right track' (Akbar, p.2, 2012).* IR teachers understood that school accountability is determined by students' obtaining scores which are at or above the standards. This teacher explained:

> "The passing grade for our school is 6.5. This means the poorest student gets that score and the best student must get 9" (Rahmat, 8, 2012)

As these teachers were obliged to engage in practices they might disagree with, they discounted the value of external assessment.

> "I do not believe in external assessment; it is fictitious, unrealistic and unbelievable". (Irma, p.8, 2012)

Their disillusionment led IR teachers to believe that assessment results were inaccurate in describing quality including student and school accountability. They believed that schools' autonomy to implement a school level curriculum had compromised schools' legitimacy in determining the KKM and in reporting student scores. Thus, although IR teachers basically agreed with the function of assessment to measure students' proficiency, they doubted the credibility of assessment results. This teacher argued that assessment could not accurately represent school accountability when:

> "We are commanded to progress students to the next level although they do not deserve it". (Eva, p.13, 2012)

Teachers in this group questioned the functions of assessment to describe accountability because they were surrounded by inconsistent policies and practices. Teachers within IR group described this phenomenon as a conspiracy among the school community. The requirement for schools to uphold their good reputation forced these teachers to present only acceptable scores to the wider school community. Teachers testified that scores were fraudulent by insisting that *'we*

make it up' (Rahmat, p.17, 2012). They signalled that they were pressured to improve scores.

> "We are not allowed to use the true scores, If we present these, most students will have to repeat a grade… we are ordered to raise the score". (Eva, p.2, 2012)

Although IR teachers felt trapped in the inflation of student scores, they appeared to be accustomed to these circumstances. Interestingly, they were likely to refer to their own authentic judgement in order to produce the *'proper final score for student reports' (Rahmat, p.18, 2012.)* In other words, the raw but authentic score, along with teachers' daily observations were the initial considerations to decide the appropriate inflation they could give to students. In short, IR teachers questioned the standards used in external assessment. These teachers were uneasy about the policy implemented for both regional and national level exams and also acknowledged that there were likely to be inconsistencies across schools.

> "9 (nine) in my context means six (6) in a town school, I bet". (Rahmat, p.16, 2012)

The above extract indicates that there is a perceived discrepancy between the quality of education in urban and rural areas in the region.

> "If I compare my assessment with the one in the city [capital of the province] where my daughter studies...they are very different". (Eva, p.14, 2012).

Teachers confessed that the semester and the National Examination did not suit rural students because they usually

> "…test some materials that we do not teach, so we have to assist students" (Rahmat, p.13, 2012)

This means that external assessment might use urban standards which were too high for rural students. IR teachers might feel guilty for being unable to complete the curriculum targets. However they could not force their students to obtain competency levels as those mastered by urban students. IR teachers presumed that using one standard for all students in the region was irrelevant.

> "How can I ask them to read or speak English while their reading in Bahasa Indonesia is not fluent?" (Irma, p.5, 2012)

This was yet another area where IR teachers indicated some disagreement over the external assessment system. Their opposition included both regional and national tests. Teachers argued that government involvement in assessment was the cause of problems and inaccuracy of assessment in education.

> "…as long as the government is not involved in the assessment process, it is reliable". (Irma, p.9, 2012)

And

> "We do not need the National Examination, or if the government considers it very important, the government should adjust the test according to local conditions and not use a single national standard". (Eva, p.6, 2012)

These teachers felt that the external assessment lead to 'public deception'. Dissimilarities in students' competencies in different geographical areas, and unequal access to learning made it hard for the community to access similar quality education.

IR case conclusion

The irrelevance group teachers favoured internal assessment through traditional assessment practices. They argued that approaches like direct observation, translating sentences or discussions of sentence structures, and traditional formats such as short-answer, matching and other pencil-paper based tests were more authentic and genuine compared to formal external tests. They agreed that external assessment should describe the accountability of students, teachers and schools but confirmed that this might be wishful thinking. They perceived themselves as powerless to resist government policy regarding high-stakes testing and the pressure from their school principal to raise students' score or to engage in unfair practices. This phenomenon led IR teachers to mistrust the system implemented at school, regional and national level.

Accountability group (AC)

Conceptions of purposes of assessment

Santi, Putri, Naya, & Angga, represent accountability conceptions of assessment. These four teachers put accountability as their first preference, improvement conception the second, and they tended not to hold irrelevance conceptions.

Teachers holding accountability conceptions believed that assessment is a valid means of establishing the accountability of a school or country in doing an educational related job. Accountability suggests that assessment records the ability of teachers and the school to improve the quality of teaching and learning.

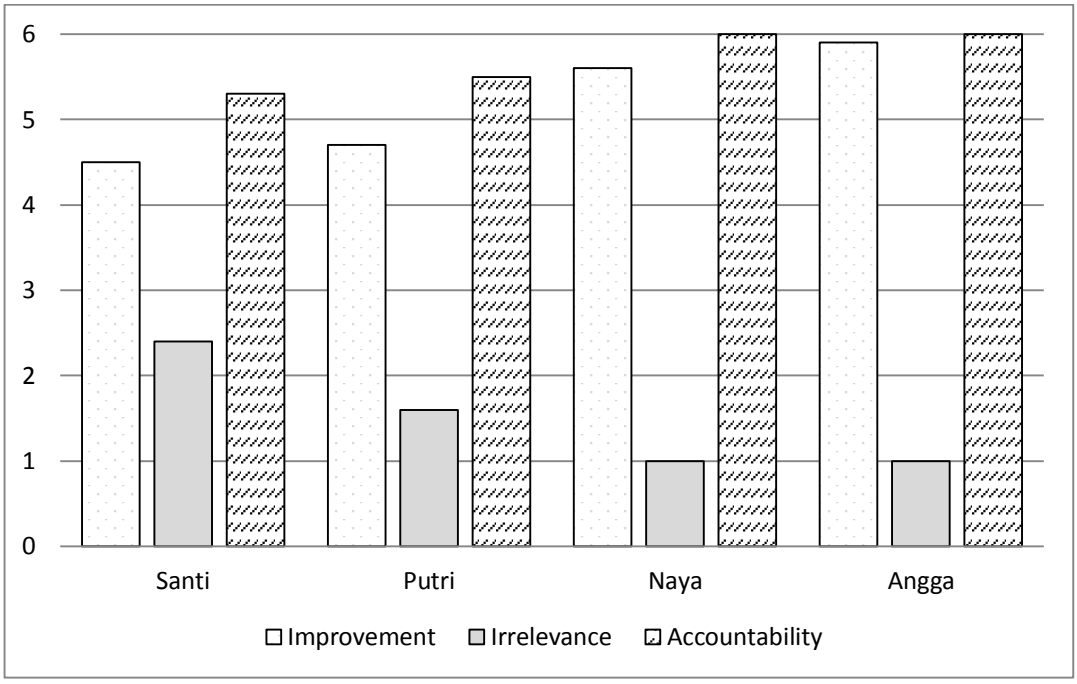


Figure 19: Profile of teachers in the accountability conception case

Similar to the improvement and irrelevance cases, the accountability group also experienced a conflict between their beliefs about internal and external assessment that generated the main themes of the case.

Internal assessment

The category of internal assessment comprises several themes such as openness to change, developing values, mixed understanding of assessment, and teaching resources and grading practices.

Openness to change

Like their colleagues in the IM and IR groups, AC teachers reported using assessment information to modify teaching. AC teachers were flexible and demonstrated an ability to utilise teachable moments:

> "I welcome and apply a sudden bright idea that comes". (Putri, p.8, 2012)

AC teachers reported flexible use of teaching resources for class instruction that was not directly connected to external examinations. Although these teachers stated that the textbooks were their main resource, they adopted and adapted other beneficial resources to support student learning:

> "Although I prepare my teaching journal based on the textbook, students' activities are compiled from many sources". (Putri, p.7, 2012).

AC teachers' use of several teaching materials encouraged them to develop internal tests:

> "I develop teacher tests of my own; sometimes I adapt them from several books". (Santi, p.6, 2012)

AC teachers were willing to adjust their teaching, in response to data from internal assessments such as teacher tests. Looking at the result of tests, AC teachers responded similarly to this teacher.

> "I question myself…I search for the causes of my failure to assist students to achieve the standards". (Putri, p.1, 2012)

To find the answer to this question, AC teachers continuously explored teaching methods or strategies that might make *'the learning process acceptable and comprehensible' (Santi, p.11, 2012)*. In other words, AC teachers understood that internal assessment was a tool that could lead them to be *'effective teachers' (Angga, p.11, 2012).* In a more detailed illustration, AC teachers reported how they adjusted their teaching if it was necessary to revisit parts of the curriculum:

> "When I find the class mastery level is less than 50% I teach particular material again by modifying my teaching like changing the strategy or using teaching aids to make my teaching a bit different". (Putri, 9, 2012)

AC teachers analysed the validity of their teacher tests by listing all test items and how students' responded to them. Through this analysis, teachers measured students' achievement against each item and evaluated students' overall comprehension against competencies tested. This practice signifies that AC teachers also support the improvement function of assessment.

Teachers insisted that the decision to re-teach particular key competencies involved careful thought and might take time. Their willingness to modify teaching signalled an attempt to meet the improvement purposes of internal assessment. Teachers stated that this lesson review was conducted after analysing the test results.

> "I analyse the test per item per student to know which part is hard". (Putri, p.15, 2012).

AC teachers' commitment to rework particular competencies signified an interest to assist students' learning. In other words, AC teachers appeared to agree with internal assessment and its dual function to benefit both teachers and students.

These teachers also felt that internal assessment could improve student-teacher relationships. They tried to approach students individually because they acknowledged that some students tended to be shy and introverted.

> "I usually ask students to write down the materials they find hard, sometimes I ask them to identify which materials they think should be re-taught…but mostly I ask them individually and assist them also individually". (Santi, p.10, 2012)

This shows that AC teachers conducted needs analyses by reviewing a lesson only with students who needed it. The quote indicates that these teachers might try to use students' reviews to ensure students obtain the standard. AC teachers also reported that they returned students' assessment papers to them for further reflection. AC teachers felt that having individual conversation with students was effective where they could illustrate *'some stories or analogies like how a young man approaches a girl' (Putri, p.8, 2012).* AC teachers regarded sharing a success story as a crucial component in teaching because it could inspire students to learn from other people.

Developing values

Similar to IM teachers, AC teachers believed that assessment was beneficial in developing students' characters. They believed that their duty as teacher:

> "Not only aims to educate students [cognitive aspect], we are required to improve their personality and behaviour [affective aspect]". (Santi, p.11, 2012).

This extract indicates an understanding that learning and assessment should focus both on cognitive and affective components. AC teachers perceived that effective

teachers should both assist students to acquire knowledge and in developing good character for future life. This understanding denoted a practice of mixing academic and non-academic components in deciding students' assessment results. AC teachers perceived that assessment facilitated:

> "Students to learn about fairness, they should be confident with their answers". (Naya, 11, 2012)

These teachers encouraged students to practise fairness and honesty during instruction. They believed that continuous application of these values could lead students to become independent learners.

> "I encourage them to stay away from cheating…learning independently is crucial for their future success". (Santi, p.10, 2012)

Teachers also cited other values such as discipline and creativity. These teachers confessed that students' creativity in completing assignments could surprise them.

> "Once I asked my students to write a letter…it was amazing, they are very creative in designing and presenting ideas". (Angga, p.10, 2012).

This teacher indicated an interest in providing more freedom for students in undertaking assessment tasks. She suggested that this strategy could maximise students' potential and motivate students to develop independent learning habits. AC teachers reported how internal assessment influences students' learning. These teachers assumed that assessment or tests were motivational tools for students.

> "Assessment motivates student to study, they become more enthusiastic". (Naya: p.2, 2012)

Thus, like the IM and the IR groups, AC teachers also conceived that positive attitudes towards learning and other values were developed through the practice of internal assessment. However, these teachers were uncertain whether students could learn and implement this value in external assessment settings.

> "Students' fairness in [external] examination is questionable, few of them do tests without cheating". (Angga, p.10, 2012)

Along with the other groups, AC teachers appeared to rate internal and external assessment differently. These teachers suggested that when they were given room

to control students and the assessment process, as they did in classroom assessments, values of fairness or honesty could be developed and implemented. The situation was different in external assessment settings, where teachers did not have access to control students. AC teachers worried about the high priority given to high scores and the strong focus on external examinations. These teachers presumed that unfair external assessment practices influenced student learning and their attitude negatively.

> "Students learn from their surroundings, let's say they observe some unfairness in examinations...so they think 'why should I challenge myself if I don't have to. If I can get 100 without studying, why shouldn't I?' They witness this in practice". (Putri, p.13, 2012)

AC teachers were concerned about this phenomenon. They were afraid that students rationalised dishonest practices in examinations. Inconsistent assessment practices led teachers to another uncertainty.

> "…my assessment is for my school only, I mean when my student gets 80 the value is lower compared to urban school students. However, students who get 80 must be the best ones in my class". (Angga, p.18, 2012).

This extract implies that AC teachers actually were not confident of their own assessment. They might be able to compare students in their classroom against other students they taught but were unsure about larger scale comparisons with other schools on a regional or national scale.

Mixed understanding of assessment

In keeping with their belief that assessment allows for accountability, AC teachers valued both formative and summative assessment practices. In some classroom activities, AC teachers gave students room for free exploration. Their assignments ranged from writing a procedure, to descriptive and recount texts. Teachers asked students to present recipes (procedural text) on a sheet of paper that included the ingredients, methods and pictures of the subject described. For a descriptive text, students described their idol in three paragraphs covering the person's physical appearance, habits, hobbies and activities. Writing a recount text challenged students to recount an incident from their own experience (Appendix F). These illustrate that AC teachers allowed students freedom to choose and indicated an interest in using portfolio tasks to guide students' learning. These teachers were

willing to engage in teaching that culminated in performance based assessment and involve their students in authentic forms of evaluation.

> "I ask students to compose a story, I ask them to make an outline and develop their story based on it [the outline]". (Putri, p.5, 2012)

AC teachers used verbal comments in addition to a score and involved students in peer assessment like evaluating one another's work. However, the practices of this performance based assessment dominated classroom activities only. When teachers wished to evaluate students' mastery of two or more competencies, for example in a teacher test, they preferred to follow assessment practices used in external assessment. AC teachers tended to focus on particular assessment formats which they believed as *'regulation' (Angga, p.7, 2012).* Teachers mainly defaulted to familiar formats that recurred in external examinations like *'essay, short answer and multiple-choice (Putri, p.6, 2012).*

> "Those formats [short answer and essay] are the most popular …besides students can easily find the answer from the text". (Naya, 5-6, 2012)

AC teachers who were assigned to teach the third grade level preferred to use multiple-choice for 'familiarising' reasons.

> "We want students familiar with multiple-choice formats". (Putri, p.15, 2012)

Making students familiar with the final test appeared to be of great concern to AC teachers so they tried to use a test structure that was *'identical to UN [National Examination]' (Angga, p.6, 2012).* To this end AC teachers' understanding of authenticity in teaching and internal assessment which was visible in teacher tests and mid-term semester tests reflected a focus on standardised testing. In other words, AC teachers indicated a mixed understanding of what might constitute authentic assessment.

These teachers mentioned observation as another measurement they used in teaching and assessment, *'we study their actual performance through class interaction' (Santi, p.1, 2012)*. AC teachers rated their own judgements highly.

> "We have to know our students well, it is one key skill for teachers…let's say an ordinary student suddenly gets a very good

> score, it signals inauthenticity; he/she must be cheating". (Angga, p.6, 2012)

AC teachers contended that observation was a pillar of internal assessment. For external assessment, AC teachers reported compliance not only regarding the format but also the materials and skills tested in the exam. Referring to the national exam, AC teachers neglected listening activities during class interaction. Some teachers contended that their school lacked supporting facilities such as a *'[language] laboratory or tape recorder' (Santi, p.7, 2012).* In different schools where this equipment was available, AC teachers complained that they *'never got any training on how to use it [laboratory]'. (Naya, p.5, 2012).*

These comments suggest that AC teachers might believe that effective teaching depends on external factors such as resources. They appeared to attribute any limitations of the curriculum on factors like equipment and facilities that were outside their individual control. This trend among AC teachers contributed to their decisions to narrow the curriculum and may have led to their choice not to teach productive skills like speaking. These teachers felt that speaking skills '*were less important than reading and writing' (Santi, p.8, 2012).* This teacher justified not teaching speaking by explaining that speaking activities did not stimulate students to express their ideas spontaneously:

> "People may call it a "speaking activity" but students actually write down their ideas on paper, memorise then report them". (Santi, p.8, 2012)

Thus AC teachers did not view common speaking activities as truly demonstrating the skill of speaking in English. This inauthenticity may have been a factor in these teachers refusal to utilise such activities in their class. Despite these other reasons, a strong focus on examination preparation seemed to dominate AC teachers' decisions to narrow the content of their teaching:

> "The national exam does not test listening and speaking skills, so I think we don't have to teach these two skills to our students…our main reference is the UN so we only need to base our teaching on it". (Angga, p.5, 2012)

This remark confirms the 'power' of the high-stakes test in AC teachers' understanding of assessment. AC teachers' strong focus on external assessment led teachers of this group to resort to traditional measurements like observation and

encouraged them to concentrate teaching on materials that would be tested in the National Examination.

Grading practices

In keeping with their focus on external assessments, AC teachers believed in the power of scoring in motivating students to learn. AC teachers particularly valued students' proficiency as captured by grades. They believed that a high score *could "increase students' confidence and improve their motivation" (Santi, p.2, 2012)*. These teachers recounted that the same rule could also be applied to students who got low scores.

> When students' get low scores, they are ashamed and reluctant to show the results to colleagues, so they try harder to improve (Angga, p.2, 2012)

With this understanding, teachers believed in the utilisation of scores to describe students' accountability *'a good student gets a good score' (Putri, p.9, 2012).* To AC teachers grading is important because the practice motivates students to learn and promotes values that are believed to be important in future life. They reported that students were eager to be scored in all tasks or assignments. In fact, students tended to ignore tasks when they realised that teachers would not mark them. This condition made teachers warn students in every activity. *'I'm going to check and record this'. (Santi, p.9, 2012).* However, in some cases, they worried about the effect of grading practices:

> "Students only think about how to pass the exam not for knowledge". (Naya, p.11, 2012)

A more extreme example of this was reported by AC teachers from rural areas where students were only interested in external assessment:

> "Students only prepare for the semester or the National Examination; they do not care about teacher test". (Angga, p.2, 2012)

Such students' beliefs suggest that there may have been resistance if teachers tried to use internal assessment to improve teaching and learning. This situation indicates students' understanding that it is only external assessment that counts in the end. This theme affirms AC teachers' beliefs that grading practices strongly influenced students to study and might denote the interrelationship between grades

and assessment beliefs in this community regardless of internal or external assessment.

External assessment

As with the other cases, AC teachers held conflicting conceptions about external assessment and held concerns about teachers' autonomy and the credibility of tests.

Conflicting conceptions

AC teachers' belief in accountability in assessment was obvious in the significance they placed on assessment practice that measured students' proficiency:

> "We want to achieve certain goals [competencies in curriculum or students' achievement level] and only assessment can give us that information". (Santi, p.2, 2012)

Teachers believed that assessment played a key role in describing learning. One AC teacher asked

> "How can we know whether a student is capable or not if we do not assess them?" (Angga, p.9, 2012)

To AC teachers, assessment is a lens through which they can judge the quality and positions of their students.

> "I can see the extent to which my students are able to absorb the teaching; it gives me a reference to recognise their level of proficiency". (Putri, p.13, 2012)

In addition, AC teachers were confident in using students' scores as an indicator of teacher quality. They believed that the score students got also measured their teachers' ability to make students comprehend the lesson. This teacher contended:

> "When students get good scores, it means I am teaching them well". (Santi, p.11, 2012)

Another purpose of external assessment according to AC teachers was to make schools accountable.

> "The school quality is determined by the National Examination". (Naya, p.11, 2012)

AC teachers understood that the National Examination was very important because the reputation of the school and its community was at stake:

> "If a school has many students who have high proficiency, it means the school is qualified". (Santi, p.3, 2012).

This teacher was not alone in stating that a school's reputation is measured by its students' scores. AC teachers regarded the high-stakes examination as essential because it informed the school's position or ranking:

> "…the government gets the information or report [about the quality] of which region or province is the best for this year". (Putri, p.2, 2012)

To this end, AC teachers demonstrated a consistent understanding of accountability. They trusted the efficacy of testing to describe quality and understood the function of standardised tests. In addition, AC teachers believed that school accountability '*depends on teacher[s'] quality' (Putri, p.9. 2012).* In other words*, 'qualified teachers make a school accountable' (Angga, p.9, 2012).* These extracts show how external assessment illustrated accountability of students, teachers and schools in the minds of AC teachers.

In spite of supporting external assessment, AC teachers argued that it was unfair. These teachers believed that the semester test was less valid or less credible because there was a mismatch between what was taught and what was tested *'they (Department of Education) test different competencies (Angga, p.12, 2012).* This teacher felt that the local government used city standards and disregarded rural and suburban student competencies. A comparable misgiving was also expressed about the National Examination that AC teachers presumed was equally unbalanced:

> "I disagree a bit that the UN (National Examination) has become the only parameter to determine graduation. The exam does not describe an on-going process [of learning]. Fine if the UN becomes the benchmark, but it is better to have a balance between teacher, school and national assessment". (Putri, p.16, 2012).

These teachers questioned the government's policy to only test particular subjects like Bahasa, English, Mathematics and Science. AC teachers viewed this policy as discrediting the other seven subjects.

> "UN does not test all subjects, it only tests four…there are students who are good at sport or arts but they are not recognised because

> the government does not acknowledge them [as] important". (Santi, p.3, 2012)

This statement described concerns about the utilisation of UN as the single high-stakes examination in the country. These teachers considered that the final examination could not portray a complete picture of student proficiency. This remark is consistent with previous comments on the need to have balanced measurement of students' cognitive and affective competence. However, it also signals a complicated perception of accountability.

Teachers' autonomy

AC teachers reported that external assessment whether conducted regionally or nationally impacted negatively on their autonomy. These teachers raised several arguments against the external examination including the Education Department's apparent distrust of teachers' competence in composing tests. They also expressed a feeling of being intimidated. Teachers particularly resented the administration of semester tests by the regional Education Department feeling that this tended to undermine teachers' professional confidence.

> "I feel like they [education department] question our capacity…they do not trust us to manage our own test". (Angga, p.12, 2012)

This policy seemed particularly unjust because semester tests were supposed to be categorised as teacher tests.

> "Most teachers question the exclusion of teacher in the development of semester tests". (Naya, p.3, 2012).

AC teachers challenged the regional policy regarding the management of semester test. They felt excluded from the practice and asserted that the regional Education Department misunderstood the national government's intended process for assessment.

> "I think they [the regional Education Department] are disobeying the rule...as far as I'm concerned 'government assessment' means assessment which is conducted by the Ministry of Education in the country, not the Education Department of a region". (Angga, 16, 2012)

This teacher understood that the MoNE expected teachers to construct semester tests for their own students. In addition, AC teachers reported another form of intimidation caused by this new test management.

> "After the semester examination is over, we are gathered in one particular school to check students' work…but the score we report in the students' report book is not the authentic score, we manipulate it". (Angga, p.18, 2012)

And

> "We are commanded to do that [raise students' scores]…the school principal told us that students' score must reach the standard…" (Santi, p.9, 2012)

Thus, teachers were ordered to follow their principal's directives. This phenomenon made these teachers powerless and they felt the '*need to follow what we are told to' (Naya, p.4, 2012)* even though these were unfair practices. Consequently, AC teachers' lack of support for this regional external assessment was caused by their exclusion from the process and the abuse of authority by school principals and the local government. AC teachers argued that the semester test would be more credible if it were returned to the control of classroom teachers.

> "I think the semester test should not be handled by the education department. Students are familiar with their teachers' test and this can help us achieving the fairness value. I reckon that fairness or honesty has become a scarce value to get since the semester test has been managed by the education department of the region". (Naya, p.13, 2012)

AC teachers complained that they received similar intimidation after the National Examination. These teachers reported that schools were so concerned about maintaining their reputation in this high-stakes examination that school principals often compromised teachers' autonomy.

> "As a leader of the school, the school principal manages to show that he/she is a good and successful manager. He/she tells his teachers about his/her intention [to have good scores] some teachers disagree with this idea but it is the school policy, so… [Hang-we have to follow it]". (Putri, p.10, 2012)

This implies that teachers were required to report good scores to assure school quality. It also confirmed the strong role of the department of education in using its power to persuade teachers and schools to perform such actions.

> "We are intimidated by the school principal; the school principal is intimidated by his boss. In a formal forum the head of the department of education explains 'last year our region got this position…can't we make it better?'...To us, the meaning is clear...they actually say 'do whatever you can to make it [the result of the national exam] look good". (Putri, p.15, 2012)

This indicates that teachers feel powerless to defend their own autonomy in the face of power shown by local and regional departments. This phenomenon influenced AC teachers' to lower their view of the credibility of assessment to meet the purpose of showing accountability. This teacher pessimistically contended that:

> "The result of UN cannot describe school accountability. Not in this region or anywhere else, well it should... but the reality tells a different story [laughing-sounding pessimistic]". (Naya, P.11, 2012)

Her comment illustrates a despondent acceptance that external assessments are less than credible. AC teachers worried that dishonest practices were becoming widespread and were concerned about the negative impact this was having on students and public trust. Sadly, this teacher bemoaned the lack of concern shown in the wider community:

> "The government, schools, teachers, parents and all elements in the community should work together to overcome this problem… however, only teachers show the greatest concern. Parents are happy with the score, they are proud when their child gets 100, they do not care about how their child gets it". (Putri, p.14, 2012)

Evidently, it was clear that even AC teachers disputed external assessment practices and their results.

> "I am not confident with the credibility of our [external] assessment". (Santi, p.9, 2012)

Teachers were uncertain about external assessment because the process and the results signified a conspiracy among the school community:

> "Everybody knows that students' won't be able to pass the examination without teachers' assistance". (Naya, p.11, 2012)

This realisation confounded these teachers despite their beliefs in the accountability purposes of the external examination. Even in situations where teachers were not directly involved in unfair practices, they acknowledged that other elements also played a role.

> "I am happy but disappointed…I mean students complain about the hard lessons but get 100 in the exam". (Putri, p.9, 2012)

These teachers were curious about the way students achieved high scores; they were suspicious about the involvement of other institutions in undermining the credibility of assessment practices.

> "They get the answer key from tutoring schools [external institutions that specifically teach examination technique]". (Putri, p.11, 2012)

This theme illustrated complex perceptions of external assessment. AC teachers' scepticism about external assessment practices appeared to be consistent across this case. The AC teachers were unhappy about the attack on teachers' autonomy, inequality, and conspiracies among school community that were visibly practiced in the context and which undermined the value of assessments.

AC case summary

AC teachers were positive that assessment should be a reflection of teaching and learning. They demonstrated enthusiasm in using assessment results to inform their teaching and make their teaching effective. They were supporters of internal assessment. However, they presented different perceptions about external assessment. Although they understood that external assessment aimed to account for the effectiveness of students, teachers and schools, they were disappointed by the implementation of external assessment in their educational setting, which affected students, schools, and the local Education Department and other external institutions.

These teachers believed that unfair examination practices decreased students' motivation for learning. They also contended that students' motivation had become skewed because students focused on obtaining a good score without considering the process. In other words, AC teachers expressed concern that those positive aspects of assessment might be compromised. AC teachers contended that principals' efforts to maintain their reputation led the school community to conduct unfair assessment practices which resulted in an ambiguous or unclear function of assessment as a means of accountability.

Chapter Five

Discussion of findings and conclusion

Introduction

In this chapter, I situate my analysis of intermediate school teachers' conceptions of assessment as expressed by a multiple case study of educators teaching in the Gowa district of South Sulawesi, Indonesia. The findings from this study are discussed within the context of the international literature on conceptions of assessment. This chapter concludes by exploring limitations, implications of this study, and recommendations for further research.

The complexity of teachers' conceptions of assessment

Teachers' conceptions and practices of assessment are contextually dependent processes that operate under the influence of multiple factors. The participants in this study reported conflict between assessment practices that they wanted to employ versus those demanded by authorities and the educational community. This research has revealed more complex components and diverse influences than previous models of conceptions proposed by earlier researchers such as Brown (2002). Brown identified three types of teachers' conceptions of assessment (TCoA) known as the improvement, the accountability and the irrelevance conceptions. These conceptions were categorised based on a meta-analysis of results reported in the international literature and were originally created to capture the conceptions of New Zealand primary school teachers. Brown's (2002) categories assumed that there were clear distinctions among the three conceptions. Teachers holding improvement conceptions agreed that the purpose of assessment is to improve teaching and learning. Those with accountability conceptions held a belief that assessment should be used for the purpose of external accountability, and teachers with irrelevance conceptions tended to view assessment as irrelevant to the work of teachers and to the life of students. Although my selection of participants was influenced by the intention of selecting participants holding a preference for one of these conceptions, my study revealed that these Indonesian teachers did not hold any particular category independently of the others. Every teacher held conceptions of assessment incorporating aspects of all three categories: improvement,

accountability and irrelevance. The results of this study reveal complex interrelationships of assessment conceptions. These findings also suggest that assessment is more than a confrontation between internal and external assessment, as previously proposed by Black and Wiliam (1998a) and Earl (2003). In my study, contributing factors interweave to construct bi-dimensional perceptions of assessment.

The methodology used in this study may have contributed to the fundamental differences reported here in comparison to previous studies of teachers' conceptions (Brown, 2002; Calveric, 2010; Segers & Tillema, 2011). In prior studies, conceptions of assessment were analysed using a single quantitative method which revealed similar findings: improvement conceptions were preferred over accountability conceptions, and teachers tended to disagree with irrelevance conceptions. This finding was replicated in the quantitative component of my research, but the qualitative analysis of individual perspectives of the teachers in each group suggested that this broad categorisation was too simplistic, at least for the Indonesian context. Whether this is so for teachers working in other national educational contexts, is a subject for future mixed methods research.

The three groups of teachers were not exclusively different, but shared similar perceptions that assessment functioned to improve teaching and learning as well as to signify accountability. Unlike teachers in other studies, they also presumed assessment could simultaneously be irrelevant. This complex understanding was clearly captured in face to face interviews where participants had the opportunity to share their reasons for this complexity. My study raises the possibility that teachers in previously published research may also have held complex conceptions; however their voices were not captured within a single methodology approach.

Results of this study reveal more than localised conceptions of assessment that differ from those reported in the published international literature. The complex nature of assessment beliefs also allows for a discrepancy between teachers' expectations and their practices. Thus, one hypothesis may be that the conceptions held by participants in my study were influenced by the Indonesian culture, educational system, assessment policies, and teacher resources. The significance of some of these factors has been raised in earlier studies. The influence of cultural factors was found to affect teachers' beliefs and practices in China, Hong Kong, and

Korea (Cheng, 2008; C.-C. Choi, 1999; I.-C. Choi, 2008). The educational system and assessment policies have also been reported in numerous studies conducted by Barnes, Clarke, and Stephens (2000), Winterbottom et al. (2008), and Remesal (2009). Other influencing factors such as teaching materials and characteristics of students were also reported as important in Yueming, Eslami, and Burlbaw's (2006) study. The relationship among various factors contributing to teacher conceptions of assessment is worth consideration, and this complexity of assessment conceptions suggests that Brown's (2002) categorisation may not be transferable to different cultural contexts.

The results of this study show this complexity principally relates to socio-ecological factors (Bronfenbrenner, 1979, 1994). My study focused on the professional life and work of teachers and enabled me to isolate the influencing factors from the macro level down to the micro level: cultural, contextual and personal. Teachers' complex perceptions and the practices and factors contributing to them are presented in the following Figure.

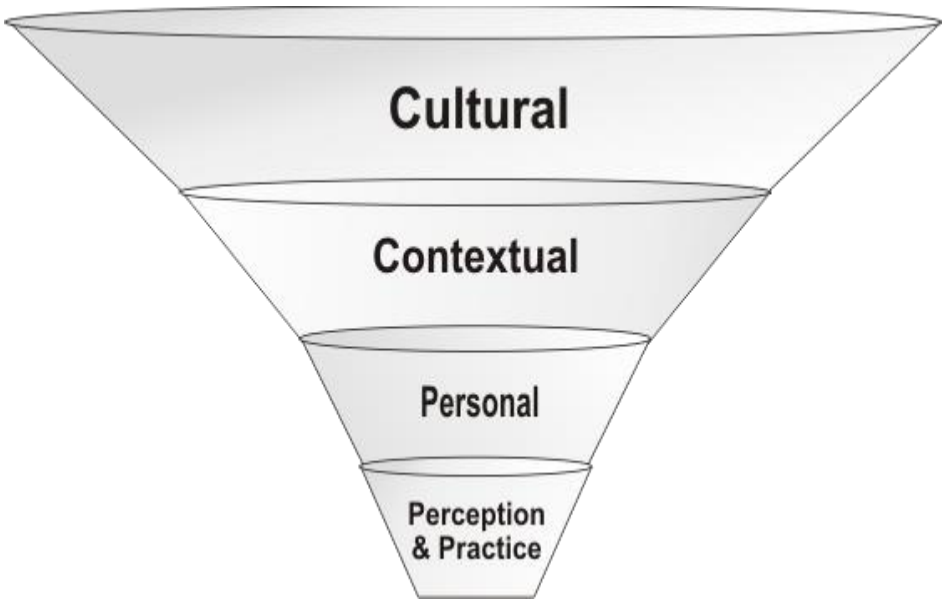


Figure 20: The interrelationship of Indonesian teachers' conceptions and practices of assessment with socio-ecological factors

This model illustrates the interactions of factors that encompass cultural, contextual and personal influences on Indonesian teachers' conceptions and practices of assessment. In this conceptual framework, factors like competition, testing and a

testing and grading culture constitute a macro system that influences the localised, national and regional contexts through institutional regulations and policies. These two layers in turn become the guide and reference to form and inform teachers' personal conceptions and practices of assessment. To understand Indonesian teachers' conceptions of assessment, then, one needs to look at these surrounding contributing factors.

Cultural factors

Cultural factors in this study denoted shared habits and beliefs regarding the practices of testing, competition and the assumption that scores are the best indicators of students' learning. Participants of the study agreed that providing assessment grades was the most effective strategy to stimulate students' learning; students were hungry for grading. The spirit of competition among students included aspiring for high grades that demonstrated their accountability and suitability for selection into further courses of education. Equally, teachers believed that grading in assessment was important because it could be used to measure student achievement and to evaluate the effectiveness of the teaching programme. Thus, teachers and students alike were motivated by grading practices (Remesal, 2011). This perception supports Brookhart's (1994) conclusion of her analysis of nineteen studies on teachers' grading practices. She claimed that grading was the most common practice of educational measurement and had become an integral part of classroom instruction both in the UK and the USA. Similar perceptions of the efficacy of grading, testing and competition shared among parents, students and community have been reported by other studies as significantly influencing social status (I.-C. Choi, 2008) as well as being the best indicator of success (Cheng, 2008).

Allen (2005) argued that grading lacks validity because it focuses more on teachers' expectation of what good students are, rather than measuring students' academic mastery of the subject matter. However, my study disagrees and suggests that teachers' long experience with grades or students' and parents' familiarity with the practice had given them an 'understanding' that grades not only inform about academic achievement, but also indicate students' efforts and motivation for learning.

This communal understanding confirms McMillan, Myran, and Workman's (2002) conclusion that the practice of grading could be interpreted in multiple ways covering knowledge and academic enabling behaviours. Grading may also affect students' attitudes as suggested by this study as well as attendance patterns (Friedman & Frisbie, 2000). In a context like Indonesia where there are large class sizes, grading could be seen as a means to lighten teachers' classroom management (Cheng & Wang, 2012). Grading practices in my study support these findings. The cultural and contextual elements in my study appear to illustrate a chicken and egg situation in that a shared understanding of grading, testing and competition may have led the country to establish a system that reflects community values. Alternatively, the multiple purposes of education in Indonesia including developing good citizenship (Jalal, Ramly, & Harianti, 2011) may also be responsible for the culture of grading, testing and competition. Nonetheless, it is clear that culture and context are the macro level factors contributing to Indonesian teachers' conceptions of assessment.

Contextual factors

Contextual factors include national, regional and institutional policies and regulations. Firstly, it is crucial to consider the strong top-down management system in Indonesia. The national education system influences regional/district policies; regional policies in turn affect school policies and objectives for teachers. Thus teachers' perceptions and practices of assessment were heavily controlled by the authorities. Bjork (2004) claimed that in Indonesia, 'the nation's system of government has exerted a powerful influence on the work of teachers' (p.134) that their work was responsible primarily to government requirements rather than to students and parents. The national policy through state-wide mandated standardised testing forced schools and teachers under their jurisdiction to adopt and comply with the policy. The National Examination is conducted to monitor the quality of education across the country, with results ranked nationally. Provincial and regional government strongly encourage schools to perform well and prepare for the national examination. When the ministry targets a pass rate of 98% students, this is a signal for provinces and districts to set similar or even higher targets.

Compliance with a National Examination system reflects the culture of the Indonesian people (Bjork, 2006) and resulted in two main impacts upon teachers: a focus on the national exam and an attack on teachers' autonomy. My study

illustrates how the National Examination overshadowed and distorted teachers' conceptions and use of assessment. IR and AC teachers only taught the skills tested in the National Examination; all participants in this study familiarised students with examination formats during classroom assessment and offered intensive programmes dealing exclusively with the tested subjects and test taking strategies before the exams. These practices clearly signalled that the National Examination has led teachers to focus on testing rather than the assessment that they believed better met the needs of students. This finding is consistent with other empirical studies including those conducted by Yueming, et al. (2006), Au (2007), and Saw (2010).

Findings from a study of teachers in four districts in Houston, Texas, USA., Yueming, et al. (2006) revealed that high-stakes assessments narrowed the curriculum by educational authorities requiring teachers to teach and assess only specific subjects and objectives to be covered in the test, particularly in the two months prior to the test. This finding was replicated in Au's (2007) meta-synthesis of forty-nine qualitative studies of high-stakes or state-mandated assessment. Saw (2010) argued that the standardised National Examination forced teachers to adopt teaching methods designed for test preparation in order to assist students to score highly. In other words, although these studies do not report against Brown's TCoA, it seems that wherever and whenever a standardised National Examination is implemented, it becomes the main reference for teaching and assessment practices. This might explain why the TCoA model is less than accurate in reporting the conceptions held by teachers working in high-stakes examination contexts like Indonesia.

These contextual factors not only expose the dangers of narrowing the curriculum, they also signify the erosion of teachers' power in the decision-making process in such an educational system. The circumstances are even more severe in the context of my study where teachers sometimes were required to collude in dishonest practices. Despite accepting the need for accountability of teacher and school, teachers felt intimidated and compromised by the control exerted by external forces.

Political drivers appeared to have a strong influence upon external accountability. A serious concern about ranking procedures or quality mapping by the Ministry of

Education (Ministry of National Education, 2005b, 2011) possibly led the regent and the head of education of the Gowa district to reassure the community that the subsidised education programme did not reduce the quality of education in the region. School principals were required to support the goal, and at the end of the day teachers had no choice but to agree with the principal's instruction to assist students in the examination. This phenomenon led teachers to feel guilty and lose face in front of students. The attack on their autonomy was clear to teachers. These teachers believed that inflating students' examination results was a demand, and they felt obliged to comply and raise scores. This exploitation of power over those in lower positions is reminiscent of Bronfenbrenner's (1979) remarks about power.

> "The greater the degree of power socially sanctioned for a given role, the greater the tendency for the role occupant to exercise and exploit the power and for those in a subordinate position to respond by increased submissions dependency, and lack of initiative". (p.92)

Inevitably, the more powerful authority influences the practice of classroom assessment. Teacher-driven assessments were highly influenced and controlled by the schools. This finding echoed Yueming, et al.'s (2006) study where teachers were powerless against school and district policies. In my study, teachers were encouraged to report that student results (score) met the minimum level of the expected performance that was established at schools regardless of their true score. Thus, obtaining the standard in this context meant meeting accountability standards for the school and even maintaining the reputation of the district.

This imbalance between teachers' and government roles in assessment placed teachers in a weak and unfavourable position to the extent that the safest way forward for teachers was to comply with the system. They responded to this situation by questioning the credibility of assessment. Some teachers argued that students' performance in the examination was suspiciously different from their daily performance in the classroom, and such manipulations of students' scores made it hard to interpret students' real performance. The conspiracy and unfairness in semester tests and the National Examination appeared to have reached a critical phase where the government disregarded the main goals of external assessment: to examine the effectiveness of the course (Nation & Macalister, 2010) or to evaluate the quality of education. Assessment in my study simply symbolised a routine check-up conducted to maintain the educational reputation of a school or districts and province or to report that the educational sector is under control.

Personal factors

A teacher's personal and professional resources may affect their understandings of assessment. Teachers who are certified or those who teach the ninth grade (the last year of intermediate school) may hold dissimilar perceptions to those held by uncertified lower grade teachers. In Indonesia, the National Examination for intermediate or middle school is conducted at Year 9. Teachers' conceptions appeared to change under the pressure of assessment as a tool for certification for students. This finding aligns with those in other studies. Barnes, et al. (2000), for example, found that secondary teachers at junior level (Years 7-10) favoured school-based assessment, whereas teachers of grades 11-12 focused on preparing students for Year 12 assessment. Other studies revealed similar findings particularly due to the different educational structures between primary and secondary schools with examination for certification at secondary level only (Bonner & Chen, 2009; Remesal, 2011). Findings from these studies revealed that primary school teachers tended to follow constructivist views of learning and perceived that assessment was a tool to improve teaching and learning, whereas secondary school teachers adopted behaviourist views and were concerned more with assessment for accountability purposes (Brown, 2002; James, 2008). My study reveals that even in the same educational structure (junior high school/intermediate level), dissimilar conceptions existed among teachers teaching at non-examination and examination years.

Teachers held different conceptions of student competency, and they made teaching adjustments in accordance with their notions of student ability. Teachers tried to contextualise the rhythms and scope of the teaching materials to meet their students' proficiency level. In this respect, my study supports Lambs' (2012) argument that outside factors affect teacher and student motivation. In his study of Indonesian junior high school students, he found that urban and rural students had different levels of motivation and English proficiency. Students from these different geographical areas also received different levels of support from the family (Lamb, 2013). Thus, teachers' perceptions of their students' proficiency may be dependent on their unique setting.

To sum up, the cultural, contextual and personal components evident in my participants and represented in my conceptual framework denote the differentiating socio-ecological factors that impacted on these Indonesian teachers'

understandings of assessment. The complexity depends on the culture, the educational system, the structure of organisation, school policies and resources of teachers and students. In this study, teachers' conceptions of assessment were found to be heavily influenced by culture and context, which persuaded them to comply with the system. They appeared to hold what Remesal (2011) termed societal conceptions. A societal conception is a view of assessment as being an extrinsic motivation device, a tool to measure achievement and provide an effective communication instrument to parents, students and the wider community. Yet, this finding from my study differs from Remesal's: the participants in my study held even more complex and extreme societal conceptions, which overrode their beliefs in the monitoring purposes of assessment. This perception in turn led my participants to question the trustworthiness of assessment. In other words, teachers concurrently held all three elements of Brown's (2002) TCoA; improvement, accountability and irrelevance. These participants' conceptions are thus unique in that the qualitative findings are inconsistent and conflict with those from the quantitative phase where improvement conceptions received the highest mean score.

Indonesian teachers' perceptions and practices of assessment

Participants' bi-dimensional perceptions and practices of assessment result from these socio-ecological influences. Figure 21 illustrates the bi-dimensional perceptions of internal and external assessment and compliance and defiance assessment practices. I use the term bi-dimensional to signify coexisting but inconsistent conceptions of assessment which existed within every teacher rather than only across teachers within a particular case cluster. As teachers revealed conflicting conceptions based on their consideration of internal (teacher-driven assessment) and external assessment (examination-based assessment), their assessment practices combined elements deriving from both. I call this compliance and defiance practice and subdivide this into five main components: authentic vs. reproductive practice, strong reliance on textbook vs. using additional teaching materials and the last component is grading practices.

Conceptions and practices are interrelated as denoted by the two-headed arrow. The strong arrow connecting conceptions to external assessment indicates teachers' focus in this assessment. The compliance and defiance practices consist of three main sets: IM, IR and AC. Each set was built up by several elements and

intersections that allow classifications and comparisons amongst the participants' practices of conducting assessment.

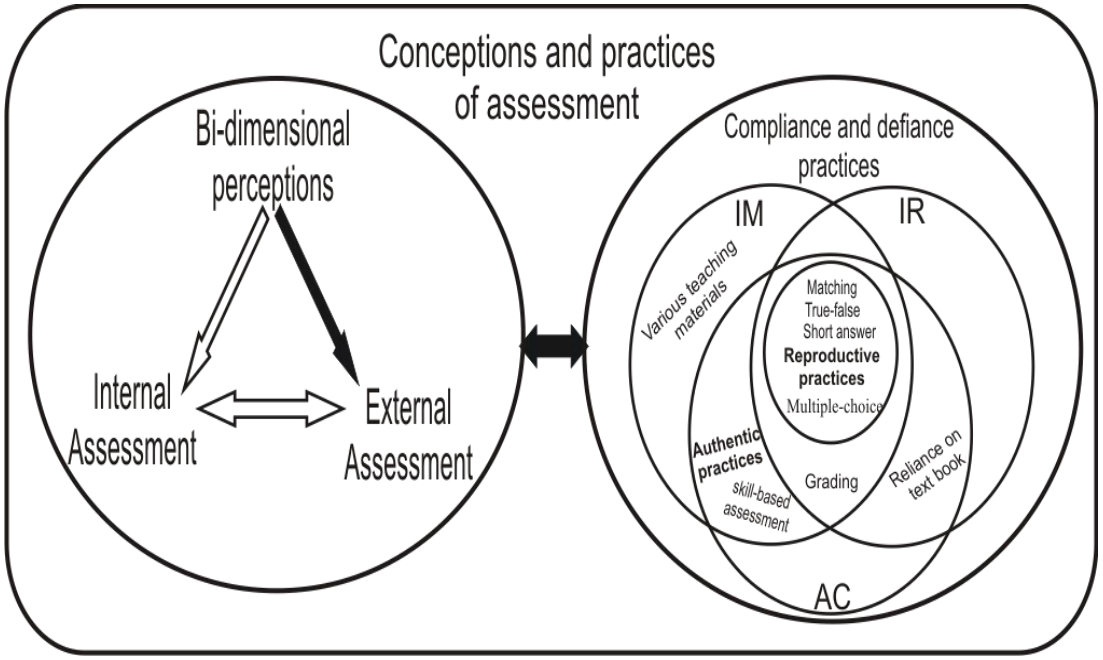


Figure 21: Indonesian teachers' perceptions and practices of assessment

Internal assessment refers to teacher driven assessment during or after instruction which is conducted for quality assurance; this assessment is also called classroom assessment (CA) (Harianti, 2005). To assure quality, teachers reported conducting CA both to evaluate the process and results of learning. This understanding was aligned with the MoNE expectation (Ministry of National Education, 2005c) which recommended evaluation through classroom tasks, teacher tests, a mid-term semester test and semester tests. This means that teachers' perceptions of CA reflected the published fitness of purposes of the assessment system in Indonesia. The focus of internal assessment shared between students as well as teachers might suggest a similarity to the improvement conception (Brown, 2002) or a pedagogical conception (Remesal, 2007, 2011). The degree of agreement across these different conceptions types is highlighted further in the next section.

Classroom assessment with external accountability focus

The Indonesian Ministry of Education suggested several strategies for classroom assessment including performance tests, paper and pencil tests, oral tests,

observation, portfolios, and self-assessment (Harianti, 2005). Therefore, the intention of the Indonesian model of classroom assessment is to monitor the on-going achievement of students and to summarise achievement at a particular time (Harlen, 2005; Segers & Tillema, 2011).

This conceptions of classroom assessment is similar to the exclusively normative use of assessment in the USA (Shepard, 2000b). In the normative use of assessment, students usually do tasks and perform 'to please the teacher or to get good grades rather than to pursue a compelling purpose' (Shepard, 2000b, p. 31). This argument perfectly illustrates the condition of classroom assessment in the context because teachers believed that students' understanding of teaching materials was reflected in their grades. They believed that the higher the grade, the better the quality of teaching, so higher grades reflected more effective teachers. When interpreted through such conceptions, the intended purposes of classroom assessment were forgotten. Shepard (2000b) contended that the compelling purpose of classroom assessment was to find shared understanding between teachers and students on what makes sense and what doesn't, and this could be used to design learning strategies. In other words, in their assessment practices, teachers in the context of the study appeared to only focus on what works rather than on what doesn't. This finding implies teachers were interested in evaluating the results of teaching rather than the process of learning.

Nonetheless classroom assessment policy was set by the MoNE and teachers simply performed their role as policy implementers. Indeed, the Indonesian classroom assessment model appeared to only partially support the empirical prototype of classroom assessment suggested by Resnick and Resnick (1992) or Airasian (1991) who proposed that classroom assessment should focus on teachers and students. Participants' classroom assessment evaluated instruction but results were rarely used to identify students' personal needs; classroom assessment seemed to only inform instructional changes but not students' learning. My participants supported internal assessments and demonstrated a preference for policies in which their assessment practices promoted classroom tasks and teachers tests. Teachers argued for a version of classroom assessment in Indonesia focused on process where they could use and develop several assessment strategies to improve teaching. This understanding reflects formative purposes of assessment where teachers could continuously make changes during

instruction (Boston, 2002) in order to improve teaching quality and students' learning (Black & Wiliam, 1998a; Harris & Brown, 2008).

However, my participants were constrained to evaluate and summarise students' achievement at particular times. These assessment practices included traditional assessments like paper and pencil tests with items covering multiple choices, true/false, matching, fill in /completion, short answers and essay tests. Any teaching adjustments were included in a remedial programme made after the teacher test which involved re-teaching the same materials with or without significant changes in strategies or giving students another chance to re-do the tests. This conception and practice contradicts earlier interpretations of classroom assessment (Resnick & Resnick, 1992; Stiggins, 1999). These authors contended that changes in teaching instruction should be conducted in day to day operation in order to maximise the diagnosable function of assessment to improve teaching and learning.

Teachers' confidence in the impact of CA on teaching was interesting because they tended to differentiate between the purposes of assessment for teaching and learning, components that to some scholars are inseparable (Black & Wiliam, 1998b, 2007; Stiggins, 1999). However, the findings from my study indicated a different understanding of connections between modifications to teaching and a possible improvement in learning. These teachers may have agreed in theory with the relationship between teaching and learning as explained in formative assessment (FA) but they interpreted these factors differently in practice. Teachers appeared to concentrate on one aspect only (teaching) assuming that teaching would automatically impact the other side (learning). This reasoning shared by participating teachers might signal the need for further investigation into the impact of assessment on teaching and learning. In other words, there is potential to separate the purposes of assessment for teaching and learning into two different components. A clear classification of the intended purposes of assessment such as its impact on teaching and learning separately is needed to avoid confusion among teachers (Frey & Schmitt, 2010). This idea (which suggests a need for PD) seems not to have been considered by policy makers in Indonesia. Therefore, my participants' conceptions that classroom assessment impacts on teaching and thus on learning could not truly be categorised as formative assessment (FA) or assessment for learning (AfL).

Assessment for learning according to Black and Wiliam (1998b)(1998b) should inform students so they can adjust their learning strategies. In the context of my study, teachers perceived that the effect of assessment was to improve teaching. They felt that improving teaching assessment would also enhance student learning. The Indonesian model of internal assessment appears to perform primarily a summative purpose whereby teachers used teacher-made tests to generate regular grading for record gathering (Harlen, 2005). Again, the use of grading in classroom assessment contradicts Mueller-Joseph (2007) who argues that classroom assessment should not determine students' grades as it is supposed to assist teachers to understand and improve learning or to diagnose students' knowledge of the topic (Tinajero & Hurley, 2001). In other words, classroom assessment should provide students with educative feedback for encouragement rather than rating their performance (Angelo & Cross, 1993). It should also continuously inform teachers about the effectiveness of their teaching (Mueller-Joseph, 2007). The limited synergy between the purposes of classroom assessment raised by scholars and the interpretations reported by my participants indicates that useful and meaningful classroom assessment again relates to the need and characteristics of teachers, students and settings to which they are applied (Angelo & Cross, 1993). Therefore the effectiveness of classroom assessment is context-specific, meaning that what works well in one context may not necessarily work in another.

Preferences for assessment practice: Compliance and defiance

The results of my study reveal that all participants favour reproductive assessment like a paper-pencil test. IR teachers were the strongest supporter of this assessment type and this may indicate their compliance with the system. Their compliance led them to use reproductive assessment over authentic forms of performance/skill-based assessment or oral tests. The latter assessment type was only demonstrated by IM teachers and did not directly relate to the external examination. Though AC teachers also belong to the same intersection, this group only partially support authentic practices as they focused primarily on measuring students' performance for writing skills. This phenomenon might denote that there is still space for teachers to interpret the Indonesian model of classroom assessment flexibly according to their interests and also their perceptions of students' proficiency. Although all teachers used reproductive practices, IM and AC teachers combined them with authentic practices like assessing students according to each skill. However, IR teachers and some AC teachers who clearly support the reproductive assessment

reported students' low competence as a hindrance factor. In contrast, IM teachers did not feel that their students' competence was a reason to limit the scope of their assessment practices.

Reproductive assessment practices relying on the use of pencil and paper were a means of familiarising students with external assessments (Frey & Schmitt, 2010). This definition meets Postareff and colleagues (2012) category of a reproductive conception. Several assessment formats like multiple-choice, matching, true/false, short answer/fill-in-the-blank and essay questions (Frey & Schmitt, 2010) were among the popular reproductive test formats used by my participants. These teachers argued that the emphasis on external accountability led them to depend on this assessment type. In order to assure their compliance, IR and AC teachers strongly relied on textbook activities published by the ministry which claimed to contain and represent all suggested key competencies of the curriculum. In other words, teachers did not design and create tailored classroom assessment activities; they selected tasks from textbook activities and past tests.

Interestingly, some defiance was demonstrated by IM teachers. This group of teachers were interested in non-reproductive assessment that includes considering students' performance of a skill or judging student product or their participation in a learning process (Brookhart, 1999). The aims of measuring skills or ability are sometimes called alternative or authentic assessment (Frey & Schmitt, 2010). IM teachers strove to combine reproductive and performance testing in their classroom assessment. Student tasks ranged from portfolio assignments such as creating posters, writing a summary of TV shows, or telling narrative stories and demonstrating procedures. Teachers also reported an interest in peer assessment. This assessment model encourages students to create, construct or respond to questions or prompts (Butler & Mc Munn, 2006). IM teachers' interest in performance/skill-based assessment signalled an interest in measuring student skills promoted by the previous competency based curriculum. In the 2004 competency-based curriculum, teachers were trained and encouraged to develop communicative competence through the teaching of text types including transactional conversations, interpersonal conversations and short functional texts (Agustien, 2006). In this curriculum, a genre-based approach was implemented following oral and written cycles and four stages of learning. In other words, IM teachers might be more comfortable with a competency-based curriculum as they

preferred that assessment practices be more authentic. In spite of their adherence to the prevalent emphasis on preparing students for examination, IM teachers' assessment practices featured a degree of defiance to the system as well as rebellion against total dependency on policy requirements.

AC teachers appeared to negotiate suitable assessment practices for students to a lesser extent. These teachers also considered authentic assessment in classroom tasks although they only focused on writing production. AC teachers tended to demonstrate an interest in 'modern' assessment by offering students an alternative to the traditional paper-pencil test. However, their use of this assessment model for written activities indicated their strong focus on external accountability. IR teachers, on the other hand, did not indicate any willingness to adapt and adopt performance-based assessment and totally complied with traditional tests in familiar formats.

Teachers' understandings of assessment and their assessment practices appeared to affect their choice of suitable teaching resources. AC teachers who used alternative assessment also reported using other teaching resources in addition to the textbook. The IM group in particular searched for materials related to listening and speaking activities like audio cassettes or DVDs that could support their teaching. Both IM and AC teachers were also interested in using authentic materials like magazines, newspapers and recycled packages to complement writing activities from the textbook. This contrasted with IR teachers who supported traditional forms of assessment, reported a strong reliance on the textbook and focused exclusively on reading and writing.

Credibility of assessment

Participants in my study indicated that internal assessments both for formative and summative purposes were more credible than external assessment because teachers were given authority to develop and control their assessment classes. In other words, teachers placed a high value on classroom activities because they realised that the process allowed them professional autonomy and enabled them to align assessment to their teaching. Moreover teachers also revealed that these internal assessments were more trustworthy than external tests, and when seriously implemented, they could contribute to teaching reasoning and values as suggested in the character education programme. My participants' confidence about the credibility of classroom assessment was related to teachers' beliefs that they were

the most trustworthy assessors of their students. According to Tinajero and Hurley (2001), teachers are the closest assessors enabling them to immediately measure and respond to students' progress and achievement. The idea that internal assessment allowed a degree of autonomy for teachers appeared to comfort participants in the study; teachers felt that they had the authority to manage their classes consistently. However, this was in reality a conditional freedom. The small portion of their teaching devoted to internal assessment led teachers to prioritise reproductive assessment and therefore limit the use of authentic assessment practices. Classroom assessment practices were dominated by paper-pencil tests meaning that other classroom assessments like performance-assessment, oral questions or portfolio were used less.

A strong focus on external assessment on the other hand represents a belief that summative or examination-based assessment functions to evaluate the curriculum, account for students learning and control teachers' practices. This conception focuses on an institution's and a society's demand for assessment. However, participants of the study were uncertain about the true purposes of external assessment and tended to refute and be sceptical of this assessment type. In the context of the study, external assessment comprised the semester test that was administered regionally and the state-wide standardised National Examination. Despite their familiarity with these forms of external assessment, teachers reported that the accountability purposes of external assessment were compromised by the excessive control of the regional government and schools' mission to produce acceptable results of assessment. This created conflicting assumptions among teachers that external assessment was a less credible and less reliable measure of student, teacher and school accountability than they wanted to believe it to be. This conflict suggests that the results of external assessment might not accurately describe the quality of education in the country.

To sum up, participants of this study superficially demonstrated an interest in FA or AfL in their classroom practices but a deeper examination of their beliefs revealed a closer affinity with summative assessment (SA). Even though these perceptions may seem inconsistent, they fit the purposes of the existing assessment system in Indonesia (Ministry of National Education, 2005a) that focuses more on external assessment. In other words, teachers' blurred perceptions that did not fully fit within parameters set by international scholars (Airasian, 1991; Black & Wiliam, 1998a;

Resnick & Resnick, 1992) may have developed as a result of the Indonesian education system. Therefore, although participants in this study reported that the purpose of assessment was to inform teaching and improve students' learning, their understandings do not precisely align with an improvement conception (Brown, 2002), a pedagogical conception (Remesal, 2011), or a transformative conception (Postareff, et al., 2012). My participants' perceptions of assessment fit more closely within a societal or mixed societal conception (Remesal, 2011) or to accountability conceptions (Brown, 2002). At the same time, unlike participants in other studies, participating teachers also balanced improvement, pedagogical and irrelevance conceptions.

Limitations

This study presents significant information pertaining to teachers' conceptions of assessment. The use of a mixed method design in this study offers new and comprehensive ways of understanding teachers' conceptions of assessment, particularly where there appear to be contradictions in how teachers view assessment. Nevertheless, there are limitations that need to be considered particularly with reference to future research needs.

Instrument

The first limitation of the study relates to the questionnaire used in the first phase. The questionnaire was not designed for use in an Indonesian context which meant that it was challenging to capture the main ideas from the TCoA in Bahasa Indonesia. As a result, there might be dimensions that are not fully captured in the Bahasa version despite my collaboration with other Indonesian scholars in an effort to gain a close translation. This version may not adequately assess participants' beliefs and practices; therefore, further revision of the TCoA might be beneficial if future scholars wish to use it in Indonesia. This study used two self-reported data sets; a questionnaire and an interview to elicit teachers' perceptions of assessment. In my study, participants' assessment practices were derived from the document analyses. Assessment documents that were derived from teachers, students, school and the ministry allowed some triangulation of teachers' assessment practices.

Sampling

Another limitation of this study comes from the selection of participants. Only junior high school teachers of English from Gowa district who were actively involved in a PD programme participated in this study. I cannot say with confidence that the sample is representative of the entire population of junior high school English teachers in the province (Creswell, 2002). This limits the generalisability of the findings because in the decentralised education system other regions might implement dissimilar policies such that different interpretations and understandings of assessment might exist. However, my participants taught in geographically dissimilar areas (urban-suburban, rural and even remote areas) and purposive sampling enabled me to capture voices from different educational and geographical contexts. More importantly, my mixed methods sampling design aimed to obtain insights into the assessment phenomenon and meaningful understandings about this underlying phenomenon within a specific context (Onwuegbuzie & Collins, 2007). This study was not designed for the purpose of generalisation.

Implications of this study

While this study has some limitations and further research would be needed for transferability, the findings have implications for policy makers in Indonesia. One important area is inconsistent government policy to manage the decentralisation of the education system. In this decentralised reform, teachers are encouraged to develop curriculum and assessment without adequate guidance, and are compromised because the government appears to exercise a virtual monopoly over school practices. Indonesian teachers strongly depend on MoNE; they are accustomed to await instruction from the central government to show their compliance. Rather than being curriculum developers, they tend to be only curriculum implementers. Policy makers should attempt to understand teachers' perceptions, knowledge, readiness and responses to any planned reform otherwise the results could be mixed if not useless. This finding supports Bjork's (2004) claim that a strong top-down authority system in Indonesia indicates that the country is not a fertile setting for reforms involving large actors for the management of public services. A decentralised system with school level management in which local staff and the school community are encouraged to manage their own affairs appeared to require a longer time to be effective. Policy makers need to explore what is crucial in the reform rather than following the 'fashion' of decentralisation and/or shifting to

classroom assessment. It is evident that the policy makers were not clear about the concept of classroom assessment which resulted in teacher confusion and inconsistent understandings of this assessment. When Government realises that teachers, students and the community are not ready to shift from summative assessment to formative assessment it is counter-productive to force teachers to implement new strategies that are not well understood. Providing administrators with sufficient training might support new policy. In addition, it may be prudent for central government to allow more room for regional and, provincial institutions to grow bigger and healthier rather than maintaining the dependency of these smaller contexts.

However, if classroom assessment is the real focus of assessment reform in the country, teachers should be more empowered in their capacity as the central and closest assessors for students. Their knowledge about classroom assessment; FA or AfL should be developed through PD. MoNE should revise their methods of training teachers, which is usually conducted within a very short time before the implementation of new policies. District and school administrators could run workshops about assessment skills and strategies to improve teachers' assessment literacy. Supporting teachers and providing them with materials and other resources could also encourage the use of classroom assessment. This suggests the need for financial support, and consideration of the effectiveness of the proposed reform. This includes a needs analysis for the appropriate reform and an awareness of teacher proficiency. A clearer and more balanced focus between internal and external assessment may be necessary for the future of assessment in the country. More importantly, such a focus may address teacher disillusionment about decreased levels of professional autonomy and facilitate the function of assessment to enhance student learning. In this way the purposes of assessment could be clearly illustrated to inform both teachers and students.

Direction for future research

This study presents a first step towards investigating teachers' conceptions of assessment in Indonesia and provides a starting point for complementary research. The study captures teachers' complex understandings of assessment that have developed due to multiple factors. Investigating the beliefs of other stakeholders in education like parents, school administrators and educational department staff could

add further information. These sources of data would enable greater perspective and reflect a wider picture of the shared assessment values among the community. Consideration of using observation could also confirm and support participants' self-reported data.

My study has focused on junior high school teachers of English in the Gowa district. Replication of the quantitative survey with a larger population of teachers of other subjects, other levels and who are located in other parts of the province and the country might allow wider comparisons. This could provide insights about whether contextual factors influenced different teachers in similar ways and whether they perceive assessment in the same way as participants in this study. Finally, this study introduces a new framework for studying teachers' conceptions of assessment. Further research using the suggested components in the new conceptual framework would enable researchers to extend the model of TCoA to better fit their own context.

Concluding Statement

This study has provided insight into the issue of teachers' conceptions of assessment. The major contribution of this study is providing a model to understand conceptions of assessment as a complex process. Figure 20 and Figure 21 indicate that teachers' conceptions are formed and informed by multiple factors. This framework for teachers' assessment conceptions is unlikely to apply uniquely to teachers of the Gowa district of South Sulawesi Indonesia and no others. Contexts with similar educational, contextual, or cultural elements might reveal similar perceptions. More importantly, the conceptual framework encourages researchers to continue searching for a suitable model that fits a specific context rather than following a particular international model. The use of a mixed methods design for this study allowed for a more complete understanding of the research problem, the study's trustworthiness, and interpretations of the findings.

The results of my study show that factors surrounding the teaching context are interrelated threads woven to form participants' conceptions. This study demonstrates that even an internationally validated survey with a number of replicated studies of assessment conceptions might not be equally valid if used in different educational contexts. One important lesson from this study may be that

one size does not fit all. The TCoA might be suited to other settings like New Zealand or Australia due to the relatively low-stakes examination context. However, when it was tested in the high-stakes examination context of Indonesia, the TCoA appeared to require adjustment. Any future measurement of teachers' conceptions of assessment should take account of socio-ecological factors in order to better capture teachers' conceptions. Understanding teachers' conceptions of assessment and the factors that influence these conceptions are essential foundations for the implementation of effective policy.

References

Abelson, R. P. (1979). Differences between beliefs and knowledge system. *Cognitive Science, 3*, 355-366.

Adamson, K., O'Brien, M., & Pasley, K. (2007). An ecological approach to father involvement in biological and stepfather families. *Fathering, 5*(2), 129-147.

Agustien, H. I. R. (2006). *Genre-based approach and the 2004 English curriculum*. Paper presented at the UPI National Seminar, Bandung, Indonesia.

Airasian, P. W. (1991). Perspective on measurement instruction. *Educational Measurement: Issues and Practice, 10*(1), 13-16.

Allen, J. D. (2005). Grades as valid mesures of academic achievement of classroom learning. *The Clearing House, 78*(5), 218-223.

Amrien, A. L., & Berliner, D. C. (2002). The impact of high-stakes test on student academic performance: An analysis of NAEP results in States with high-stakes test and ACT, SAT, and AP test results in States with high school graduation exams. *Education Policy Studies Laboratory*.

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers*. San Francisco: Jossey-Bass Publishers.

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher, 36*(5), 258-267. doi: 10.3102/0013089x07306523

Azis, A. (2012). Teachers' conceptions and use of assessment in students learning. *Indonesian Journal of Applied Lingusitics, 2*(1), 41-51.

Barnes, M., Clarke, D., & Stephens, M. (2000). Assessment : The engine of systemic curricular reform. *Journal of Educational Studies, 32*(5), 623-650.

Bergman, M. M. (2008). The straw men of the qualitative-quantitative. In M. M. Bergman (Ed.), *Advances in mixed methods research* (1st ed., pp. 11-21). London: Sage Publication.

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: 10.1080/096959X.1010.527701

Bjork, C. (2004). Decentralisation in education, institutional culture and teacher autonomy in Indonesia. *International Review of Education, 50*(3-4), 245-262.

Bjork, C. (2006). Transferring authority to local school communities in Indonesia: Ambitious plans, mixed results. In C. Bjork (Ed.), *Educational Decentralization: Asian Experiences and Conceptual Contributions* (pp. 129-147). Dordrecht, The Netherlands: Springer.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2002). *Working inside the black box: Assessment for learning in the classroom*. London: Department of Education and Professional Studies, Kings College.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education : Principles, Policy & Practice, 5*(1), 7-74. doi: 10.1080/0969595980050102

Black, P., & Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment*. London: Kings College.

Black, P., & Wiliam, D. (2007). Large-scale assessment systems design principles drawn from international comparisons. *Measurement, 5*(1), 1-53.

Bonner, S. M., & Chen, P. P. (2009). Teacher candidates' perceptions about grading and constructivist teaching. *Educational Assessment, 14*(2), 57-77.

Borko, H. (2004). Professional development and teacher learning : Mapping the terrain. *American Educational Research Association, 33*(8), 3-15.

Boston, C. (2002). The concept of formative assessment. *Practical Assessment, Research & Evaluation, 8*(9).

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Broadfoot, P. (2007). *An introduction to assessment*. New York: Continuum International

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*, 513-531.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1994). Ecological model of human development *International Encyclopedia of Education* (2 nd ed., Vol. 3, pp. 1643-1647). Oxford: Elsevier.

Brookhart, S. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education, 7*(4), 167-183.

Brookhart, S. (1999). The art and science of classroom assessment. *ASHE-ERIC Higher Education Report, 27*(1), 35-57. doi: 10.1002/aehe.3640270104

Brookhart, S. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practices, 30*(1), 3-12.

Brown, G. T. (2002). *Teachers' conceptions of assessment*. Dissertation. University of Auckland, NZ. Auckland.

Brown, G. T. (2003). *Teachers' instructional conceptions: Assessments' relationship to learning, teaching and teacher efficacy*. Paper presented at the Conference of the Australian and New Zealand Associations for research in Education (AARE/NZARE), Auckland, New Zealand.

Brown, G. T. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. New York: Nova Science Publishers.

Brown, G. T., Hui, S. K. F., & Yu, F. W. M. (2010). *Teachers' conceptions of assessment: Developing a model for teachers in Hong Kong*. Paper presented at the Biannual conference of the International Test Commission.

Brown, G. T., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conception and practices of assessment. *Assessment in Education: Principles, Policy & Practice, 16*(3), 347-363. doi: 10.1080/09695940903319737

Brown, G. T., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education, 27*, 210-220.

BSNP. (2006). *Standar isi untuk satuan pendidikan dasar dan menengah: Standar kompetensi dan kompetensi dasar SMP/MTs*. Jakarta: Badan Standar Nasional pendidikan.

Bumen, N. (2009). Possible effects of professional development on Turkish teachers' self-efficacy and classroom practice. *Professional Development in Education, 35*(2), 261-278.

Butler, S. M., & Mc Munn, N. D. (2006). *A teacher's guide to classroom assessment: Understanding and using assessment to improve student learning*. San Francisco: Jossey-Bass.

Butterfield, S., Williams, A., & Marr, A. (1999). Talking about assessment: Mentor-student dialogues about pupil assessment in initial teacher training. *Assessment in Education, 6*(2), 225-246.

Calderhead, J. (1996). Teachers: Belief and knowledge. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 709-725). New York: McMillan.

Calveric, S. B. (2010). *Elementary teachers' assessment beliefs and practices.* Doctor of Philosophy dissertation, Virginia Commonwealth University, Virginia.

Chan, J. K.-s. (2007). *We have various forms of assessment but only summative assessment assessments count: Case studies of the implementation of an innovative assessment policy in Hong Kong*. Paper presented at the Redesigning Pedagogy-Culture, Understanding and Practice, Singapore.

Cheng, L. (2008). The key to success: English language testing in China. *Language Testing, 25*(1), 15-37.

Cheng, L., Rogers, W. T., & Wang, X. (2008). Assessment purposes and procedures in ESL/EFL classrooms. *Assessment & Evaluation in Higher Education, 33*(1), 9-32. doi: 10.1080/02602930601122555

Cheng, L., & Wang, X. (2012). Grading, feedback, and reporting in ESL/EFL classrooms. *Language Assessment Quarterly, 4*(1), 85-107. doi: 10.1080/15434300701348409

Choi, C.-C. (1999). Public examinations in Hong Kong. *Assessment in Education, 6*(3), 405-417.

Choi, I.-C. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing, 25*(1), 39-62.

Christians, C. (2005). Ethics and politics in qualitative reasearch. In D. N.K & L. Y.S (Eds.), *The handbook of qualitative research* (3rd ed.). Thousand Oaks: Sage Publication.

Cohen, D. K., & Hill, H. C. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record, 102*(2), 294-343.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education*. New York: Routledge.

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddler River: Pearson Education.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* Thousand Oaks: Sage Publications.

Creswell, J. W. (2005). *Educational research planning: Conducting and evaluating quantitative and qualitative research* (2nd ed.). New Jersey: Pearson Education.

Creswell, J. W. (2007). *Qualitative Inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks: Sage Publications.

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks: Sage Publications.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks: Sage Publications.

Creswell, J. W., Plano Clark, V. L., & Garret, A. L. (2008). Methodological issues in conducting mixed methods research design. In M. M. Bergman (Ed.), *Advances in mixed methods research* (1st ed., pp. 66-83). London: Sage Publications.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in the behavioural and social sciences* (pp. 209-240). Thousand Oaks, CA: Sage Publications.

Danielson, C. (2008). Assessment for learning: For teachers as well as students. In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 191-213). New York: Taylor & Francis.

Degbey, M. (2009). *Using assessment to improve teaching and learning in upper secondary schools*. master thesis. University of Turku. Turku.

Delanshere, G., & Jones, J. H. (1999). Elementary teachers' beliefs about assessment in mathematics: A case of assessment paralysis. *Journal of Curriculum and Supervision, 14*(3), 216-240.

Denzin, N. K. (1989). Interpretive interaction. In G. Morgan (Ed.), *Beyond method: Strategies for social research* (pp. 129-146). Beverly Hills, CA: Sage Publications.

Dole, S., Nisbet, S., Warren, E., & Cooper, T. J. (1999). Teacher collaboration in developing rich assessment task in mathematics as a professional development activity. *Mathematics Teacher Education and Development, 1*, 38-49.

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximise student learning*. Thousand Oaks: Corwin Press.

Erzberger, C., & Kelle, U. (2003). Making inferences in mixed methods: The rules of integration. In A. Tashakkori & C. Teddlie (Eds.), *Handboook of mixed methods in social and behavioral research* (pp. 457-489). Thousand Oaks: Sage Publications.

Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: the effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis, 20*(2), 95-11. doi: 10.3102/01623737020002095

Fowler, L. T. S., Banks, T. I., Anhalt, K., Hinrich Der, H., & Kalis, T. (2008). The association between externalizing behavior problems, teacher-student relationship quality, and academic performance in young urban learners. *Behavioral Disorders, 33*(3), 167-183.

Frey, B. B., & Schmitt, V. L. (2010). Teachers' classroom assessment practices. *Middle Grades Research Journal, 5*(3), 107-117.

Friedman, S. J., & Frisbie, D. A. (2000). Making report cards measure up. *The Education Digest, 65*(5), 45-50.

Gipps, C., McCallum, B., & Hargreaves, E. (2000). *What makes a good primary school teacher? Expert classroom strategies*. London: Routledge Falmer.

Gowa, P. K. (2011). Kondisi geografis dan demografis kabupaten Gowa Retrieved 19 July, 2011, from http://www.gowakab.go.id

Gowa, P. K. (2012). Generasi emas Gowa keluaran SKTB Retrieved 8 February 2013, 2012, from www.gowakab.go.id

Grant, S. G., Derme-Insinna, A., Gradwell, J. M., Pullano, L., Lauricella, A. M., & Tzetzo, K. (2002). When increasing stakes need not mean inscreasing standards: The case of the New York state global history and geography exam. *Theory and Research in Social Education, 30*(4), 488-515. doi: 10.1080/00933104.2002.10473208

Gray, D. E. (2009). *Doing research in the real world* (2nd ed.). London: Sage Publications.

Green, T. F. (1971). *The activities of teaching*. New York: McGraw-Hill.

Gregory, K., & Clarke, M. (2003). High-stakes assessment in England and Singapore. *Theory into Practice, 42*(1), 66-74.

Griffiths, T., Gore, J., & Ladwig, J. (2006). *Teachers' fundamental beliefs, commitment to reform, and the quality of pedagogy*. Paper presented at the Australian Association for Research in Education Annual Conference.

Grunbaum, N. N. (2007). Identification of ambiguity in the case study research typology: What is a unit of analysis? *Qualitative Market Research: An international Journal, 10*(1), 78-97. doi: 10.1108/13522750710720413

Guba, E. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology: A journal of theory, research & development, 29*(2), 75-91.

Guba, E., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park: Sage Publications.

Guskey, T. R. (2003). How classroom assessment improve learning. *Educational Leadership, 60*(5), 6-11.

Guthrie, J. T. (2002). Preparing students for high-stakes test taking in reading. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 370-391). Newark: International Reading Association.

Hamel, F. L. (2003). Teacher understanding of student understanding: Revising the gap between teacher conceptions and students' ways with literature. *Research in The Teaching of English, 38*(1), 49-84.

Hammersley, M. (2008). Troubles with triangulation. In M. M. Bergman (Ed.), *Advances in mixed methods research* (1st ed., pp. 22-36). London: Sage Publications.

Hanson, W. E., Creswell, J. W., Plano Clark, V. L., Petska, K. S., & Creswell, J. D. (2005). Mixed methods research designs in counselling psychology. *Journal of Councelling, 52*(2), 224-235. doi: 10.1037/0022-0167.52.2.224

Hargreaves, A., Earl, L., & Schmidt, M. (2002). Perspectives on alternative assessment reform. *American Educational Research Journal, 39*(1), 69-95. doi: 10.3102/00028312039001069

Harianti, D. (2005). *Model penilaian kelas* Jakarta: Departemen Pendidikan Nasional.

Harlen, W. (1998). Classroom assessment: A dimension of purposes and procedures. In C. Ken (Ed.), *SAMEpapers* (pp. 75-97). Hamilton New Zealand: Center for Science, Mathematics and Technology Educational Research, University of Waikato.

Harlen, W. (2005). Teachers' summative practice and assessment for learning-tensions and synergies. *The Curriculum Journal, 16*(2), 207-223.

Harlen, W. (2007). *Assessment of Learning*. London: Sage Publications.

Harris, L., & Brown, G. T. (2008). *New Zealand teachers' conceptions of the purposes of assessment: Phenomenographic analyses of teachers' thinking*. Paper presented at the Australian Assiciation for Research in Education (AARE), Brisbane, Australia.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Hesse-Biber, S. N. (2010). *Mixed methods research : Merging theory with practice*. New York: The Guilford Press.

Hewson, C. (Ed.) (2006) The SAGE dictionary of social research methods. London: Sage Publications.

Hill, M. (2000). *Remapping the assessment landscape: Primary teachers reconstructing assessment in self-managing schools*. doctoral dissertation. University of Waikato. Hamilton, NZ.

Isaac, T. (2010). Educational assessment in England. *Assessment in education: Principles, policy and practices, 17*(3), 315-334. doi: 10.1080/0969594X.2010.491787

Isaksson, U. (2013). Validity and reliability testing of the Swedish version of Melbourne decision making questionnaire. *Scandinavian Journal of Caring Science*, 1-8. doi: 10.1111/scs.12052

Ivankova, N. V. (2004). *Students' persistence in the University of Nebraska-Lincoln distributed doctoral programme in educational leadership in higher education: A mixed methods study.* Doctor of Philosophy, University of Nebraska, Lincoln.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequestial explanatory design: From theory to practice. *Field Methods, 18*(3), 2-20. doi: 10.1177/1525822X05282260

Jalal, F., Ramly, M., & Harianti, D. (2011). *Panduan pelaksanaan pendidikan karakter (Guide to character education)*. Jakarta: Pusat Kurikulum dan Perbukuan Balitbang.

James, M. (2008). Assessment and learning. In S. Swaffield (Ed.), *Unlocking assessment*. New York: Routledge.

Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches* (3rd ed.). Thousand Oaks: Sage Publications.

Kennedy, K. J., Chan, J. K. S., Fok, P. K., & Yu, W. M. (2008). Forms of assessment and their potential for enhancing learning: Conceptual and cultural issues. *Educ Res Policy Prac, 7*, 197-207. doi: 10.1007/s10671-008-9052-3

Khan, S., & VanWynsberghe, R. (2008). Cultivating the under-mined: Cross-case analysis as knowledge mobilization. *Forum: Qualitative Social Research, 9*(1), 1-21 article 34.

Lamb, M. (2012). A self system perspective on young adolescents' motivation to learn English in urban and rural settings. *Language Learning, 64*(4), 997-1023. doi: 10.1111/j1467-9922.2012.00719.x

Lamb, M. (2013). 'Your mum and dad can't teach you!': Constraints on agency among rural learners of English in the developing world. *Journal of Multilingual and Multicultural Development, 34*(1), 14-29. doi: 10.1080/01434632.2012.697467

Leech, N. L., Dellinger, A. B., Brannagan, K. B., & Tanaka, H. (2009). Evaluating mixed research studies: A mixed methods approach. *Journal of Mixed Methods Research, 4*(1), 17-31. doi: 10.1177/1558689809345262

Li, W. S., & Hui, S. K. F. (2007). Conceptions of assessment of mainland China college lectures: A technical paper analyzing the Chinese version of COA-III. *The Asia Pacific-Education Researcher, 16*(2), 185-198.

Lickona, T. (1991). *Educating for character: How our schools can teach respect and responsibility*. New York: Bantam Books.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4-16.

Ma, H.-I., Hwang, W.-J., & Chen-Sea, M.-J. (2005). Reliability and validity testing of a chinese-translated version of the 39-item parkinson's disease questionnaire (PDQ-39). *Qual Life Res, 14*, 565-569.

McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practices, 20*(1), 20-32.

McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research, 95*(4), 203-213.

McMillan, J. H., & Schumacker, S. (2010). *Research in Education Evidence-Based Inquiry* (7th ed.). New Jersey: Pearson.

Merriam, S. B. (1988). *Case study research in education: A qualitative approach*. San Francisco: Jossey-Bass.

Merriam, S. B. (2001). *Qualitative research and case study application in education*. San Francisco: Jossey-Bass.

Mertens, D. M. (2003). Mixed methods and the politics of human research: The transformative-emancipatory perspective. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 135-164). Thousand Oaks: Sage Publications.

Mertens, D. M. (2005). *Research and evaluation in education and psychology. Integrating diversity with qualitative, quantitative, and mixed methods* (2nd ed.). London: Sage Publications.

Ministry of National Education. (2005a). *Indonesian Government Policy Act No. 19. 2005: Standard of Education*. Jakarta: Ministry of National Education.

Ministry of National Education. (2005b). *Peraturan menteri pendidikan nasional Republik Indonesia nomor 20 tahun 2005 tentang ujian nasional tahun pelajaran 2005/2006 Minister of National Education regulation no. 20 year 2005 on National Examination for academic year of 2005/2006* Jakarta: Biro Hukum dan Organisasi Kemdiknas.

Ministry of National Education. (2005c). *Peraturan pemerintah Republik Indonesia nomor 19 tahun 2005 tentang standar nasional pendidikan Government regulation number 19 year 2005 on national standards of education*. Jakarta: MoNE.

Ministry of National Education. (2008). *Lampiran permen diknas No. 20 tahun 2007 tentang standar penilaian pendidikan : Attachment and explanation of The Minister regulation no. 20/2007 on educational assessmentthe standards*. Jakarta: Badan Standar Nasional pendidikan.

Ministry of National Education. (2011). *Tanya-jawab pelaksanaan ujian nasional 2011 FAQ about national examination 2011*. Jakarta: Kementrian Pendidikan Nasional.

Mueller-Joseph, L. (2007). Classroom assessment. *Journal of Security Education, 2*(4), 47-53.

Muijs, D. (2011). *Doing quantitative research in education with SPSS* (2nd ed.). London: Sage Publications.

Nation, I. S. P., & Macalister, J. (2010). *Language curriculum design*. New York: Routledge.

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Curriculum Studies, 19*(4), 317-328.

Noor, N. B. M., Muniandy, M. K., Krishnan, S. K., & Mathai, E. J. (2010). Upper primary teachers' perceptions of PSLE English oral assessment. *English Language Teaching, 3*(4), 142-151.

Onwuegbuzie, A. J., & Collins, K. M. T. (2007). A typology of mixed methods sampling design in social science research. *The Qualitative Report, 12*(2), 281-316.

Onwuegbuzie, A. J., & Tedddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 351-383). Thousand Oaks: Sage Publications.

Ozerk, K., & Whitehead, D. (2012). The impact of national standards assessment in New Zealand, and national testing protocols in Norway on indigenous schooling. *International Electronic Journal of Elementary Education, 4*(3), 545-561.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307-332.

Philippou, G., & Christou, C. (1997). Cypriot and Greek primary teachers' conceptions about mathematical assessment. *Educational Research and Education, 3*(2), 140-159.

Postareff, L., Virtanen, V., Katajavuori, N., & Lindblom-Ylanne, S. (2012). Academics' conceptions of assessment and their assessment practices. *Studies in Educational Evaluation, 38*(4), 84-92.

Pratt, D. D. (1992). Conceptions of teaching. *Adult Education Quarterly, 42*(4), 203-220. doi: 10.1177/074171369204200401

Ravitch, D. (1995). *National standards in American education: A citizen's guide*. Washington DC: Brookings.
Remesal, A. (2007). Educational reform and primary and secondary teachers' conceptions of assessment: The Spanish instance, building upon Black and Wiliam (2005). *The Curriculum Journal, 18*(1), 27-38. doi: 10.1080/09585170701292133
Remesal, A. (2009). *Spanish students teachers' conceptions of assessment when starting their career*. Paper presented at the Perceptions and conceptions of assessment in the classroom: Different national perspectives, Amsterdam.
Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment : A qualitative study. *Teaching and Teacher Education, 27*, 472-482.
Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessment: Alternative views of aptitude, achievement, and instruction* (pp. 37-75). Boston: Kluwer Academic Publishers.
Ryan, K., & Bohlin, K. E. (1999). *Building character in schools: Practical ways to bring moral instruction to life*. San Francisco: Josey-Bass
Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education, 5*(1), 77-84.
Saw, L. O. (2010). Assesment profiles of Malaysia: High-stakes external examinations dominate. *Assessment in Education: Principles, Policy & Practice, 17*(1), 91-103. doi: 10.1080/09695940903319752
Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation, 37*(1).
Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2), 63-75.
Shepard, L. A. (2000a). The role of assessment in learning culture. *Educational Researcher, 29*(7), 4-14. doi: 10.3102/0013189X029007004
Shepard, L. A. (2000b). The role of assessment in teaching and learning (G. S. o. E. I. S. Center for the Study of Evaluation, Trans.) *CSE Teachnical report 517* (pp. 1-85). Los Angeles: University of California
Singh, K. (2007). *Quantitative social research methods*. New Delhi: Sage Publication India.
Smith, M. L., & Fey, P. (2000). Validity and accountability in high-stakes testing. *Journal of Teacher Education, 51*(5), 334-344. doi: 10.1177/0022487100051005002
Stake, R. E. (2000). Case studies. In D. N.K & L. Y.S (Eds.), *Handbook of qualitative research* (2nd ed., pp. 435-454). Thousand Oaks: Sage Publications.
Stake, R. E. (2010). *Qualitative Research: Studying how things work*. New York: Guilford Press.
Stiggins, R. J. (1999). Assessment, students, confidence, and school success. *Phi Delta Kappan, 81*(3), 191-198.
Sudibyo, B. (2007). *Lampiran peraturan menteri pendidikan nasional republik Indonesia Nomor 20 tahun 2007 tetang standar penilaian pendidikan*. Jakarta: Departemen Pendidikan Nasional Retrieved from http://bsnp-Indonesia.org.
Suhardi, D. (2009). *Buku saku kurikulum tingkat satuan pendidikan (KTSP) sekolah menengah pertama*. Jakarta: Direktorat Pembinaan Sekolah Menengah Pertama.
Suyanto. (2009). Urgensi pendidikan karakter Retrieved 8 November, 2012, from http://www.mandikdasmen.depdiknas.go.id/web/pages/urgensi.html

Sweeting, A., & Vickers, E. (2007). Language and the history of colonial education: The case of Hong Kong. *Modern Asian Studies, 41*(1), 1-40.

Swick, K. J., & Williams, R. D. (2006). An analysis of Bronfenbrenner's bioecological perspective for early childhood educators: Implications for working with families experiencing stress. *Early Childhood Education Journal, 33*(5), 371-378. doi: 10.1007/s10643-006-0078-y

Tabachnick, B., & Fidell, L. (2007). *Using multivariates statistics*. Boston: Pearson/Alllen Bacon.

Tashakkori, A., & Tedddlie, C. (2003). The past and future of mixed methods research: From data triangulation to mixed model design. In A. Tashakkori & C. Tedddlie (Eds.), *The handbook of mixed methods in social & behavioral research* (pp. 671-701). Thousand Oaks: Sage Publications.

Thomas, G. (2011). *How to do your case study: A guide for students and researchers*. London: Sage Publications Ltd.

Thompson, A. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127-146). New York: McMillan

Thrupp, M. (2008). National standards for New Zealand's primary and intermediate school pupils. *New Zealand Annual Review of Education, 17*, 199-218.

Tinajero, J. V., & Hurley, S. R. (2001). Assessing progress in second-language acquisition. In S. R. Hurley & J. V. Tinajero (Eds.), *Literacy Assessment of Second Language learners*. Boston: Allen and Bacon.

van den Berg, R. (2002). Teachers' meaning regarding educational practice. *Review of Educational Research, 72*(4), 577-625.

Vardar, E. (2010). *Sixth, seventh, and eighth grade teachers' conception of assessment.* Master of science Master, Middle East Technical University, Ankara.

Webb, N. L. (1992). Assessment of students' knowledge of mathematics: Steps toward a theory. In D. A. Grouws. (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 661-683). New York: MCMillan

White, R. (2007). Balance in assessment. *Measurement: Interdisciplinary Research and Perspectives, 5*(1), 65-67.

Wiliam, D., & Black, P. (1996). Meaning and consequences: A basis for distingushing formative and summative functions of assessment. *British Educational Research Journal, 22*, 537-548.

Winterbottom, M., Brindley, S., Taber, K. S., Fisher, L. G., Finney, J., & Riga, F. (2008). Conceptions of assessment: Trainee teachers' practice and values. *The Curriculum Journal, 19*(3), 193-213.

Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1991). To use their minds well : Investigation new forms of students' assessment. *Review of Research in Education, 17*, 31-74.

Yin, R. K. (2003a). *Applications for case study research* (2nd ed.). London: Sage publications.

Yin, R. K. (2003b). *Case study research : Design and methods* (3rd ed.). London: Sage Publications.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed. Vol. 5). Thousand Oaks: Sage Publications.

Yueming, J., Eslami, Z. R., & Burlbaw, L. M. (2006). ESL teachers' perceptions and factors influencing their use of classroom-based reading assessment. *Bilingual Research Journal, 30*(2), 407-430.

Zulfikar, T. (2009). The making of Indonesian education: An overview on empowering Indonesian teachers. *Journal of Indonesian Social Sciences and Humanities, 2*, 13-39.

Appendices

Appendix A: Key competencies for English subjects

Year 7 Semester one

Key competencies	Basic competencies
Listening	
1. To understand the meaning of simple interpersonal and transactional conversation to interact in a daily life context.	1.1 To respond to the meaning of simple interpersonal and transactional conversation using a variety of simple oral speech accurately, fluently and suitable to interact in a daily life context. This includes greetings, introducing someone, giving command and forbid someone. 1.2 To respond to the meaning of simple interpersonal and transactional conversation using variety of simple oral speech accurately, fluently, and suitable to interact in a daily life context. This includes asking and giving information, thanking, forgiving, and expressing idea in a polite manner.
2. Understand the meaning of simple oral function text to interact in a daily life context.	2.1 To respond to the meaning of speech act of simple oral functional text accurately, fluently and in appropriate manner to interact in a daily life context. 2.2 To respond to the meaning of simple oral functional text accurately, fluently and in appropriate manner to interact in a daily life context.
Speaking	
3. Expressing the meaning of short interpersonal and transactional conversation to interact in a daily life context	3.1 To interact with the nearest environment covering the speech acts like greetings, introducing someone, giving command and forbid someone. 3.2 To interact with the nearest environment covering the speech acts like asking and giving information, thanking, forgiving, and expressing idea in a polite manner.
4. Expressing the meaning of short interpersonal and transactional conversation to interact in a daily life context	4.1 To express the meaning of speech act of simple oral functional text accurately, fluently and in appropriate manner to interact with the nearest environment. 4.2 To express the meaning of idea of simple oral functional text accurately, fluently and in appropriate manner to interact with the nearest environment.
Reading	
5.To understanding the meaning of simple written functional text related to the daily life context	5.1 To read meaningful words, phrases, and sentences loudly with good pronunciation, stressing and intonation. 5.2 To respond to the meaning of simple written functional text accurately, fluently, in an appropriate manner.
Writing 6. To express the meaning of simple written functional text related to daily life	6.1 To express the meaning of simple written functional text using a variety of written text accurately, fluently and in appropriate manner.

context.	6.2 To express the rhetoric structure of simply written functional text using a variety of written text accurately, fluently and in appropriate manner.

Year 7 Semester two

Key competencies	**Basic competencies**
Listening	
7. To understand meaning of simple interpersonal and transactional conversation to interact in a daily life context.	7.1 To respond to meanings of simple short transactional (*to get things done*) and interpersonal (to socialise) conversations accurately, fluently and comprehensively to interact in the daily life context involving the speech act of asking and giving services, asking and giving something, asking and giving information. 7.2 To respond to meanings in simple short transactional *(to get things done*) and interpersonal (to socialise) conversations accurately, fluently and comprehensively to interact in daily life context involving the speech act of asking and giving opinion, expressing like and dislike, responding to something
8. To understand the meaning of simple short functional spoken text and monologue in the forms of descriptive *and* procedure *to* interact in a daily life context.	8.1 To respond to meaning in simple short functional spoken texts accurately, fluently and comprehensively to interact in a daily life context 8.2 To respond the meaning in simple short monologue accurately, fluently and comprehensively to interact in the daily life context in the forms of descriptive and procedure.
Speaking	
9. To express the meaning of simple short transactional and interpersonal conversation to interact in a daily life context.	9.1 To express meanings in simple short transactional (*to get things done*) and interpersonal (to socialise) conversations accurately, fluently and comprehensively to interact in the daily life context involving the speech act of asking and giving services, asking and giving something, asking and giving information. 9.2 To express meanings in simple short transactional *(to get things done*) and interpersonal (to socialise) conversations accurately, fluently and comprehensively to interact in the daily life context involving the speech act of asking and giving opinion, expressing like and dislike, responding to something
10. To express meaning of simple short functional spoken text and monologue in the forms of descriptive and procedure to interact in daily life context	10.1 To express meaning of simple short functional spoken texts accurately, fluently and comprehensively to interact with surrounding environment 10.2 To express meaning of simple short monologue using spoken language accurately, fluently and comprehensively to interact in daily life context in the form descriptive and procedure.

	10.3 To read aloud simple short functional written text and essay in the form of descriptive and procedure with acceptable pronunciation, stress and intonation
Reading	
11. To express meaning of simple short functional written texts and essays in the forms of descriptive and procedure to interact in daily life context	11.1 To respond to the meaning of simple short functional written texts accurately, fluently and comprehensively to interact in daily life context. 11.2 To respond to meaning and rhetoric steps of simple short functional texts accurately, fluently and comprehensively to interact in the daily life context in the form of descriptive and procedure. 11.3 To read aloud simple short functional written text and essay in the form of descriptive and procedure with acceptable pronunciation, stress and intonation
Writing	
To express the meaning of simple short functional written text and essay in the forms of descriptive and procedure to interact in daily life context.	12.1 To express short functional written texts using written language accurately, fluently and comprehensively to interact in daily life context 12.2 To express meaning and rhetoric steps of simple short functional texts accurately, fluently and comprehensively to interact in the daily life context in the form of descriptive and procedure

Appendix B: TCoA Survey

Part 1: Demographic information

Would you provide the following personal information?

1) What is your gender? (*Tick one only)*

- ❒ Female
- ❒ Male

2) What is your age? (*Tick one only)*

- ❒ Between 23-30
- ❒ Between 31-40
- ❒ Between 41-50
- ❒ More than 50 years old

3) What is your education level? (*Tick one only)*

- ❒ Diploma
- ❒ Bachelor
- ❒ Master
- ❒ Doctor

4) For how many years have you taught? (*Tick one only)*

- ❒ Less than 3
- ❒ Between 3-10
- ❒ Between 11 and 20
- ❒ More than 20

5) For which year/level of the school are you teaching? (*Tick one only)*

- ❒ Year 7
- ❒ Year 8
- ❒ Year 9

6) What is your certification status? *(Tick one only)*

- ❒ Certified
- ❒ Not certified

7) What training in educational assessment have you had? (*Tick all that apply)*

- ❒ None
- ❒ ½ to 1 day Workshop or Seminar conducted at school
- ❒ ½ to 1 day Workshop or Seminar conducted by other institution
- ❒ Completed undergraduate Paper
- ❒ Completed postgraduate Paper
- ❒ Other: (*give details*) _____

Thank you for your help. Your cooperation is appreciated.

Part 2: TCoA Questionnaire

This survey asks about your beliefs and understandings about ASSESSMENT, whatever that term means to you. Please answer the questions using your own understanding of assessment.

Use the following rating scale and choose the one response that comes closest to describing your opinion.

- **Strongly disagree (1)**
- **Mostly disagree (2)**
- **Slightly agree (3)**
- **Moderately agree (4)**
- **Mostly agree (5)**
- **Strongly agree (6)**

Once you have completed the survey return it to **Astuti Azis**, Victoria University of Wellington, for analysis. If you wish to know your scores please put your name on the last page of this questionnaire. If you have any queries please do not hesitate to contact Tuty on 463 5233 ext. 9401.

No	Conceptions of assessment	Please circle one for each statement					
		Strongly disagree (1)	Mostly disagree (2)	Slightly agree (3)	Moderately agree (4)	Mostly agree (5)	Strongly agree (6)
1.	Assessment helps students improve their learning	☐	☐	☐	☐	☐	☐
2.	Assessment determines if students meet qualification standards	☐	☐	☐	☐	☐	☐
3.	Assessment information helps modifies on-going teaching of students	☐	☐	☐	☐	☐	☐
4.	Assessment result are sufficiently accurate	☐	☐	☐	☐	☐	☐
5.	Assessment prepares students for examinations	☐	☐	☐	☐	☐	☐
6.	Assessment is used by school leaders to police what teachers do	☐	☐	☐	☐	☐	☐
7.	Assessment has little impact on teaching	☐	☐	☐	☐	☐	☐
8.	Assessment results can be depended on	☐	☐	☐	☐	☐	☐
9.	Assessment helps student succeed in authentic/real world experience	☐	☐	☐	☐	☐	☐
10.	Assessment is used to provoke students to be interested in learning	☐	☐	☐	☐	☐	☐
11.	Assessment cultivates in students a positive attitudes towards life	☐	☐	☐	☐	☐	☐
12.	Assessment results are filed and ignored	☐	☐	☐	☐	☐	☐
13.	Assessment stimulates students to think	☐	☐	☐	☐	☐	☐
14.	Assessment is assigning a grade or level to students work	☐	☐	☐	☐	☐	☐
15.	Assessment is an imprecise process	☐	☐	☐	☐	☐	☐
16.	Assessment measures the worth or quality of schools	☐	☐	☐	☐	☐	☐

17.	Assessment fosters students' character	☐	☐	☐	☐	☐	☐
18.	Assessments interferes with teaching	☐	☐	☐	☐	☐	☐
19.	Assessment teaches examinations-taking techniques	☐	☐	☐	☐	☐	☐
20.	Assessment indicates how good a teacher is	☐	☐	☐	☐	☐	☐
21.	Teachers should take into account error and imprecision in all assessment	☐	☐	☐	☐	☐	☐
22.	Assessment sets the schedule or timetable for classes	☐	☐	☐	☐	☐	☐
23.	Assessment helps students gain good scores in examinations	☐	☐	☐	☐	☐	☐
24.	Assessment selects students for future education or employment opportunities	☐	☐	☐	☐	☐	☐
25.	Assessment results contribute to teachers' appraisals	☐	☐	☐	☐	☐	☐
26.	Assessment helps students avoid failures on examinations	☐	☐	☐	☐	☐	☐
27.	Assessment forces teachers to teach in a way that is contrary to their beliefs	☐	☐	☐	☐	☐	☐
28.	Assessment results should be treated cautiously because of measurement error	☐	☐	☐	☐	☐	☐
29.	Assessment results are trustworthy	☐	☐	☐	☐	☐	☐
30.	Assessment is an accurate indicator of a school's quality	☐	☐	☐	☐	☐	☐
31.	Assessment familiarises student with examination format	☐	☐	☐	☐	☐	☐

Thank you very much for spending time completing this questionnaire; it is much appreciated.

Please read the attachment

Attachment: Invitation to be interviewed

Complete this section only if you are willing to be interviewed

I am looking for volunteers to take part in an interview. This will take about 45 minutes of your time and would be arranged at a time and place to suit you.
If you would be willing to be interviewed, to talk further about your response to this questionnaire please give me the name you are known by and details of the preferred way you wish to be contacted (email address, mobile or home phone).

Please notice that you are under no obligation to go through with the interview when contacted. You can change mind at any time. Your responses will be confidential.

Name : _____

Contact details: _____

Appendix C: Information Sheet

VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui

Information Sheet for Participants

Title of Project: Investigating Indonesian junior high school teachers' conceptions of assessment: A mixed methods study

Researcher: Astuti Azis, School of Educational Psychology and Pedagogy, Victoria University of Wellington (VUW), New Zealand

I am a doctoral student at Victoria University of Wellington (VUW). As part of my PhD, I am conducting research on teachers' conceptions of assessment. The aim of this research is to understand conceptions of assessment held by Indonesian junior high school teachers of English, to explore factors contributing to teachers' conceptions and to investigate how teachers' perceive their conceptions of assessment reflect their assessment practices. This research is supervised by Dr. Margaret Gleeson and Prof. Luanna H. Meyer.

The results from this study will provide rich data towards international understanding teachers' conceptions of assessment in different contexts and different cultures. This study will add to the literature on Indonesian junior high school teachers' conceptions of assessment in the learning of English as a Foreign Language.
I would like to invite you to participate in this research study. Your participation is voluntary and you will be identified under a pseudonym. Your name will not be revealed and it will not be possible for you to be identified personally. This research has been assessed and approved by the Faculty of Education Ethics Committee.

Participants' involvement
As a junior high school teacher of English, you are invited to take part in my study and I would appreciate any assistance you can offer me. Your assistance would involve the following: In the first phase of my study, to capture your conceptions of assessment, I would like you to fill in a questionnaire. The questionnaire completion should take no longer than 15 minutes. If you would like to participate further, please add your contact details to the survey so that I can contact you and invite you to participate in the second phase of the study.

The second phase involves asking you to grant me permission to view your assessment folder and then conducting an interview with you. Firstly, I will ask you to allow me to view and analyse your assessment documents. I would like to draw on materials from your assessment folder as prompts for an interview with you. Your materials will enable me to delve more deeply into your conceptions of assessment, factors contributing to your conceptions, as well as how you perceive your conceptions reflect your assessment practices.

Lastly, I would like to interview you for approximately 45 minutes. The interview will be held in a setting of your choice in or near your school at a time convenient to you. I will audio-record the interview and then transcribe the interview verbatim. Any information taken during interview and when using your documents will be checked with you for accuracy at the end of the session.

Sharing of information
The data collated and examined will be part of my PhD thesis that will be available through Victoria University of Wellington library, Education Department of South Sulawesi Province and Education Department of Gowa Regency. None of the information you share with me at any stage, including your identity, will be shared with your fellow teachers or the PD leader. The only persons who will have access to the data will be my supervisors and me. All collected data (recordings, transcripts and notes) will be kept on password protected system and destroyed three years after the end of the research. A summary of the research findings will be sent to you at your request. The results of this research may be published in academic journals or presented at academic conferences.

Confidentiality
Your participation in this research is voluntary. You have the right to withdraw from the research at any time before data analysis commences without providing any explanation. The information gathered from you will be destroyed after your withdrawal.

If you have any questions about this project please contact me at astuti.azis@vuw.ac.nz. If you have any ethical concerns please contact Dr. Allison Kirkman (Allison.Kirkman@vuw.ac.nz), Chair of Victoria University of Wellington, human ethics committee.

Your contribution to this research as a participant will provide valuable information on the importance of teachers' conceptions of assessment.
Thank you for your cooperation.

Supervisors: Prof. Luanna H. Meyer Victoria University of Wellington, School of Educational Psychology and Pedagogy Tel: 04 463 9598 Email: luanna.meyer@vuw.ac.nz	 Margaret Gleeson, PhD Victoria University of Wellington, School of Educational Psychology and Pedagogy Tel: 04 463 9563 Email: margaret.gleeson@vuw.ac.nz
Researcher: Astuti Azis Victoria University of Wellington, School of Educational Psychology and Pedagogy Tel: 0220243502 Email: astuti.azis@vuw.ac.nz	

VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui

Information Sheet for the Head of Education Department of Gowa Regency

Title of Project: Investigating Indonesian junior high school teachers' conceptions of assessment: A mixed methods study

Researcher: Astuti Azis, School of Educational Psychology and Pedagogy, Victoria University of Wellington (VUW), New Zealand

I am a doctoral student at Victoria University of Wellington (VUW). As part of my PhD, I am conducting research on teachers' conceptions of assessment. The aim of this research is to understand conceptions of assessment held by Indonesian junior high school teachers of English, to explore factors contributing to teachers' conceptions and to investigate how teacher' perceive their conceptions of assessment reflect their assessment practices. This research is supervised by Dr. Margaret Gleeson and Prof. Luanna H. Meyer.

Please accept this letter as my written request for your permission to invite junior high school English teachers in your department to participate in my study. The results from this study will provide rich data towards understanding teachers' conceptions of assessment in different contexts and different cultures. This study will add to the literature on Indonesian junior high school teachers' conceptions of assessment in the learning of English as a Foreign Language.

Teachers in the study will be asked to participate through completing a questionnaire, allowing me to view their assessments and taking part in interviews. I need your consent to contact the English teachers of junior high school in your department in order to select participants. I also need your consent to allow me to conduct a survey, analyse teachers' assessment folders and conduct interviews.

Participants' involvement
Teachers' contribution to this research as a participant will provide valuable information on how Indonesian junior high school teachers of English understand conceptions of assessment and how they perceive their conceptions reflect their assessment practices. All participants who volunteer for this study will be required to give written informed consent.

All participants involved in this research have the right to decline participation and withdraw themselves at any time before data analysis begins without providing any information. The participants can ask questions about the study at any time and have their questions answered to their satisfaction. They will receive a summary of the research findings when the research is concluded, if they wish to.
Data gathered in this study will be kept confidential. None of the information teachers share with me at any stage, including their identity, will be shared with others. The only persons who will have access to the data will be my supervisors and me. All collected data (recordings, transcripts and notes) will be kept on a password protected system and destroyed three years after the end of the research.

Sharing of information
The data collated and examined will be part of my PhD thesis that will be available through Victoria University of Wellington library, the Department of Education of South Sulawesi Province, and the Department of Education of Gowa Regency. The results of this research may be published in academic journals or presented at academic conferences.

If you have any questions about this project please contact me at astuti.azis@vuw.ac.nz. If you have any ethical concerns please contact Dr. Allison Kirkman (Allison.Kirkman@vuw.ac.nz), Chair of Victoria University of Wellington, human ethics committee.

Thank you for your cooperation.

Supervisors:	
Prof. Luanna H. Meyer Victoria University of Wellington, School of Educational Psychology and Pedagogy Tel: 04 463 9598 Email: luanna.meyer@vuw.ac.nz	Margaret Gleeson, PhD Victoria University of Wellington, School of Educational Psychology and Pedagogy Tel: 04 463 9563 Email: margaret.gleeson@vuw.ac.nz
Researcher: Astuti Azis Victoria University of Wellington, School of Educational Psychology and Pedagogy Tel: 0220243502 Email: astuti.azis@vuw.ac.nz	

Appendix D: Consent Form

VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui

CONSENT TO PARTICIPATE IN RESEARCH (Phase 1)

Title of Project: Investigating Indonesian junior high school teachers' conceptions of assessment: A mixed methods study

Researcher: Astuti Azis, School of Educational Psychology and Pedagogy, Victoria University of Wellington (VUW), New Zealand.

Please read each of the following statements carefully before acknowledging your participation in the research.

- ☐ I acknowledge that I have been provided with enough information about the nature and objectives of this research and I have been given the opportunity to seek further clarification.
- ☐ I understand that my identity will remain confidential.
- ☐ I understand that any information I provide will be kept at a secure location, and will only be available to the researcher and her supervisors
- ☐ I understand that I can withdraw from the project until one week after the survey is completed
- ☐ I understand that all research notes and data will be destroyed three years after the conclusion of the research
- ☐ I understand that the information I have provided will be used only by Astuti Azis for this research project, publications and presentation arising from this research
- ☐ I understand that the final thesis will be kept at the Victoria University library, Education Department of South Sulawesi Province and Education Department of Gowa Regency and may be used in publications and conferences
- ☐ I understand that I may contact the researcher or her supervisors if I require further information or to make a complaint relating to my involvement in the research.
- ☐ I agree to complete a research questionnaire for this project.

Signed: _____
I would like to receive a summary of the results of this research when it is completed **YES/NO**

My mailing or email address for the summary:
Mailing/email address_____

VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui

CONSENT TO PARTICIPATE IN RESEARCH (Phase 2)

Title of Project: Investigating Indonesian junior high school teachers' conceptions of assessment: A mixed methods study

Researcher: Astuti Azis, School of Educational Psychology and Pedagogy, Victoria University of Wellington (VUW), New Zealand.

Please read each of the following statements carefully before acknowledging your participation in the research.

- ☐ I have been provided with enough information about the nature and objectives of this research and I have been given the opportunity to seek further clarification
- ☐ I understand that my identity will remain confidential.
- ☐ I consent to the use of my assessment folder for document analysis
- ☐ I consent to a digital audio- recording of my interview
- ☐ I understand that the researcher will give me access to the transcript of my interview so that I can check it for accuracy
- ☐ I understand that I may withdraw from the research without providing any reason up until the data analysis begins (which will be one week after the interview takes place).
- ☐ I understand that any information I provide will be kept at a secure location
- ☐ I understand that any information shared with the researcher will only available to the researcher and her supervisors
- ☐ I understand that all research notes and data will be destroyed three years after the conclusion of the research
- ☐ I understand that if I withdraw from the research all the information I have provided will be destroyed
- ☐ I understand that the information I have provided will be used only by Astuti Azis for this research project, publications and presentation arising from this research
- ☐ I understand that I may contact the researcher or her supervisors if I require further information or to make a complaint relating to my involvement in the research.
- ☐ I understand that the final thesis will be kept at the Victoria University library, Education Department of South Sulawesi Province and Education Department of Gowa Regency and may be used in publications and conferences.
- ☐ I agree to be interviewed by the researcher

Signed: _____

I would like to receive a summary of the results of this research when it is completed **YES/NO**

Mailing/email address_____

Appendix E: Interview Guide

Values and Factors

1. In your current teaching, what is the purpose of assessment?
2. When we are talking about educational assessment in general, what do you think are the main functions of doing assessment?
3. What is the role of assessment in your student learning?
4. What in your experience lead you to this understanding of assessment?

Practice

5. How do you assess your students/What methods do you use in assessing your students?
6. Do different methods of assessment you use have different purposes? tell me about these
7. What criteria do you use to determine your students' grades
8. What strategies do you use in assessing your students? why?/tell me more

Impact

9. What is the impact of assessment on your student learning /on your teaching/on your school accountability?
10. What do you think students learn/should learn from assessment?
11. What do you learn from it?
12. What do you perceive as challenges or problems in assessing your students?

Accuracy

13. To what extent do you perceive that assessment results provide an accurate measure of students' performance?
14. Can you suggest a more accurate/reliable assessment format?
15. How do you use assessment results?

Appendix F: Assessment Documents

National Examination

MATA PELAJARAN

Mata Pelajaran : Bahasa Inggris
Jenjang : SMP/MTs

WAKTU PELAKSANAAN

Hari/Tanggal : Selasa, 24 April 2012
Jam : 08.00 – 10.00

PETUNJUK UMUM

1. Isilah Lembar Jawaban Ujian Nasional (LJUN) Anda sebagai berikut:
   a. Nama Peserta pada kotak yang disediakan, lalu hitamkan bulatan di bawahnya sesuai dengan huruf di atasnya.
   b. Nomor Peserta, Tanggal Lahir, dan Paket Soal (lihat kanan atas sampul naskah) pada kolom yang disediakan, lalu hitamkan bulatan di bawahnya sesuai dengan angka/huruf di atasnya.
   c. Hitamkan bulatan pada kolom Nama Mata Ujian yang sedang diujikan.
   d. Nama Sekolah, Tanggal Ujian, dan Bubuhkan Tanda Tangan Anda pada kotak yang disediakan.
2. Tersedia waktu 120 menit untuk mengerjakan Paket Soal tersebut.
3. Jumlah soal sebanyak 50 butir, pada setiap butir soal terdapat 4 (empat) pilihan jawaban.
4. Periksa dan laporkan kepada pengawas ujian apabila terdapat lembar soal yang kurang jelas, rusak, atau tidak lengkap.
5. Tidak diizinkan menggunakan kalkulator, HP, tabel matematika atau alat bantu hitung lainnya.
6. Periksalah pekerjaan Anda sebelum diserahkan kepada pengawas ujian.
7. Lembar soal boleh dicoret-coret.

SELAMAT MENGERJAKAN

MATA PELAJARAN

Mata Pelajaran : Bahasa Inggris
Jenjang : SMP/MTs

WAKTU PELAKSANAAN

Hari/Tanggal : Selasa, 24 April 2012
Jam : 08.00 – 10.00

PETUNJUK UMUM

1. Isilah Lembar Jawaban Ujian Nasional (LJUN) Anda sebagai berikut:
   a. Nama Peserta pada kotak yang disediakan, lalu hitamkan bulatan di bawahnya sesuai dengan huruf di atasnya.
   b. Nomor Peserta, Tanggal Lahir, dan Paket Soal (lihat kanan atas sampul naskah) pada kolom yang disediakan, lalu hitamkan bulatan di bawahnya sesuai dengan angka/huruf di atasnya.
   c. Hitamkan bulatan pada kolom Nama Mata Ujian yang sedang diujikan.
   d. Nama Sekolah, Tanggal Ujian, dan Bubuhkan Tanda Tangan Anda pada kotak yang disediakan.
2. Tersedia waktu 120 menit untuk mengerjakan Paket Soal tersebut.
3. Jumlah soal sebanyak 50 butir, pada setiap butir soal terdapat 4 (empat) pilihan jawaban.
4. Periksa dan laporkan kepada pengawas ujian apabila terdapat lembar soal yang kurang jelas, rusak, atau tidak lengkap.
5. Tidak diizinkan menggunakan kalkulator, HP, tabel matematika atau alat bantu hitung lainnya.
6. Periksalah pekerjaan Anda sebelum diserahkan kepada pengawas ujian.
7. Lembar soal boleh dicoret-coret.

SELAMAT MENGERJAKAN

The following text is for questions number 1 and 2.

Hi, Ayu

Feeling better now? By the way, we have a project in Biology class, cassava fermentation. It sounds interesting. We are in one team. I hope you get well soon and come to school next week. We have to do some preparation for it.

Miss you,
Sandra

1. What does Sandra inform Ayu? She ...
   A. says that Ayu is not in her team.
   B. tells Ayu about the biology project.
   C. says that Ayu will miss the biology class.
   D. asks Ayu to do the preparation by herself.

2. From the text we know that ...
   A. Ayu is not sick.
   B. Sandra is Ayu's classmate.
   C. they will have a biology examination.
   D. Sandra is getting better next week.

The following text is for questions number 3 and 4.

To: Anderson

Congratulation, you have bought a bigger house you've always dreamt about. I hope you will feel more comfortable to stay in it. Don't forget to invite me on the house-warming party

Panji

3. Why Anderson bought a bigger house?
   A. He wanted to invite all his friends.
   B. He wants to feel more relax.
   C. He always wanted to have a bigger house.
   D. He dreamt to invite his friends to a party.

4. "I <u>hope</u> you will feel more comfortable to stay in it."
   The underlined word can be replaced with ....
   A. think
   B. say
   C. wish
   D. spend

The following text is for questions number 5 and 6.

5. Who will host The New Year's Eve Party?
   A. Jimmy.
   B. Fletcher's family.
   C. Sandy and Bill Fletcher.
   D. Sandy, Bill, and their friends.

6. From the text above, we can conclude that ....
   A. Jimmy will not celebrate the new year party
   B. Sandy and Bill will invite you to their house
   C. the party will be celebrated on the street
   D. you should call Jimmy if you want to come

Read the following text to answer question number 7.

DUE TO MANY SHARP CORALS,
THE VISITORS ARE PROHIBITED TO SWIM ALONG THIS AREA

7. Where do we usually find the text?
   A. At the lake.
   B. At the beach.
   C. At the river bank.
   D. At the swimming pool.

The following text is for questions number 8 to 10.

Zoofari Beach Party

The Washington Park Zoo will hold a new annual fundraiser "Zoofari Beach Party"
Saturday June 25th, 5:30 pm
auction, BBQ, Buffet Dinner, feeding animals and free flow wine and softdrink
Limited advanced tickets are available at the zoo ticket booth
Adult - $20
Children (6-12) - $10
fun for the whole family
get your ticket now!!!

Committee

8. The committee hold a beach party because ....
   A. they will celebrate zoofari day
   B. they will welcome the visitors
   C. they want to collect fund
   D. they want to have fun

9. "... are <u>available</u> at the zoo ticket booth."
   The underlined word means ....
   A. kept
   B. bought
   C. sold
   D. seen

10. The announcement tells us about ....
   A. the Zoofari fundraiser program
   B. the facilities on the beach
   C. the beach party visitors
   D. the price of the tickets

The following text is for questions number 11 and 12.

A BURST OF NATURE

NatraBurstTM is a powerful food source consisting of a wide variety of premium ingredients, which work together to provide a variety of health benefits. Each serving contains the equivalent of more than 6 vegetables and fruits, with an ORAC value of 4.000+. Additionally, NatraBurstTM help support lean muscle mass with its concentration of high quality protein.

- Increase energy
- Reduce appetite
- Promote Fat Loss
- Anti-Aging
- Promote Healthy Skin

BUY 2 GET 1 FOR FREE ON SELECTED STORE

11. What's the benefit of consuming NatraBurst?
    A. Increase appetite
    B. Decrease energy
    C. Promote fat loss
    D. Promote healthy hair

12. The purpose of the writer writes the text is ....
    A. to warn the reader the danger of the product
    B. to describe the composition of the product
    C. to persuade people to consume the product
    D. to tell the reader how to use the product

The following text is for questions number 13 to 15.

Palembang, January 21, 2012

Dear Paula,

Hello Paula, how are you? It's been a long time since we last met. Well, I just want to tell you that I have a new pocket camera. My uncle, Rudi, arrived from Singapore last month and bought it for me. It is a digital camera. I am very happy to have it.

Because of that, I have a new hobby, photography. I bring my camera everywhere and I take pictures of people or something interesting. Someday, I want to join national or even international photo contest. I have already had a number of photo collections by now.

Here, I give you some of my works. Give me your comment about my work. I would like to hear an honest opinion from you.

Ok, I think that's all from me, write to me soon, OK?

Regards,
Nadia

13. Who bought the camera?
    A. Paula.
    B. Nadia.
    C. the writer's uncle.
    D. Paula's uncle.

14. What is the main idea of the second paragraph?
    A. The writer has a new hobby.
    B. Nadia always takes pictures everywhere.
    C. The writer joins some competitions.
    D. Nadia needs some comment from Paula.

15. The text generally talks about ....
    A. the writer's new hobby
    B. Paula's new hobby
    C. a pocket camera
    D. Nadia's camera

The following text is for questions number 16 to 18

How to Treat a Burn Wound

Treating burn wound can be easy and simple but can also be very dangerous. Here is how you treat them:

1. Start by immediately cooling the burn wound with either an ice pack or cold running water for 5 – 10 minutes.
2. Evaluate the severity of the wound, if only the top layer of skin is affected, the burn treatment can be done at home.
3. After keeping the wound under water for another 20 – 30 minutes, cover the wound loosely with a bandage and take a pain-reliever for however long the wound continues to hurt.
4. The burn should heal itself completely in 10 – 14 days. If it doesn't you will need to see a physician.

16. In what condition you don't need professional care?
    A. When the wound takes longer time to heal
    B. When only top layer of skin is affected
    C. When infection occurs in the wound
    D. When you have already put ice pack on the wound

17. From the text above, we can conclude that ....
    A. the burn wound is cooled with fresh water
    B. the wound is covered tightly with a bandage
    C. the burn wound is simple to treat
    D. the burn victim should be hospitalized

18. "If <u>it</u> doesn't, you will need to see a physician"
    The underlined word refers to ....
    A. the wound
    B. the bandage
    C. the physician
    D. the pain reliever

The following text is for questions number 19 to 21.

I live in chinatown area or *Pecinan* in Semarang. It is a very crowded place. There are many kelenteng (temples) here. One of them is *Tay Kak Sie* Temple in Gang Lombok which people visit regularly.

Not far from the temple, there is a very popular tourist attraction which can only be found at night, *Semawis* night market. The name *Semawis* is derived from a Javanese term which means everything is available here. Yes, we can really find many kinds of things here, ranging from specialty foods, to *Feng Shui* consultation, to karaoke.

However, my favourite place that I always visit when I have a spare time is *Pak* Suharto's workshop. The workshop is next to the Chinatown gate. In his workshop, he teaches his students to paint using Chinese ink. The ink is made from pine wood. People come to *Pak* Suharto's workshop to order paintings, too. I love watching *Pak* Suharto drafting and painting his objects.

19. Where is *Pak* Suharto's workshop located?
    A. Behind the temple
    B. Not far from the temple
    C. At the centre of pecinan
    D. Next to the Chinatown gate

20. The night market is called *Semawis* because ....
    A. people can find various of things there
    B. it is located inside the Chinatown
    C. people visit it reguarly
    D. it is near the temple

21. "... to *Feng Shui* consultation ...." (Paragraph 2)
    The underlined word means ....
    A. a talk with the chief of the group
    B. a discussion over something for at least one day
    C. a meeting with an expert in a particular field to get advice
    D. a secret meeting conducted by a group of people to set a plan

The following text is for questions number 22 to 24.

EASTER ISLAND

Easter Island is in the South Pacific, 3.700 km from the coast of Chile. On this island, there are 600 large statues. We don't know who built them, but they are probably constructed between 1150 and 1500. We don't really know why they are there too.

A Norwegian explorer, Thor Heyerdhal, believed that they were built by people from South America. To prove this, he made a simple raft and sailed there, all the way from Peru.

Archeologists think that the statues represent dead tribal leaders. We don't know why the statues were left alone on the island. Perhaps the builder were killed by a disease or war. Perhaps they used all the natural resources on the island. There are many unanswered questions about Easter Island.

22. "... believe that <u>they</u> were built by ..." (paragraph 2)
The underlined word refers to ....
   A. unanswered questions
   B. natural resources
   C. the huge statues
   D. the ancient builders

23. The main idea of the last paragraph is ....
   A. the archeologist's opinion about the statues
   B. the reasons for making the statues
   C. the tribal leaders' statues
   D. the disease and war

24. The text generally talks about ....
   A. the location of Easter Island
   B. Easter Island with its statues
   C. the journey of a Norwegian explorer
   D. the statues represent tribal leaders

The following text is for questions number 25 to 27.

Last Saturday, I woke up early. I didn't get up from my bed because there was no school. Suddenly, my telephone rang. It was my friend, Fanny. She asked me to go out at 10.00 o'clock. She wanted to buy something in the traditional market. So I hopped from the bed and headed to the shower.

Finally, we were out. In the street, I saw a piece of pink paper. I took it and we read it. We were shocked, it was a voucher for a four night tour to Lombok. The expired date was that day. To our surprised, the name was Fanny Fenita and the birth date was exactly the same as Fanny, my friend, and it was also valid for two persons. My God! We were thinking that maybe the paper just fell from the sky for us.

We were hurried to the tour agency that issued the voucher. The tour agency took care of everything. We went home and still couldn't believe what was going on. Two days later we were on the Senggigi Beach, Lombok.

25. What did the writer see in the street?
    A. A voucher.
    B. A friend.
    C. Telephone.
    D. Market.

26. Why were the writer and her friend shocked?
    A. She took the pink paper.
    B. Her wish came true.
    C. The voucher fell from the sky.
    D. The voucher has her friend name.

27. What is the main idea of the first paragraph?
    A. The writer didn't get up.
    B. The writer's friend asked her out.
    C. The writer's friend telephoned her.
    D. The writer didn't go to school that day.

The following text is for questions number 28 to 30.

I took my sister, Katy for her first flight last month. When we arrived at the airport, she wanted to know why the official weighed the luggage. Then she wondered why she needed a passport. She didn't understand what the security check was for. Then she asked who the flight attendant was. When she saw the plane, she couldn't understand what the captain did and where he sat.
What a lot of questions! I was very pleased when we landed.

28. "..., what the captain did and where he sat."
The word 'he' refers to ....
    A. the official
    B. the seceurity
    C. the officoal
    D. the captain

29. After the airplane landed, ...
    A. the writer is very pleased
    B. the sister weighed the luggage
    C. the sister asked the flight attendant
    D. the writer needed a passport

30. What did the text tell us about?
    A. Katy and her sister's hobby
    B. Katy's first flight
    C. The activity at the airport
    D. The flight attendant's job

The following text is for questions number 31 to 34.

A gaunt wolf was almost dead with hunger when he met a house-dog. "Ah cousin", said the dog.

"I knew how it would be, your irregular life will finish you. Why don't you work steadily as I do, and get your food regularly given to you".

"I would have no objection", said the wolf. "If I could only get a place".

"I will easily arrange that for you" , said the dog.

"Come with me to my master and share my work".

So the wolf and the dog went towards the town together. On their way, the wolf notice that the hair of the dog's neck was very much worn away. So he asked him what happened.

"Oh, it's nothing, this is only the place where the collar is put on at night to keep me chained up, but you will get to used to it".

"Really? So good bye to you master dog". Then the wolf walk away.

31. Why did the wolf change his mind?
    A. He is afraid of the dog's master.
    B. He wants to have his freedom.
    C. He doesn't want to lose his hair.
    D. He is worried to live in a town.

32. The main idea of the last paragraph is ....:
    A. the dog was chained up at night
    B. the dog loss his hair around his neck
    C. the wolf decided not to follow the dog
    D. the wolf was afraid of working hard

33. The text mostly tells about ....
    A. a hungry wolf
    B. the dog and the wolf
    C. a hard worker dog
    D. a wealthy dog

34. What can we learn from the text?
    A. Better live in the jungle than in a city.
    B. It's not good to share work with a family member.
    C. Performance is everything in life.
    D. Better live free than be a slave.

The following text is for questions number 35 to 38.

Qui Jun and the Arrogant Monk

Once upon a time, there was a monk called Shan lives in a village in China. He was very popular. However, this made him very arrogant.

Qui Jun heard about his arrogance and wanted to teach the monk a lesson. He went to meet Shan. Dressed up as a commoner, He was ignored by Shan. Just then a servant of the monk came with a message. "The son of an army officer is here to see you".

The monk said, "I will go and greet him".

Shan welcomed the son of the army officer respectfully.

After the army officer's son had departed, Qui Jun asked Shan the reason for his double-faced behaviour. "Why is it that you greeted the army officer's son so respectfully, yet behaved so arrogantly towards me?"

The monk had a quick reply, "Please don't get me wrong. For me greeting means not greeting, as not greeting means greeting."

Qui Jun understood the monk's mischief and hit him on his head with his stick.

"According to your logic, beating you means not beating you and not beating you means beating you. Therefore, I have to give you a beating", said Qui Jun.

Shan immediately realized the folly of his actions and started showing respect to everyone he met, irrespective of their status.

35. Who came with a message?
    A. The son of an army officer.
    B. Shan's servant.
    C. Shan, the monk.
    D. Qui Jun.

36. Why did Qui Jun say that Shan had double-faced behaviour?
    A. He neither greeted nor acknowledged him.
    B. He welcomed the son of the army officer disrespectfully, but not him.
    C. Qui Jun heard about his arrogance and wanted to teach him a lesson.
    D. He greeted the army officer's son, but not him.

37. "However, this made him very <u>arrogant</u>." (Paragraph 1)
    What does the underlined word mean?
    A. Polite to everybody.
    B. Friendly to other people.
    C. Showing respect to others.
    D. Feeling better than others.

38. From the story we can learn that we should ...
    A. respect everyone, no matter what their status.
    B. be arrogant because we are popular.
    C. respect others based on their status.
    D. have double-faced behaviour.

Read the text and answer questions 39 to 42.

Jellyfish are not really fish. They are invertebrate animals. This means that unlike fish or people, they have no backbones. In fact, they have no bones at all.

Jellyfish have stomachs and mouths, but no heads. They have nervous systems for sensing the world around them, but no brains. They are made almost entirely of water, which is why you can look through them.

Some jellyfish can glow in darkness by making their own light. The light is made by a chemical reaction inside the jellyfish. Scientists believe jellyfish glow for several reasons. For example, they may glow to scare away predators or to attract animals they like to eat.

Most jellyfish live in salt water, apart from a few types that live in fresh water. Jellyfish are found in oceans and seas all over the world. They live in warm, tropical seas and in icy waters near the North and South poles.

39. Which one creates Jellyfish's light?
    A. White blood.
    B. Nervous system.
    C. Chemical reaction.
    D. Salt water.

40. Based on the text, we know that ....
    A. they belong to invertebrate animals.
    B. they have heads like other animals.
    C. their brain helps them find the food.
    D. they cannot live in fresh water.

41. What is the text about?
    A. Jellyfish.
    B. Kinds of all fish.
    C. All invertebrate animals.
    D. Some kinds of sea animals.

42. "Some jellyfish can glow in darkness by making their own light." (paragraph 3)
    The word "glow" in the sentence means ....
    A. move
    B. produce
    C. appear
    D. shine

For questions 43 to 45, complete the following text with the correct answer.

One of the my most memorable trip was a tour ... (43) a Portuguese graveyard. The tour was ... (44) by the caretaker. The wealthy ... (45) had marble mausoleums while the poor people rent a plot for 10 years.

43. ....
    A. to
    B. for
    C. from
    D. into

44. ....
   A. give
   B. gave
   C. given
   D. giving

45. ....
   A. friends
   B. workers
   C. relatives
   D. families

This text is for number 46 to 48.

My Bedroom

My bedroom is (46) ... . It'a about four meters long and three meters wide. There is a bed across the door. A desk and a chair stand near the window. On the corner, stand three door wooden cupboard to (47) ... all of my stuff and my clothes. On the center of the ceiling, a twenty-watt spiral lamp gives enough light for the entire of the room and functions as a reading lamp as well. I really thanked to my father who designed the room because I have enough space and I stay (48) ... in it.

46. ....
   A. fabulous
   B. spacious
   C. hazardous
   D. tremendous

47. ....
   A. keep
   B. keeps
   C. kept
   D. keeping

48. ....
   A. comfortably
   B. awkwardly
   C. clumsily
   D. quietly

49. Arrange the following jumbled words into a good sentence.

Cinderella	was	called	a	there	girl	once	beautiful
1	2	3	4	5	6	7	8

   A. 7 – 5 – 3 – 4 – 6 – 8 – 1 – 2
   B. 7 – 5 – 2 – 4 – 8 – 6 – 3 – 1
   C. 7 – 5 – 2 – 1 – 4 – 8 – 6 – 3
   D. 7 – 5 – 2 – 4 – 6 – 8 – 1 – 3

50. **Arrange the following sentences into a good paragraph.**
    1. We started our day with performances.
    2. The performance that I was in was Labamba
    3. We had performances, food stalls, displays, a raffle ticket draw.
    4. Then we played games like tug war and football
    5. They came from Australia, Asia, Arab, and Greece.
    6. We had an International Day at my school yesterday.
    7. After our school activities, we had lunch.
    8. There were food stalls that serve food from foreign countries.

    A. 6 – 3 – 5 – 1 – 2 – 8 – 7 – 4
    B. 6 – 3 – 1 – 2 – 4 – 7 – 8 – 5
    C. 6 – 3 – 1 – 7 – 8 – 5 – 4 – 2
    D. 6 – 3 – 2 – 7 – 1 – 4 – 8 – 5

Semester test for grade 7

PEMERINTAH KABUPATEN GOWA
DINAS PENDIDIKAN OLAH RAGA DAN PEMUDA
Alamat : Jl. Mesjid Raya No. 30 Sungguminasa Kabupaten Gowa , Kode Pos. 9111 Telp.. (0411) 867774

SOAL ULANGAN SEMESTER GANJIL TP.2011/2012

SATUAN PENDIDIKAN	: SMP / MTs
MATA PELAJARAN	: **Bahasa Inggris**
KELAS	: VII (Tujuh)
BENTUK SOAL	: Essay
HARI/TANGGAL	: ...../..... 2011
JUMLAH SOAL	; 15 Nomor
WAKTU	: 90 Menit

Petunjuk : Jawablah pertanyaan di bawah ini dengan Singkat,jelas dan tepat !

Reading Text

Hi Nia,
My name is Sheila.I am from Yogyakarta.I am thirteen years old.This is my photograph. I am in the middle .We are students.Ida and Andy are my classmates.They are from Bandung.
Tell me about you, Nia.

Write soon,
Love,
Sheila

Answer the questions :

1.Who writes the letter ?
2.Who is the letter for ?
3.What is the writer ?
4.How old is the writer?
5.Where is the writer from ?
6.Read the text carefully. Then fill in the following form.
Arham is a new student at SMP 2 Sungguminasa.He wants to be a member of boyscouts in his school.But before, he must fill in the form that consist of his data.Your job is to help Arham to fill in the form.

Eagle Boyscout
SMP 2 Jl.Jend Soedirman Sungguminasa

Students' Data
1. Full Name :
2. Date of birth :
3. Address :
4. Phone Number :
5. Interests :

6. Parents Name :
7. Blood Type :
8. E-mail Address (if any) :

Sungguminasa,...../...../2011

(your name)

Complete the following 'greeting card' using the words provided

For Mira ,
best friend

Hi Mira ! How are you ?(7).....since the (8).....we played basketball in my school yard.
Let's meet again and (9).....,,,at my house.How about having an "ayam bakar"
(10) .....with my family?
Can't (11).....

Love and Hug,

Nina

-spend some time	-party	-wait to see
-It's been two years	- last time	

12. Rearrange these words into good sentences
    a. had – wonderful – a – I – holiday
    b. today – is – the weather – rice
    c. buying – is – book – the girl – a
13. Complete the sentences .Use the words provided

– takes	-brush
–studies	-watch

    a. I always .... my teeth.
    b. She usually ....her lesson in the evening.
    c. He often .....a nap in the afternoon.
    d. We usually ..... television in the evening.

Complete each message below. The clues will guide you.

14. Write a note to your father. Ask him to pick you up at school at 02.00 p.m.Tell him that you will be waiting at the gate.

Dad,

..... at 02.00 pm. I'll be.....

15. Write a note to your sister,Risa. Tell her that you have to early.Ask her to come to school.

Risa,

I have .....,just .....

Semester test for grade 8

PEMERINTAH KABUPATEN GOWA
DINAS PENDIDIKAN OLAH RAGA DAN PEMUDA
Alamat : Jl. Mesjid Raya No. 30 Sungguminasa Kabupaten Gowa , Kode Pos. 9111 Telp. (0411) 867774

SOAL ULANGAN SEMESTER GANJIL TP.2011/2012

SATUAN PENDIDIKAN	: SMP / MTs
MATA PELAJARAN	: **Bahasa Inggris**
KELAS	: VIII (Delapan)
BENTUK SOAL	: Essay
HARI/TANGGAL	: ...../..... 2011
JUMLAH SOAL	: 15 Nomor
WAKTU	: 90 Menit

Jawablah soal di bawah ini dengan singkat,jelas dan tepat !

I .Read the text below to answer questions no 1 – 3 .

Dear friend,

Please come to my Birthday Party on :
Saturday, August 26, 2011.
At 5 pm
At my house, Jl Mayang 17.
Please come ! We'll have lots of fun.

Mira

1. Who sends the invitation ? siapa mengirim ini
2. What time will the party start ?
3. Where will be the Birthday Party ?

II. Read the text below to answer quetions no 4 – 7.

Singapore is an island city of about 4 million people. It's a beatiful city with lots of parks and open spaces. It's also a clean city.

Most of the people live in high-rise flats in different parts of the island. The business district is very modern, with lots of tall new office buildings. Singapore also has some nice older sections. In Chinatown there are rows of old shop houses. The government buildings in Singapore are very beautiful and date from the colonial days.

Singapore is famous for its shops and restaurants. There are many goods shopping centers. Singapore retaurants sell Chinese, Indian, Malay and European food, and the prices are quite reasonable.

4. How big is the population of Singapore ?
5. Where do most people live ?
6. What can we find in Chinatown ?
7. What is the main idea of the last paragraph ?

III. Read the text below to answer questions no 8 – 11 .

Several days ago, I was ill, so I went to the public Health Center near my house. After waiting a while, a doctor met me. Then, the doctor examined me. The doctor asked me to be injected because my illness was quite serious. I agreed to be injected. After that the doctor gave me advice to Have a rest at home for several days. He also wrote a prescription for me. He ordered me to take the medicine as soon as possible, before I left his room.

8. What happened to the writer several days ago ?
9. Where did the writer go ?
10. Did the writer agree to be injected ?
11. What does the word "**He**" in the text refer to ?

IV. Complete the following text using the words provided in the box.

You're invited

Occassion : ..... (12)
Date : March 13
Time : 6:30 p.m.
Place : .....(13)

..... (14). This is a surprised party.

Regrets only : Marie Salinger 238-1722

- Please arrive on time
- 44 Devoe Road
- Sunday
- Kim's Birthday

15. Arrange the following sentences into good order.

- On Saturday morning, I helped my sister making a birthday cake.
- On Friday, I went to the mall to buy a nice shirt for my grandpa.
- It was my grandfather's birthday last Sunday.
- On Sunday evening, my uncle and aunt came to my house, they brought flowers for grandpa .
- Finally grandpa blew the candle and cut the cake while we were singing "Happy Birthday" song .
- Then, we sat together in the living room.

Semester test for grade 9

PEMERINTAH KABUPATEN GOWA
DINAS PENDIDIKAN OLAH RAGA DAN PEMUDA

Alamat : Jl. Mesjid Raya No. 30 Sungguminasa Kabupaten Gowa , Kode Pos. 9111 Telp.. (0411) 867774

SOAL ULANGAN SEMESTER GANJIL TP.2011/2012

SATUAN PENDIDIKAN	: SMP / MTs
MATA PELAJARAN	: **Bahasa Inggris**
KELAS	: IX (Sembilan)
BENTUK SOAL	: Pilihan Ganda
HARI/TANGGAL	: ...../..... 2011
JUMLAH SOAL	; 30 Nomor
WAKTU	: 90 Menit

Petunjuk : Pilihlah Jawaban Yang tepat di bawah ini !

1. Ami : are sure this Mr.Bram's house ?
   Fany : . . . (expression of certainty)
   a. Definetely ? c.I'm quite certain
   b. That's Ok d. Really ?
2. Complete the dialogue below with uncertainty expression
   Fadel : What are you doing this weekend ?
   Adi : _____
   a. Sure c. I'm certain
   b. Certainly d. I'm not sure

Text 1

To make some pineapple stew you will need 1 pineapple cut into dice, 3 glasses of water, 1 glass of granulated sugar, 2 sticks of cinnamons, and 10 whole clovers.

The first thing that you need to do is boil the water and bring it to boil. Then, you have to the add sugar, cinnamon, and clovers. Once it is boiled, add the pineapples. After that you need to wait until you can smell the pineapple aroma. Finally, remove it from heat and serve cold

For number 3- 9 answer the questions based on the text above !

3. What is the text about ?
   a. Serving pineapple stew c. Describing pineapple stew
   b. Making pineapple stew d. Explaining pineapple stew
4. What do we need to make pineapple stew?
   a. 1 pineapple, 2 sticks of cinnamons,1 glasses of water,3 glass of granulated sugar. and 10 whole clovers.
   b. 1 pineapple . 3 glasses of water, 1 glass of granulated sugar, 2 sticks of cinnamons, and 10 whole clovers.
   c. 1 pineapple . 1 glasses of water, 2 glass of granulated sugar, 3 sticks of cinnamons, and 10 whole clovers.

d. 3 pineapple , 3 glasses of water, 2 glass of granulated sugar, 1 sticks of cinnamons, and 10 whole clovers.

5. How many ingredients do we need to make pineapple stew ?
   a. 5 b. 7 c. 6 d. 8
6. What you do before add sugar, cinnamon and clovers?
   a. Heat the water until boil c. Smell the pineapple aroma
   b. Add the pineapples d. remove from the heat and serve cold
7. How should you serve the pineapple stew?
   a. Boil b. cold c. Hot d. frozen
8. What is the main idea of the second paragraph ?
   a. How to make pineapple stew c. The ingredients we need to make pineapple stew
   b. How to serve pineapple stew d. The procedure how to make pineapple stew
9. Finally, remove *it* from heat and serve cold . What does the word "it" here refer to ?
   a. Water c .Sugar
   b. Pineapple stew d. Pineapple

ANNOUNCEMENT

In order to prepare the committee of story telling and speech contest, all OSIS board members need to have a meeting on November 24, 2011 after school at the library. It is very important not to miss the meeting.

Please come on time!

Elvinda

The OSIS Leader

10. What does the announcement tell you ?
    a. A meeting on November 24,2011
    b. About Speech contest
    c. The committee preparation of story telling and speech contest
    d. It is very important not to miss the meeting
11. The purpose of writing the announcement is to . . . .
    a. invite all OSIS board members to have a meeting
    b. prepare the story telling and speech contest
    c. ask all OSIS board members to come on time on every meeting
    d. describe the committee of competition
12. The OSIS board members will . . . on the meeting.
    a. compete for all OSIS board members c. study together at the Library
    b. form a committee of the contest d. practice story telling
13. Where did the meeting be held ?
    a. At the OSIS room c. At the school
    b. At the Committee room d. At the library
14. It is very important <u>not to miss</u> the meeting. The underlined phrase means to . . . .
    a. stand b. lose c. hold d. Come

Text 3

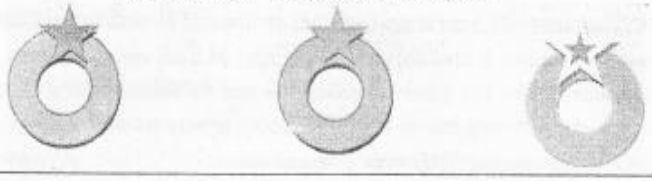
Banks

A bank is a financial institution that that provide banking and other financial services. It is an institution that holds a banking license.

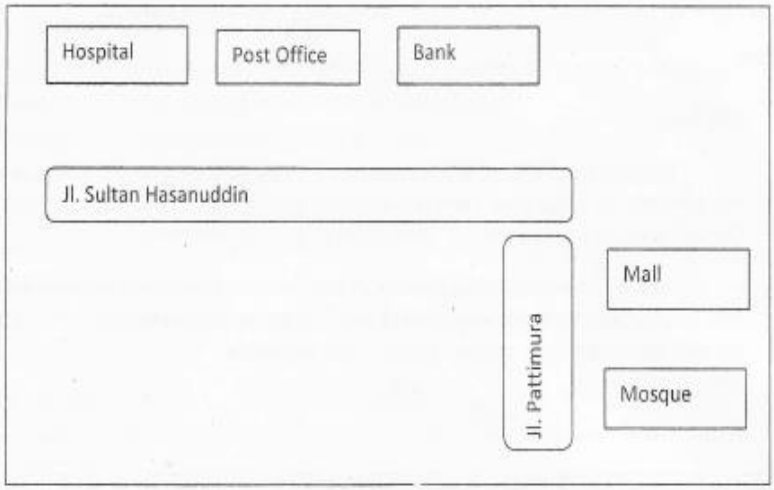
The word bank derived from Italian banca. This word derived from German and means bench. Typically, a bank generates profit from transaction fees on financial services or the interest from a loan. The type of services offered by a bank depends on the type of bank and the country. However, bank services usually include lending out money to companies and individuals, issuing checking and saving accounts, chasing check, facilitating money transactions such as wire transfer, issuing credit card, ATM, debit card and also storing valuable things in save deposit boxes.

15. Where do the bank store the valuable things ?
    a. ATM
    b. Deposit boxes
    c. Credit card
    d. Debit card
16. The first paragraph tells us about ?
    a. The facilities provided by bank .
    b. The kind of bank.
    c. The function of bank
    d. The definition of bank .
17. Bank generates profit from transaction fees on financial services or the interest from a loan.
    The word "generate" here has as similar meaning with ....
    a. Keep
    b. Develop
    c. Protect
    d. Save
18. Mother : Could you turn on the television ?
    Puput : _____
    Mother : Turn on the television ,please .
    a. I don't know
    b. I don't agree with you
    c. Sorry ,what did you say ?
    d. heally ?
19. Rudi : Have you hear about Chandra ?
    Tony : What's wrong with him ?
    Rudi : He had an accident last night
    Rudi : Really ?_____
    a. I'm sorry to hear that
    b. It's just kidding
    c. It's Ok
    d. Don't worry
20. Rizki : I got 10 for Mathematics yesterday.
    Tika : *Really? Fantastic*!
    The italic word is showing .
    a. attention b. admiral c. agreement d. Repetition

Choose a beautiful gift from the stunning designer diamond
Jewellery Collections by Vivienne Westood, Jasper Conran, Bruce Oldfield and Versace
At Cool Diamonds Europe's largest on-line jewelers. With over 5 million hits each month
this website has revolutionized the way we buy diamonds.
The bespoke service also offer a 30–day money-back guarantee.
View in 3D online, or by appointment at 16 Greville Street, EGIN8SQ
020 7045 5322 www,cooldiamonds,com



21. These right statement from the text is ...
    a. Vivienne Westwood is a diamond jewellery seller
    b. The cool Diamonds website has revolutionized the way to buy diamonds
    c. Jasper Conran is not the stunning diamond jewellery designer
    d. We can't buy the diamonds jewellery on-line
22. " Choose a beautiful gift from the stunning designer diamond jewellery "
    What is the meaning of the underlined word ?
    a. Very surprising
    b. Very annoying
    c. Very challenging
    d. Very impressing
23. What is the advertisement about ?
    A. To describe the collection diamond
    B. To tell the people how to get the diamond
    C. To show the diamond to the people
    D. To get the diamond from the designer
24. Lela told poppy not to disturb her politely.What does she say ?
    Lela : _____
    Poppy : Sorry.
    a. Don't bother me, please !
    b. Disturb me, please !
    c. Don't worry, be happy !
    d. Don't leave me alone !
25. Dedy : *Excuse me, Can I borrow the newspaper ?*
    Rudy : Here, you are.
    The italic utterance expresses _____
    a. Gratitude b. politeness c. asking opinion d. introduction
26. The following is a sentence that expresses politeness
    a. Take the book !
    b. Excuse me, let me pass this way !
    c. Come on, don't cry baby !
    d. Shut up !



27. What building do you find on the corner of Jl. Sultan Hasanuddin before turn to jalan Patimura

a. Hospital b. Post office c. Bank d.Mall

28. What building do you find on the corner of Jl. Sultan Hasanuddin before turn to jalan Patimura

a. Hospital b. Post office c. Bank d.Mall

29. What building do you find after you get the hospital on Jln. Sultan Hasanuddin ?

a. Post office b. Bank c. Hospital d. Mosque

30. Give a direction how to get Mosque from The hospital

a. walk along Jl. Pattimura and turn left to Jln. Hasanuddin and the hospital is before Post office

b. Walk along Jl. Hasanuddin and turn rigt to Jln.Pattimura and the mosque is after the Mall

c. Walk along Jl. Hasanuddin and turn rigt to Jln.Pattimura and the mosque is before the Mall

d. walk along Jl. Pattimura and turn left to Jln. Hasanuddin and the hospital is beside Post office

How to Make a cup of Tea

a) Add boiling water

b) Add sugar ,milk or lemon to taste

c) Leave to infuse 2 to 3 minutes

d) Put one tea bag in a cup

31. The best arrangement for the direction above is _____

a) a – b – c – d b). d- c – b – a c.) d – a – c – b d). d – c – a – b

Jalan Gunung Bawakaraeng Makassar

4th December 2010

Dear Bunga

Bunga ,I'm so glad knowing that you and your family will come here next holiday. We haven't met each other for a long time . I miss you so much. If you come here, I will show a new and the biggest indoor park in the world that we call Trans Studio. It is a nice place .

You said you and your family would go by ship. However. I don't know when you will come here exactly. Please tell the time, so I can pick you up at Soekarno- Hatta harbor. Okay, that's all for you now. Please reply soon. My love to auntie, uncle and Guntur.

Love

Pretty

32. When will Bunga's family come to Pretty's town ?

a. Last month
c. Before new year
b. Next holiday
d. In December

33. What will Pretty show to Bunga's family ?

a. A new park
c. a new and the biggest indoor park
b. Her new place
d. a harbor

34. How would Bunga's family go to Pretty's town ?

a. By car b. By bus c. By train d. By ship

35. What do you think the relation between Bunga and Pretty ? They are ...

a. Schoolmates b. classmates c. sisters d. cousins

36. It is a ***nice*** place . What is the synonym of the word " nice " in the sentence ?

a. Pretty b. Sweet c. wonderful d. bitter

A radio can be played into an electrical outlet or get electric power from batteries. The main parts of a radio are antenna, tuner, and speaker.

The antenna, or aerial, is a piece of wire or metal.. An antenna can pick up radio waves. The Antenna may be inside the radio, or part of it may be outside the radio ,but connected to it.

A radio's turner lets the radio pick up signals from individual radio stations. Different radio station give off different electric signal, called frequencies.

37. The text talks about . . .

a. The description of a radio
c. The parts of radio
b. How to operate a radio
d. The use of a radio

38.The main parts of a radio consists of . . .

a. Batteries, antenna, and turner
c. antenna, turner and speaker
b. Electric power, turner and speaker
d. antenna, waves and frequencies

39. Pick up radio wage is the function of

a. A radio 's antenna
c. a radio 's turner
b. A radio's speaker
d. A radio's aerial

40. What is the main idea of the first paragraph ?

a. A radio can be played into an electrical outlet or get electric power from batteries.
b. The main parts of a radio are antenna, tuner, and speaker.
c. The antenna, or aerial, is a piece of wire or metal
d. A radio's turner lets the radio pick up signals from individual radio stations.

41. The main idea of the second paragraph is . . .

a. The antenna is a piece of wire or metal
b. The antenna may inside the radio
c. the antenna can pick up radio waves
d. the antenna may be outside the radio

42. . . . but connected to it (paragraph 2 line 2). What does the word "it " here refer to ?

a. Radio
b. Turner
c. Antenna
d. speaker

Mug Bean Porridge

Ingredients :

- One cup of mug bean
- 4 tablespoonfuls of sugar
- ½ teaspoonful of salt
- 5 cups of water

Steps :

1. Wash the mug bean
2. Boil the water
3. Put the mug bean into the boiling water
4. Let the mug bean become tender
5. Add the salt and the sugar
6. Simmer for 30 minutes
7. Now you can serve it

43. What should we do after we let the mug bean became tender ?

a. Add five cup of water
b. Add the salt and the sugar
c. Simmer for 30 minutes
d. Put the mug bean into the boiling water

44 .How many steps do we need to make bean Porridge ?

a. Four b. Five c. Six d. Seven

45 " Let the mug bean become tender " (step 4). The underlined words means ....

a. Easy to cut of chop
b. Very hard to chop
c. good taste and smell
d. very done

. 46. "Now you can serve it " (Step 7) .What does the word it here refer to ?

a. Cup of mug bean
b. Mug bean porridge
c. Boiling water
d. table spoonful of sugar

47. What is the text about ?

a. Serving mug bean porridge
b. Making mug bean porridge
c. Describing mug bean porridge
d. Explaining mug bean porridge

How to play Snake and Ladder game

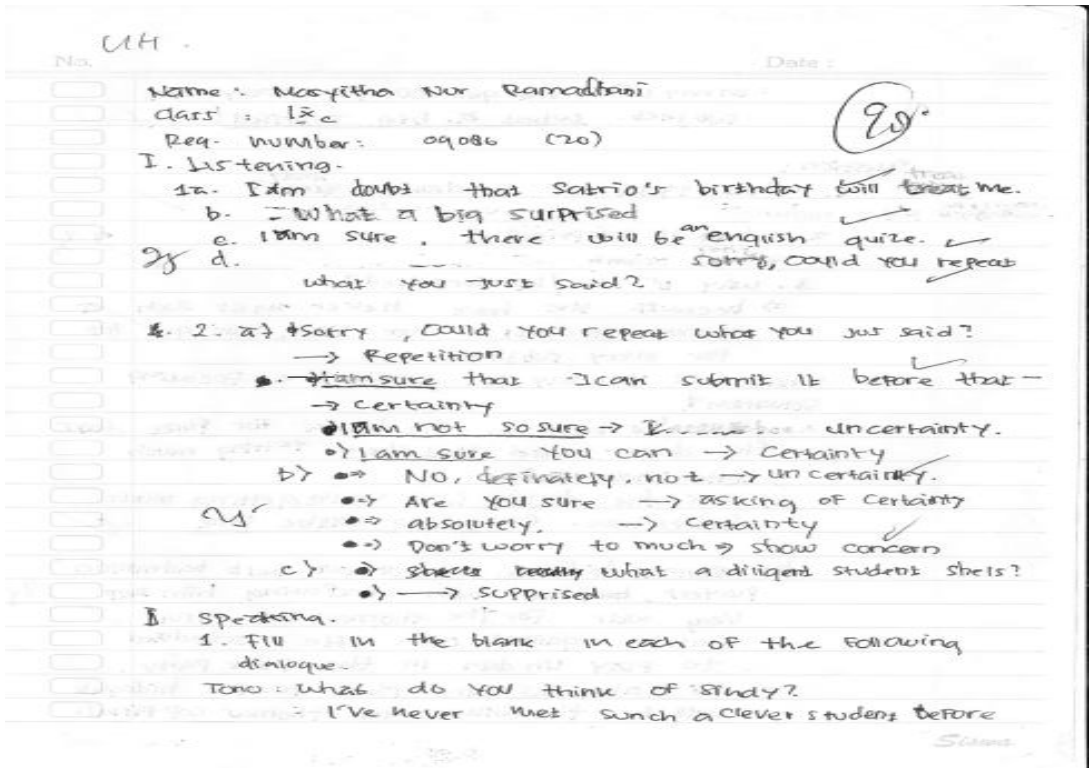
To play the game we need : a board, some counters and a dice

How to play it :

- First of all, give every player a counter.
- Next, make a toast to decided which player will start the turn.
- Then the one who wins the toast throw the dice. If he throws four, he will jump to square for but if he throws two jump to square two. When there is a ladder in the square he can go up the ladder but when there a snake he must go down.
- Have the second player to throw the dice make step based on the throw.
- Make steps until the finish. The one who gets the finish square first is the winner.

48. What is the purpose of the text ?
    a. Describing a kind of game.
    b. Telling the way to play a game.
    c. Informing a children game.
    d. Giving entertainment to reader.
49. What do you do after making a toast?
    a. Opening the board.
    b. Moving the Ladder.
    c. Killing the snake.
    d. Throwing the dice.
50. What does " dice " in the text mean ?
    a. Small cube with dots on each side.
    b. Big cubes with picture on it.
    c. Small box with spots.
    d. Big box with carving on it.

Teacher test IM group



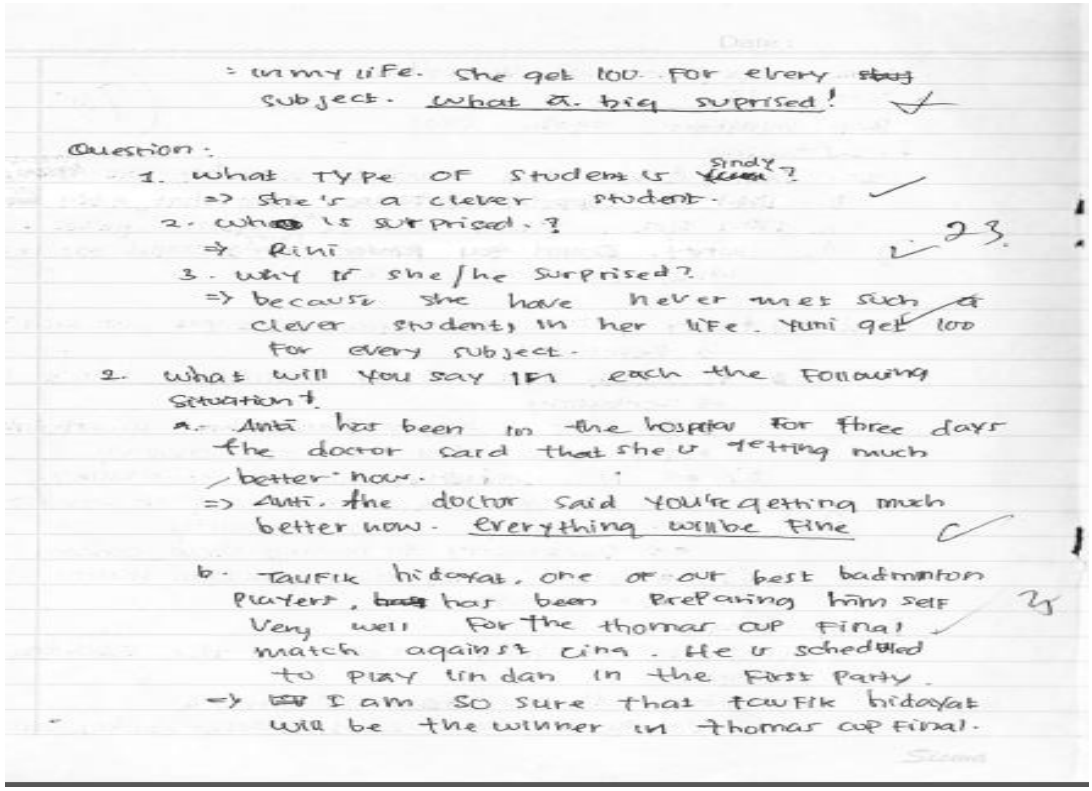
UH

Name : Nasyitha Nur Ramadhani
Class : IX c
Reg. number : 09086 (20)

I. Listening.
1a. I am doubt that Satrio's birthday will treat me.
b. What a big surprised
c. I am sure, there will be an english quize.
d. Sorry, could you repeat what you just said?

2. a) Sorry, could you repeat what you just said?
→ Repetition
b. I am sure that I can submit it before that —
→ certainty
• I am not so sure → uncertainty.
• I am sure you can → Certainty
b) • No, definately, not → uncertainty.
• Are you sure → asking of certainty
• absolutely. → Certainty
• Don't worry to much → show concern
c) • What a diligent student she is?
• → Supprised

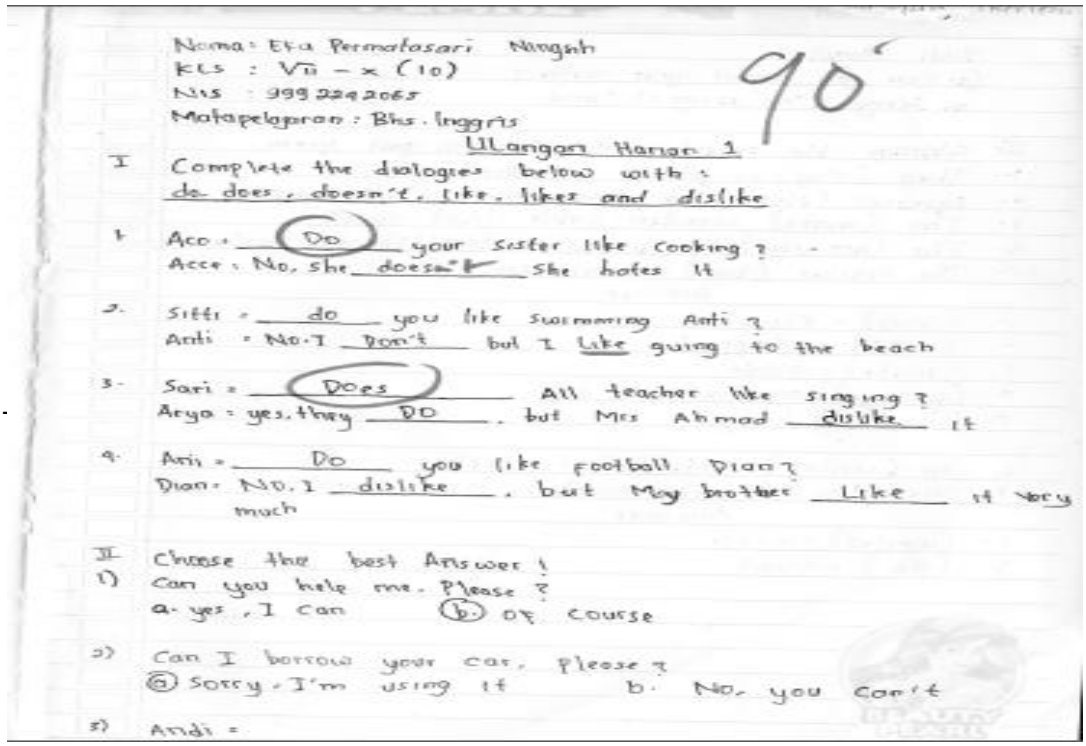
II Speaking.
1. Fill in the blank in each of the following dialoque.
Tono : What do you think of Sindy?
: I've never meet Sunch a clever student before



: in my life. She get 100 for every subject. What a big suprised!

Question :
1. What type of student is Sindy?
=> She's a clever student.
2. Who is surprised.?
=> Rini
3. Why is she/he surprised?
=> because she have never meet such a clever student, in her life. Yuni get 100 for every subject.
2. What will you say if each the following situation?
a. Anti has been in the hospital for three days the doctor said that she is getting much better now.
=> Anti, the doctor said you're getting much better now. Everything will be fine
b. Taufik hidayat, one of our best badminton players, has been preparing himself Very well For the thomar cup final match against Cina. He is scheduled to play lin dan in the first party.
-> I am so sure that taufik hidayat will be the winner in thomar cup final.

Teacher test IR group



Nama : Efa Permatasari Nasution
Kls : VII – x (10)
Nis : 999 229 2065
Mata pelajaran : Bhs. Inggris
90'
Ulangan Harian 1

I Complete the dialogues below with :
do, does, doesn't, like, likes and dislike

1. Aco : Do your sister like cooking ?
Acce : No, she doesn't she hates it

2. Sitti : do you like swimming Anti ?
Anti : No. I don't but I like going to the beach

3. Sari : Does All teacher like singing ?
Arya : yes, they DO . but Mrs Ahmad dislike it

4. Ani : Do you like football Dian ?
Dian : No. I dislike . but My brother Like it very much

II Choose the best Answer !
1) Can you help me. Please ?
a. yes, I can (b.) of course

2) Can I borrow your car, Please ?
(a) Sorry, I'm using it b. No, you can't

3) Andi =

Budi : Mama,
(a) Can you repeat your name ?
b. Sorry I'm using it now

III Change the verbs in blacket into past Form
1. They (play) in jakarta yesterday
2. Scientist (stop) the experiment last year
3. The (make) wroden table last week
4. The (arrange) the books in the library yesterday
5. The teacher (buy) a microssoft

Answer
1. (Play) = Played
2. (stop) = stopped
3. (make) = made
4. (arrange) = arranged
5. (buy) = bought

6. He (writes) letter last week
7. We (go) to the market last night

Answer
6. (writes) = wrote
7. (go) = went

SOAL ULANGAN HARIAN I

SEMESTER II TP. 2010/2011

1. Darma : ***Excuse me, can I borrow the magazine?***
   Rudy : Here, you are.
   The italic utterance expresses _____
   a. Gratitude b. politeness c. asking opinion d. asking repetition
2. The following is a sentence that expresses politeness
   a. Please, come on time!
   b. Could I pass this way?
   c. Come on, don't cry baby!
   d. What did you say?
3. Laila told poppy not to disturb her politely
   The following might be her utterance
   a. Don't bother me, please!
   b. Disturb me, please!
   c. Don't worry, be happy!
   d. Don't leave me alone!

Dear Lila

How are you ? I'm doing good here. Last weel I went to Batam Island for business. I'm so sorry I could not come to Jakarta because it was only a short visit. I only stayed for one night. My colleague picked me up at the airport. Then, we continued our journey to the hotel. It was a very hot day. I remembered that I drank five cold drinks in one hour. I also visited Nagoya,Batam centre and Galang Island.I didn't forget to go shopping there.

Okay, Lila. That 's enough for now. Don't forget to write me.

Love,

Susan

B. For no.4 -7 answer these questions below based on the text above

4. Susan didn't come to Jakarta because . . .
   a. Her colleague picked her up at the airport
   b. She had a lot of things to do in Batam
   c. she didn't know Jakarta
   d. it was only a short visit
5. From the passage we know that Susan . . .
   a. Had a vacation in Batam Island
   b. Went to Batam for business
   c. went to Jakarta for business
   d. stayed in Batam
6. "I didn't forget to go shopping there.
   "What does the word "there "in the sentence refer to?
   a. Jakarta
   b. The airport
   c. Batam Island
   d. Nagoya, Batam centre and Galang Island
7. We continued our ***journey*** to the hotel. The synonym of the word " journey " is . . .
   a. Come b. trip c. visit d. stay

C. For no. 8 - 9 write utterance using expression of politeness based on the provided situation!

8. You want to borrow your friend's money.
9. The room is very cold. You ask your friend to switch off the fan.
10. Write a letter to your teacher tells him/ her that you can't come to school because you are sick.

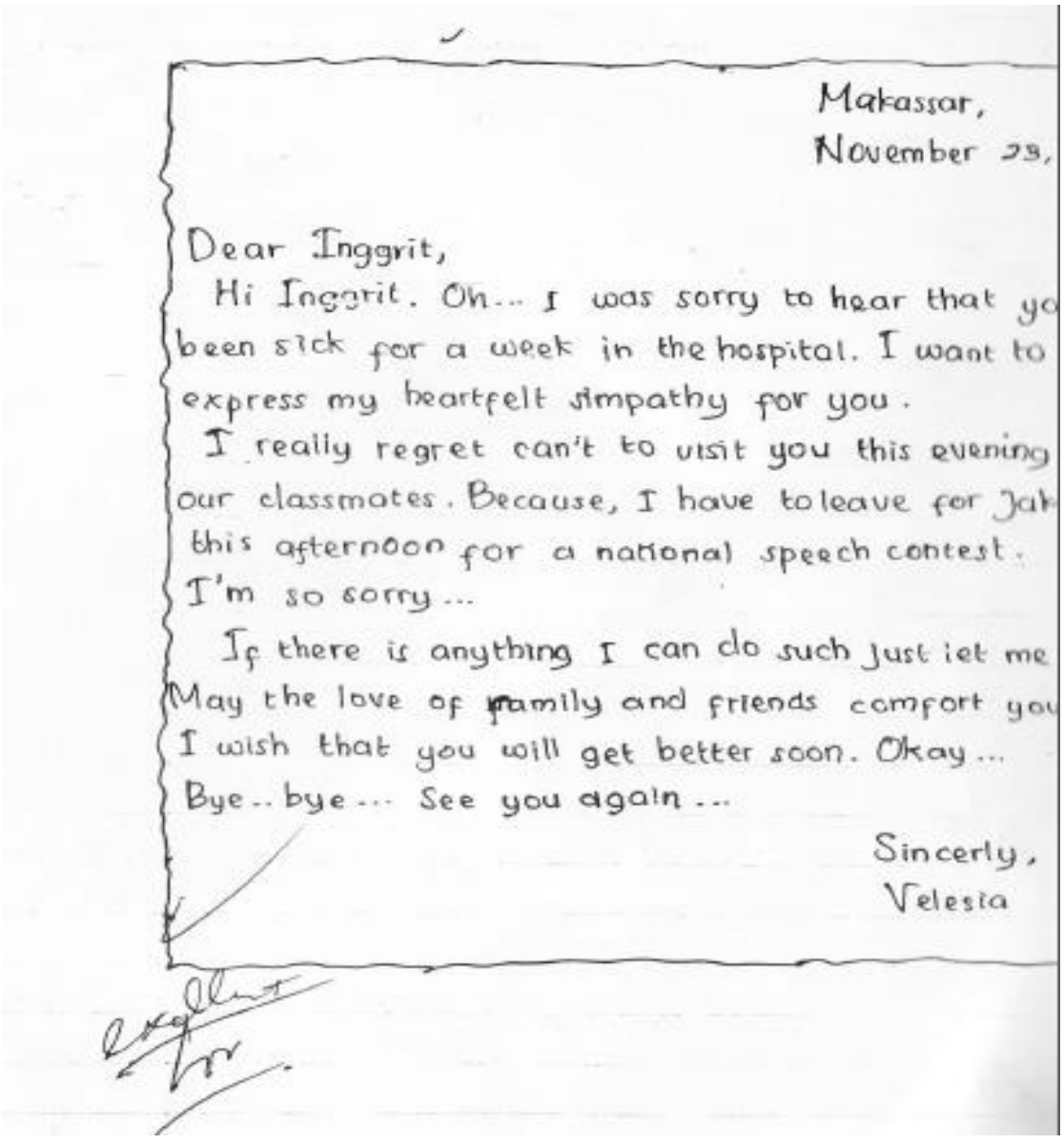
Student assignment IM group

Activity. 4. make a letter.

One of your classmates has been sick in the hospital for a week. The class has decided to visit him / her this evening, but you cannot go with them because you have to leave for Jakarta this afternoon for a national speech contest. Write a letter telling your regret and wish that she / he will get better soon.

Answer:

→



Makassar,
November 29,

Dear Inggrit,

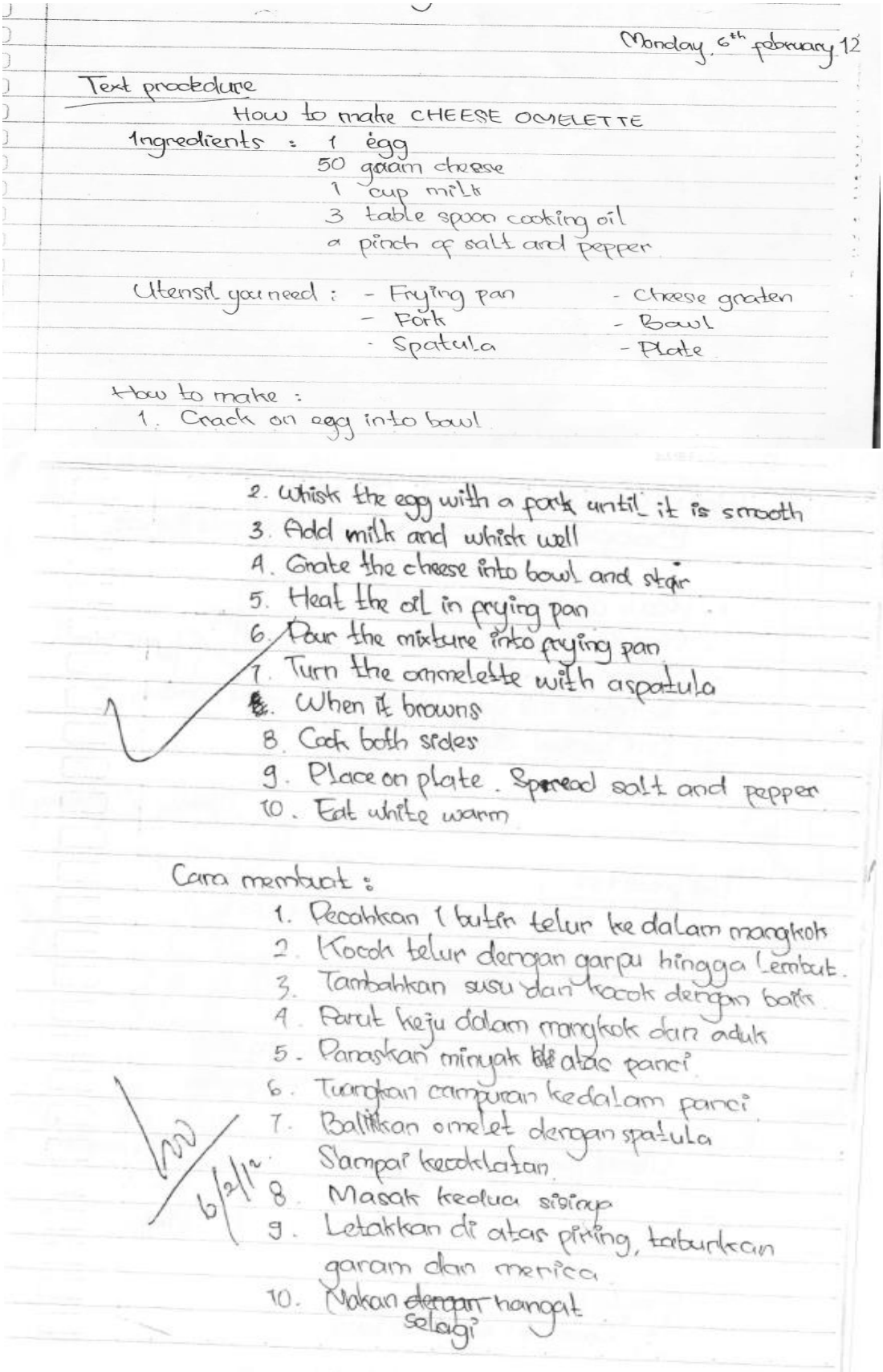
Hi Inggrit. Oh... I was sorry to hear that yo
been sick for a week in the hospital. I want to
express my heartfelt simpathy for you.

I really regret can't to visit you this evening
our classmates. Because, I have to leave for Jak
this afternoon for a national speech contest.
I'm so sorry...

If there is anything I can do such just let me
May the love of family and friends comfort you
I wish that you will get better soon. Okay...
Bye.. bye... See you again...

Sincerly,
Velesia

Student assignment IR group



Monday, 6th february 12

Text procedure

How to make CHEESE OMELETTE

Ingredients : 1 egg
50 garam cheese
1 cup milk
3 table spoon cooking oil
a pinch of salt and pepper

Utensil you need : - Frying pan - Cheese graten
- Pork - Bawl
- Spatula - Plate

How to make :
1. Crack on egg into bawl
2. Whisk the egg with a pork until it is smooth
3. Add milk and whisk well
4. Grate the cheese into bowl and stair
5. Heat the oil in frying pan
6. Pour the mixture into frying pan
7. Turn the ommelette with aspatula
8. When it browns
8. Cook both sides
9. Place on plate. Spread salt and pepper
10. Eat white warm

Cara membuat :
1. Pecahkan 1 butir telur ke dalam mangkok
2. Kocok telur dengan garpu hingga lembut.
3. Tambahkan susu dan kocok dengan baik
4. Parut keju dalam mangkok dan aduk
5. Panaskan minyak di atas panci
6. Tuangkan campuran kedalam panci
7. Balikkan omelet dengan spatula
Sampai kecoklatan
8. Masak kedua sisinya
9. Letakkan di atas piring, taburkan garam dan merica
10. Nikmati hangat selagi

Student assignment AC group

VOLLY BALL

When I was fourteen years old, where I was second years of junior high school , I always got ridicule from my friends . because I don't dominate in sports subject, every sport subject I never join it. I don't like with ridicule from my friends, I often shy with them

Two days ago, my perents call me for talking about my problems and actually I don't honest to them, because push from my parents, fanally I telling about my trouble after listening my story, they were gave suggestions for me. So, they were want me had keep up with the volley team in my school. They had presented something for me.

At one time. I hate sport but now I very like it, although the changed occurred because be forced homever I always satisfied.

Appendix G: Ethical Approval

FACULTY OF EDUCATION TE WHĀNAU O AKO PAI
DONALD STREET PO Box 17 310, Karori 6147, Wellington, New Zealand
Phone +64-4-463 9500 Fax +64-4-463 9649 Website www.vuw.ac.nz/education

12 October 2011

Astuti Azis
PhD Student
Victoria University of Wellington Faculty of Education
C/- School of Educational Psychology and Pedagogy
Donald Street
Wellington

Dear Astuti

RE: Ethics application SEPP/2011/70: RM 18851

I am pleased to advise you that your ethics application **'Investigating Indonesian junior high school teachers' conceptions of assessment: A mixed methods study'**, with requested amendments, has been approved by the Victoria University of Wellington Faculty of Education Ethics Committee. Please note that the approval for your research to commence is from the date of this letter.

Best wishes for your research.

Yours Sincerely

Dr Sue Cornforth

Co-Convener
Victoria University of Wellington Faculty of Education Ethics Committee

Appendix H: Permission letter from the Province

PEMERINTAH PROVINSI SULAWESI SELATAN
BADAN PENELITIAN DAN PENGEMBANGAN DAERAH
Jalan Urip Sumohardjo No. 269 Telp. 436936-436937 FAX. 436934
Makassar (90231)

Makassar, 14 Oktober 2011

Kepada

Nomor : 070.5.1/ 1063g /Balitbangda
Lampiran : -
Perihal : **Izin/Rekomendasi Penelitian**

Yth. Bupati Gowa

di-

Sungguminasa

Berdasarkan surat School of Educational Psychology and Pedagogy Faculty of Education Victoria Univ. of Wellington, New Zealand tanggal 14 Agustus 2011 perihal tersebut diatas, mahasiswa/peneliti dibawah ini :

Nama	: **Astuti Azis, M.Ed**
Nomor Pokok	: 300228460
Program Studi	: School of Educational Psychology and Pedagogy Faculty of Education Victoria Univ. of Wellington, New Zealand
Pekerjaan	: Mahasiswa (S3)
Alamat	: Donald St. Karori, Wellington 6147, New Zealand

Bermaksud untuk melakukan penelitian di daerah/kantor saudara dalam rangka penyusunan skripsi/tesis, dengan judul :

"INVESTIGATING INDONESIAN JUNIOR HIGH SCHOOL TEACHERS' CONCEPTIONS OF ASSESSMENT : A MIXED METHODS STUDY (KONSEP GURU BAHASA INGGRIS SMP TENTANG PENILAIAN)"

Yang akan dilaksanakan dari : Tgl. 17 Oktober 2011 s/d 02 April 2011

Sehubungan dengan hal tersebut diatas, pada prinsipnya kami ***menyetujui*** kegiatan dimaksud dengan ketentuan :

1. Sebelum *dan* sesudah melaksanakan kegiatan, kepada yang bersangkutan melapor kepada Bupati/Walikota Cq. Kepala Bappeda/Balitbangda, apabila kegiatan dilaksanakan di Kab./Kota;
2. Penelitian tidak menyimpang dari izin yang diberikan;
3. Mentaati semua peraturan perundang-undangan yang berlaku dan mengindahkan adat istiadat setempat;
4. Menyerahkan 2 (dua) eksemplar copy hasil penelitian kepada Gubernur Sulsel.Cq. Kepala Badan Penelitian dan Pengembangan Daerah Propinsi Sulawesi Selatan;
5. Surat izin akan dicabut kembali dan dinyatakan tidak berlaku apabila ternyata pemegang surat izin ini tidak mentaati ketentuan tersebut di atas.

Demikian disampaikan untuk dimaklumi dan dipergunakan seperlunya.

a.n. KEPALA BADAN
Sekretaris

Drs. H. M. Pangurisang Parawansa M.Si
Pangkat : Pembina Tk. I
NIP : 19620414 198812 1 001

TEMBUSAN : Kepada Yth :
1. Gubernur Sulawesi Selatan di Makassar (sebagai laporan);

Appendix I: Permission letter from the Regent of Gowa

PEMERINTAH KABUPATEN GOWA
BADAN KESATUAN BANGSA, POLITIK & LINMAS
Jln. Mesjid Raya No. 30. Telepon. 884637. Sungguminasa – Gowa

Sungguminasa, 17 Oktober 2011

Kepada

Nomor : 070/1545 /BKB.P.L/2011
Lamp :
Perihal : ***Rekomendasi Penelitian***

Yth. Kadis Diknas, Olahraga dan
Pemuda Kab.Gowa
Di-
Sungguminasa

Berdasarkan Surat Kepala Badan Penelitian dan Pengembangan Daerah (BALITBANGDA) Nomor : 070.5.1/10639/Balitbangda tanggal 14 oktober 2011 tentang Rekomendasi Penelitian.

Dengan ini disampaikan kepada saudara bahwa yang tersebut dibawah ini :

Nama	: **ASTUTI AZIS, M.Ed**
Tempat/Tanggal Lahir	: Sungguminasa, 2 Juni 1975
Jenis Kelamin	: Perempuan
Pekerjaan	: Mahasiswa (S3)
Alamat	: Bumi Pallangga Mas B1/18 Mangalli Kec.Pallangga

Bermaksud akan mengadakan Penelitian/Pengumpulan Data dalam rangka penyelesaian Disertasi di wilayah tempat/ saudara yang berjudul : ***"INVESTIGATING INDONESIAN JUNIOR HIGH SCHOOL TEACHER'S CONCEPTIONS OF ASSESSMENT : A MIXED METHOD STUDY (KONSEP GURU BAHASA INGGRIS SMP TENTANG PENILAIAN)".***

Selama : 17 Oktober s/d 2 April 2012
Pengikut : Tidak ada

Sehubungan dengan hal tersebut diatas, maka pada prinsipnya kami dapat menyetujui kegiatan tersebut dengan ketentuan :

1. Sebelum dan sesudah melaksanakan kegiatan kepada yang bersangkutan harus melapor kepada Bupati Cq. Kepala Badan Kesatuan Bangsa, Politik dan Linmas Kab. Gowa
2. Penelitian tidak menyimpang dari izin yang diberikan
3. Mentaati semua peraturan perundang-undangan yang berlaku dan mengindahkan adat istiadat setempat
4. Menyerahkan 1 (Satu) Eksemplar copy hasil kegiatan kepada Bupati Gowa Cq. Kepala Badan Kesatuan Bangsa, Politik dan Linmas Kab. Gowa

Demikian disampaikan dan untuk lancarnya pelaksanaan dimaksud diharapkan bantuan seperlunya.

An. KEPALA BADAN
SEKRETARIS,

Drs. A. ARIFIN
Pangkat : Pembina Tk.I
NIP : 19551025 197603 1 003

Tembusan :
1. Bupati Gowa (Sebagai Laporan)
2. Muspida Kab. Gowa
3. School of Educational Psychology and Pedagogy Faculty of Education

Appendix J: Permission letter from the Education Department

PEMERINTAH KABUPATEN GOWA
DINAS PENDIDIKAN OLAH RAGA DAN PEMUDA KAB. GOWA
Jl. Mesjid Raya No. 30 ☎ (0411) 867774 Sungguminasa – Gowa

Sungguminasa, 18 oktober 2011

Nomor : 870/1214 Dikorda/X/2011
Lamp :
Perihal : Rekomendasi Penelitian

K e p a d a
Yth. Para Ketua MGMP BERMUTU
Bahasa Inggris
Di- Sungguminasa

Berdasarkan Surat Kepala Badan Kesatuan Bangsa, Politik & Linmas (Kesbang) Nomor : 070/2545/BKB.P.L/2011 Kesbang tanggal 17 Oktober 2011 tentang Rekomendasi Penelitian.

Dengan ini diampaikan kepada saudara bahwa yang tersebut dibawah ini :

Nama	: **Astuti Azis, M.Ed**
Tempat/Tanggal Lahir	: Sungguminasa, 2 Juni 1975
Jenis Kelamin	: Perempuan
Pekerjaan	: Mahasiswa (S3)
Alamat	: Bumi Pallangga Mas B1/18 Mangngalli Kec. Palangga

Bermaksud mengadakan Penelitian/Pengumpulan data dalam rangka penyelesaian disertasi di kelompok MGMP saudara yang berjudul : **"Investigating Indonesian Junior High School Teachers' Conceptions of Assessment: A Mixed Methods Study (Konsep Guru Bahasa Inggris SMP Tentang Penilaian)".**

Waktu	: 19 Oktober s/d 2 April 2012
Pengikut	: Tidak ada

Sehubungan dengan hal tersebut diatas, maka pada prinsipnya kami dapat menyetujui kegiatan tersebut dengan ketentuan:

1. Sebelum dan sesudah melaksanakan kegiatan, yang bersangkutan harus melapor kepada Kepala Dinas Pendidikan Olahraga dan Pemuda Kab. Gowa
2. Penelitian tidak menyimpang dari izin yang diberikan
3. Mentaati semua peraturan perundang-undangan yang berlaku dan megindahkan adat istiadat setempat
4. Menyerahkan 1 (satu) eksemplar copy hasil kegiatan kepada Kepala Dinas Pendidikan Olahraga dan Pemuda Kab. Gowa cq. Kepala Bidang Pendidikan Dasar Dikorda Kab. Gowa

Demikian disampaikan dan untuk lancarnya pelaksanaan dimaksud diharapkan bantuan seperlunya.

Kepala Dinas Pendidikan, Olahraga
dan Pemuda Kabupaten Gowa,

H. Idris Faisal Kadir, S.H
Pangkat : Pembina Utama Muda
NIP : 19640629189121001

Tembusan:
1. Bupati Gowa (Sebagai Laporan)
2. School of Educational Psyhology and Pedagogy Faculty of Education

Appendix K: Letter to PD Leader

VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui

September 2011

Dear……,

My name is Astuti Azis and I am doing my PhD in Education at Victoria University of Wellington New Zealand.

I am currently working on my Doctor of Philosophy dissertation. The focus of my study is on "Indonesian junior high school teachers' conceptions of assessment'. My research question is: *How do Indonesian junior high school teachers understand the role of assessment and how do they perceive these understandings to promote students learning?*

I have been granted permission by the Head of Education Department of Gowa Regency to survey and interview junior high school teachers of English. I would appreciate if you would allow me to conduct a survey with teachers while they are meeting to undertake Professional Development (PD) in your group. I would like to spend approximately 30 minutes with these teachers at a time convenient to you.

Please indicate your agreement by emailing me at astuti.azis@vuw.ac.nz I will follow up this letter with an email in a week time.

Many thanks for your support.

Yours sincerely

Astuti Azis

Appendix L: Letter to teacher (phase 1)

VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui

September 2011

Dear……,

My name is Astuti Azis and I am doing my PhD in Education at Victoria University of Wellington New Zealand.

I am currently working on my Doctor of Philosophy dissertation and would like you to take part in the interview of my research. The focus of my study is on 'Indonesian junior high school teachers' conceptions of assessment'. My research question is: '*How do Indonesian junior high school teachers understand the role of assessment and how do they perceive these understandings to promote students learning'?*

I have been granted permission by the Head of Education Department of Gowa Regency to survey and interview junior high school teachers of English. I have also informed your PD leader regarding this permission. I would like to invite you to participate in the first phase of my study: a 15 minutes survey on teachers' conceptions of assessment.

Your participation is entirely voluntary, and you may withdraw from this study at any time before data analysis begins. Your confidentiality is strictly assured and any names used in the final report will be pseudonyms. All data will be stored in a locked cabinet and will be destroyed after 3 years.

I attach an information sheet with further details. Please read this sheet and if you are willing to participate, please complete the consent form. I will be back in a week's time to confirm whether or not you wish to participate. I will collect the consent form.

Should you wish further information please contact me, on 085242480530,
tuty_azis@yahoo.com or my supervisors, Dr. Margaret Gleeson,
Margaret.gleeson@vuw.ac.nz or Prof. Luanna H Meyer, luanna.meyer@vuw.ac.nz at the Faculty of Education, Victoria University of Wellington, New Zealand.

Many thanks for your support.

Yours sincerely

Astuti Azis

Appendix M: Letter to teacher (phase 2)

VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui

Letter to Teacher (Phase 2)
September 2011
Dear……,

My name is Astuti Azis and I am doing my PhD in Education at Victoria University of Wellington New Zealand.

I am currently working on my Doctor of Philosophy dissertation and would like you to take part in the interview of my research. The focus of my study is on 'Indonesian junior high school teachers' conceptions of assessment'. My research question is: '*How do Indonesian junior high school teachers understand the role of assessment and how do they perceive these understandings to promote students learning'?*

Your response on the questionnaire shows a strong preference in one particular conception of assessment. Such a preference indicates that you will be able to provide rich information for my study. I would like to invite you to participate in the second phase of my study: to understand your values of assessment, factors contributing to your conceptions and your perceptions on how your conceptions reflect your assessment practices.

The interviews will be audio-recorded, and might be followed by additional interviews to clarify points that arise. A summary of your interviews will be available to you to check for accuracy. Your confidentiality is assured and any names used in the final report will be pseudonyms. All data will be stored in a locked cabinet and will be destroyed after 3 years.

I attach an information sheet with further details. Please read this sheet and if you are willing to participate, please complete the consent form. I will phone you in a week's time to confirm whether or not you wish to participate. I will collect the consent form. Should you wish further information please contact me, on 085242480530, astuti.azis@vuw.ac.nz or my supervisors, Dr. Margaret Gleeson, Margaret.gleeson@vuw.ac.nz or Prof. Luanna H Meyer, luanna.meyer@vuw.ac.nz at the Faculty of Education, Victoria University of Wellington, New Zealand.

Many thanks for your support.

Yours sincerely

Astuti Azis