The practice of formative assessment by EFL teachers in secondary high schools in Indonesia

Maya Defianty

Thesis submitted for the degree of Doctor of Philosophy in Education

University of Canberra

2018
The mediocre teacher tells
The good teacher explains
The superior teacher demonstrates
The great teacher inspires

William Arthur Ward
Abstract

Based on Black and William’s seminal publication in 1998, a number of countries have implemented formative assessment as a means of improving student learning. Formative assessment was adopted in Indonesia more than two decades ago. Recently, the notion has been strengthened in the current curriculum, Kurikulum 2013, which explicitly emphasises students’ role in the process of assessment. Unfortunately, the notion of formative assessment, as stipulated in the Kurikulum 2013, was criticised by teachers, who claimed that it is a heavy burden and too complicated to implement. These criticisms may derive from a number of factors; the most conspicuous of which is that teachers lack information about formative assessment. To date, there have been only a few studies that focus on formative assessment in English language teaching in Indonesia, and no studies that specifically investigate how formative assessment can be implemented in the Indonesian context.

This study aims to explore how formative assessment can be carried out and to find examples of good practice which can inspire other teachers to adopt or adapt the practice. A multiple case study design was employed to investigate teachers’ understanding of formative assessment, their application of this understanding to their teaching practice, and the way they overcome any challenges encountered. Eight secondary school English teachers participated in this study: four teachers from schools under the Ministry of Education and Culture (MoEC) and four teachers from schools under the Ministry of Religious Affairs (MoRA). The teachers who agreed to participate were selected based on the following criteria: they have had training and experience in implementing formative assessment, and they were recommended by their peers in the secondary high school teachers’ association committee. Data from interviews, classroom observations, and artefacts were analysed using a framework derived from Wiliam and Leahy’s (2015) categorisation of formative assessment strategies, which included clarifying, sharing, and understanding learning
intentions and successful criteria; engineering effective discussions, tasks and activities that elicit evidence of learning; providing feedback that moves learning forward; activating students as learning resources for one another; and activating students as owners of their own learning (Oswalt, 2013; Thompson & Wiliam, 2008; Wiliam & Leahy, 2015).

The findings from the study showed that teachers have diffuse understandings in defining formative assessment; however, they agree on other concepts of formative assessment such as its potential benefits and key elements. Furthermore, they add another element which can only be found in the area of formative assessment in ELT: the integration of language macro skills. This study has shown that teachers understand the notion of formative assessment in two different ways: firstly, teachers use the results of formative assessment to adjust their teaching in order to enhance learning, and secondly, they implement formative assessment strategies in line with their teaching and learning context, thus developing their own unique formative assessment culture. This study has found that the teachers who participated in this study were able to provide excellent examples of how formative assessment can be carried out with the aid of instruments such as three-lines writing, public speaking, video-conferencing and so on. In addition, this study has identified that teachers were able to make room for formative assessment despite the testing culture that is entrenched in the Indonesian education system. Nevertheless, some of the teachers’ practices needed to be refined in order to gain the potential benefits of formative assessment. This study also found that teachers were able to overcome challenges that they encountered during the implementation of formative assessment, such as large class sizes and limited time, by integrating educational technology and modifying teaching methods. Therefore, the results of this study suggest that formative assessment can permeate teaching practice, and that teachers in Indonesia have the potential to reap its benefits.
# Table of Contents

ABSTRACT .................................................................................................................................................. iii  
CERTIFICATE OF AUTHORSHIP OF THESIS .......................................................................................... v  
ACKNOWLEDGEMENTS ........................................................................................................................... vii  
TABLE OF CONTENTS ........................................................................................................................... xi  
LIST OF TABLES .......................................................................................................................................... xv  
LIST OF FIGURES ......................................................................................................................................... xvii  
LIST OF ABBREVIATIONS ......................................................................................................................... xix  

## CHAPTER 1 INTRODUCTION .................................................................................................................... 1  
1.1 BACKGROUND OF THE STUDY ............................................................................................................ 5  
1.2 STATEMENT OF THE RESEARCH PROBLEM .................................................................................... 11  
1.2.1 RESEARCH AIMS AND RESEARCH QUESTIONS ........................................................................... 14  
1.2.2 SIGNIFICANCE OF THE STUDY ................................................................................................... 15  
1.3 RESEARCH METHODOLOGY ............................................................................................................. 16  
1.4 STRUCTURE OF THE THESIS ............................................................................................................ 17  

## CHAPTER 2 LITERATURE REVIEW ......................................................................................................... 19  
2.1 OUTLINING ASSESSMENT .................................................................................................................... 21  
2.2 UNDERSTANDING FORMATIVE ASSESSMENT ............................................................................... 26  
2.3 IMPLEMENTING FORMATIVE ASSESSMENT ...................................................................................... 36  
2.3.1 CLARIFYING, SHARING, AND UNDERSTANDING LEARNING INTENTIONS AND SUCCESSFUL CRITERIA .......................................................................................................................... 37  
2.3.2 ENGINEERING EFFECTIVE CLASSROOM DISCUSSION, QUESTIONS, AND LEARNING TASKS .................................................................................................................................................. 39  
2.3.2.1 QUESTIONING .............................................................................................................................. 39  
2.3.2.2 TESTS ........................................................................................................................................ 42  
2.3.3 PROVIDING EFFECTIVE FEEDBACK THAT MOVES LEARNING FORWARD .................................. 44  
2.3.3.1 FORMATIVE FEEDBACK ............................................................................................................ 46  
2.3.4 ACTIVATING STUDENTS AS INSTRUCTIONAL RESOURCES FOR ONE ANOTHER .................. 50  
2.3.5 ACTIVATING STUDENTS AS THE OWNERS OF THEIR OWN LEARNING .................................. 55  
2.4 IDENTIFYING FORMATIVE ASSESSMENT CHALLENGES AND THEIR POTENTIAL SOLUTIONS ....................................................................................................................................... 59  
2.4.1 UNCOVERING FORMATIVE ASSESSMENT CHALLENGES .......................................................... 60  
2.4.2 COPING WITH CHALLENGES ....................................................................................................... 64
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.3.1 BELIEFS ABOUT FEEDBACK</td>
<td>158</td>
</tr>
<tr>
<td>4.2.3.2 GIVING FEEDBACK</td>
<td>161</td>
</tr>
<tr>
<td>4.2.4 ACTIVATING STUDENTS AS LEARNING RESOURCES FOR ONE ANOTHER</td>
<td>168</td>
</tr>
<tr>
<td>4.2.5 ACTIVATING STUDENTS AS THE OWNERS OF THEIR OWN LEARNING</td>
<td>172</td>
</tr>
<tr>
<td>4.3 CHALLENGES AND SOLUTIONS</td>
<td>125</td>
</tr>
<tr>
<td>4.3.1 IDENTIFYING FORMATIVE ASSESSMENT CHALLENGES</td>
<td>176</td>
</tr>
<tr>
<td>4.3.2 OVERCOMING FORMATIVE ASSESSMENT CHALLENGES</td>
<td>178</td>
</tr>
<tr>
<td>4.4 SUMMARY</td>
<td>181</td>
</tr>
<tr>
<td>CHAPTER 5 DISCUSSION</td>
<td>183</td>
</tr>
<tr>
<td>5.1 SURVIVING THE FUZZINESS OF DEFINING FORMATIVE ASSESSMENT</td>
<td>184</td>
</tr>
<tr>
<td>5.2 APPLYING FORMATIVE USE OF TESTS</td>
<td>190</td>
</tr>
<tr>
<td>5.3 BUILDING FORMATIVE ASSESSMENT CULTURE</td>
<td>197</td>
</tr>
<tr>
<td>5.4 OVERCOMING FORMATIVE ASSESSMENT BARRIERS</td>
<td>203</td>
</tr>
<tr>
<td>5.4.1 IDENTIFYING THE ACTUAL FORMATIVE ASSESSMENT BARRIERS</td>
<td>203</td>
</tr>
<tr>
<td>5.4.2 PROVIDING SOLUTIONS</td>
<td>208</td>
</tr>
<tr>
<td>5.5 MISSING FORMATIVE ASSESSMENT OPPORTUNITIES</td>
<td>202</td>
</tr>
<tr>
<td>5.5.1 MISSING FORMATIVE ASSESSMENT OPPORTUNITIES: QUESTIONING</td>
<td>213</td>
</tr>
<tr>
<td>5.5.2 MISSING FORMATIVE ASSESSMENT OPPORTUNITIES: PEER ASSESSMENT</td>
<td>215</td>
</tr>
<tr>
<td>5.5.3 MISSING FORMATIVE ASSESSMENT OPPORTUNITIES: CLARIFYING, SHARING, AND UNDERSTANDING EVIDENCE OF LEARNING AND SUCCESS CRITERIA</td>
<td>217</td>
</tr>
<tr>
<td>5.6 SUMMARY</td>
<td>224</td>
</tr>
<tr>
<td>CHAPTER 6 EMBRACING FORMATIVE ASSESSMENT</td>
<td>227</td>
</tr>
<tr>
<td>6.1 RESEARCH SUMMARY</td>
<td>227</td>
</tr>
<tr>
<td>6.2 SUMMARY OF RESEARCH FINDINGS</td>
<td>230</td>
</tr>
<tr>
<td>6.2.1 TEACHERS’ UNDERSTANDING OF FORMATIVE ASSESSMENT</td>
<td>230</td>
</tr>
<tr>
<td>6.2.2 FORMATIVE ASSESSMENT IMPLEMENTATION</td>
<td>230</td>
</tr>
<tr>
<td>6.2.3 FORMATIVE ASSESSMENT CHALLENGES AND SOLUTIONS</td>
<td>231</td>
</tr>
<tr>
<td>6.3 RESEARCH CONTRIBUTIONS</td>
<td>232</td>
</tr>
<tr>
<td>6.4 IMPLICATIONS OF THE STUDY</td>
<td>234</td>
</tr>
<tr>
<td>6.5 LIMITATIONS OF THE STUDY</td>
<td>236</td>
</tr>
</tbody>
</table>
6.6 RECOMMENDATIONS ........................................................................ 237
6.7 REFLECTION .................................................................................. 239

REFERENCES ......................................................................................... 241
APPENDIX 1 APPROVAL FROM ETHICS COMMITTEE .......................... 271
APPENDIX 2 SAMPLE OF INSTITUTION APPROVAL .............................. 273
APPENDIX 3 INFORMATION SHEET .................................................. 275
APPENDIX 4 CONSENT FORM FOR TEACHER .................................... 277
APPENDIX 5 CONSENT FORM FOR STUDENTS ................................. 279
APPENDIX 6 INTERVIEW PROTOCOLS ............................................. 281
APPENDIX 7 OBSERVATION PROTOCOLS ......................................... 283
APPENDIX 8 SAMPLE OF ASSESSMENT CRITERIA APPLIED BY ONE OF THE PARTICIPANTS .......................................................... 285
APPENDIX 9 SAMPLE OF LESSON PLAN .......................................... 287
APPENDIX 10 ASSESSMENT INSTRUMENTS APPLIED BY PARTICIPANTS ... 297
APPENDIX 11 SOME EXAMPLES OF FORMATIVE ASSESSMENT PRACTICE APPLIED BY ONLY ONE PARTICIPANT .......................... 299
List of Tables

TABLE 1.1: SCHOOLING SYSTEMS IN INDONESIA ............................................. 6
TABLE 2.1: TESTING CULTURE VERSUS ASSESSMENT CULTURE ............ 24
TABLE 2.2: DIFFERENCES BETWEEN KURIKULUM 1994 AND KBK ........... 187
TABLE 3.1: PARTICIPANTS’ PROFILES .......................................................... 107
TABLE 4.1: PARTICIPANTS’ UNDERSTANDING OF FORMATIVE ASSESSMENT ...................................................................................... 125
TABLE 4.2: FORMATIVE ASSESSMENT INSTRUMENTS APPLIED BY PARTICIPANTS ................................................................. 142
TABLE 5.1 ASSESSMENT SYSTEMS IN SECONDARY SCHOOLS IN INDONESIA ................................................................................... 193
TABLE 5.2: CONFUSED AND CLARIFIED LEARNING INTENTIONS AND SUCCESS CRITERIA ............................................................. 219
TABLE 5.3: SAMPLE OF LEARNING OBJECTIVE ....................................... 223
### List of Figures

| FIGURE 2.1 | TEST, ASSESSMENT, AND TEACHING AND LEARNING | 21 |
| FIGURE 2.2 | GRAPHICAL REPRESENTATION OF FORMAL AND INFORMAL FORMATIVE ASSESSMENT | 31 |
| FIGURE 2.3 | THE VARIATION OF FORMATIVE ASSESSMENT | 32 |
| FIGURE 2.4 | VARIATIONS OF FORMATIVE ASSESSMENT CHARACTERISTICS | 33 |
| FIGURE 2.5 | CHECKLIST FOR ASKING APPROPRIATE QUESTIONS | 43 |
| FIGURE 2.6 | FORMATIVE FEEDBACK LOOP | 49 |
| FIGURE 2.7 | PEER ASSESSMENT MODEL | 53 |
| FIGURE 2.8 | ASSESSMENT: STREAMS, TYPES, AND INSTRUMENTS | 78 |
| FIGURE 2.9 | CURRICULUM IN INDONESIA | 86 |
| FIGURE 3.1 | THE PROCESS OF DATA ANALYSIS | 120 |
| FIGURE 4.1 | FORMATIVE ASSESSMENT STRATEGIES | 134 |
| FIGURE 5.1 | LEARNING OBJECTIVES | 222 |
### List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AfL</td>
<td>Assessment for Learning</td>
</tr>
<tr>
<td>CLT</td>
<td>Communicative Language Teaching</td>
</tr>
<tr>
<td>EAL</td>
<td>English as an Additional Language</td>
</tr>
<tr>
<td>EFL</td>
<td>English as a Foreign Language</td>
</tr>
<tr>
<td>ELT</td>
<td>English Language Teaching</td>
</tr>
<tr>
<td>ESL</td>
<td>English as a Second Language</td>
</tr>
<tr>
<td>FUST</td>
<td>Formative Use of Summative Test</td>
</tr>
<tr>
<td>ICT</td>
<td>Information Communication and Technology</td>
</tr>
<tr>
<td>POSTT</td>
<td>Pedagogy of Science Teaching Test</td>
</tr>
<tr>
<td>KBK</td>
<td>Kurikulum Berbasis Kompetensi (Competence-based curriculum)</td>
</tr>
<tr>
<td>KBM</td>
<td>Kegiatan Belajar Mengajar (Teaching and learning activities)</td>
</tr>
<tr>
<td>KLT</td>
<td>Keeping Learning on Track</td>
</tr>
<tr>
<td>KKG</td>
<td>Kelompok Kerja Guru (Primary teacher working group)</td>
</tr>
<tr>
<td>KKM</td>
<td>Kriteria Ketuntasan Minimal (Criteria for minimum completion)</td>
</tr>
<tr>
<td>KKM</td>
<td>Kelompok Kerja Madrasah (Teacher working group for Madrasa teachers)</td>
</tr>
<tr>
<td>KMOFAP</td>
<td>King’s Medway-Oxfordshire Formative Assessment Project</td>
</tr>
<tr>
<td>KTSP</td>
<td>Kurikulum Tingkat Satuan Pendidikan (School-based curriculum)</td>
</tr>
<tr>
<td>MGMP</td>
<td>Musyawarah Guru Mata Pelajaran (Panel of secondary school subject teachers)</td>
</tr>
<tr>
<td>MoEC</td>
<td>Ministry of Education and Culture</td>
</tr>
<tr>
<td>MoRA</td>
<td>Ministry of Religious Affairs</td>
</tr>
<tr>
<td>SKL</td>
<td>Standar Kompetensi Lulusan (Graduate standard competence)</td>
</tr>
<tr>
<td>TFALTQ</td>
<td>Teacher Formative Assessment Literacy Questionnaire</td>
</tr>
<tr>
<td>UAS</td>
<td>Ujian Akhir Semester (End of semester examination)</td>
</tr>
<tr>
<td>UN</td>
<td>Ujian Nasional (National examination)</td>
</tr>
<tr>
<td>UTS</td>
<td>Ujian Tengah Semester (Mid-semester examination)</td>
</tr>
<tr>
<td>ZPD</td>
<td>Zone of Proximal Development</td>
</tr>
</tbody>
</table>
Acknowledgements

This PhD journey has not only been remarkable and enlightening, but also challenging. I would like to extend my gratitude to many individuals who have helped me throughout this quest.

My first and foremost gratitude goes to my primary supervisor, Kate Wilson, Ph.D. You are an inspirational mentor. Your constructive feedback and mind-stimulating questions, comments and advice have aroused my critical thinking. Thank you for your continuous guidance – a perfect balance between providing direction and encouraging independence. I would also like to thank Dr Jeremy Jones for his brilliant comments and suggestions.

My gratitude also goes to the Indonesian Ministry of Religious Affairs for their financial support which has enabled me to realize my dream of having a doctorate degree. I would like to thank Louise Baird for her translation service for the interviews and classroom interactions. I also would like to thank Kathie Brown from A+ Academic Editing, Canberra for her copyediting service. My sincere appreciation is directed to all teachers who participated in this study, thank you for your time and for allowing me to gain insight into your teaching practice. I also thank my friends and colleagues with whom I have shared this wonderful journey of being a PhD student.

Finally, I am very grateful to my beloved family. Thank you to my parents whose love, support, and endless prayers have sustained me throughout this journey. My special thanks go to my husband. How grateful I am to have you beside me, thank you for always believing in me and putting up with me through the toughest moments in this journey. Words cannot express how I cherish your love, support, and encouragement. To my daughter, you have always been my greatest strength and motivation; I dedicate this thesis to you.
Chapter 1
Introduction

My interest in assessment was aroused while I was still at school. Growing up in a country where high-stakes testing was administered at the end of every learning level meant that I took the test at the end of primary, low-secondary, and high secondary school, in addition to taking it to gain entrance to university. My constant encounters with testing and assessment coloured my whole experience of schooling. In some ways, assessment often motivated and focused my attention on the learning materials. Nevertheless, I frequently experienced anxiety, frustration, and disappointment because I felt that the single, unexplained score that the teachers assigned did not define my ability, and it discouraged me from learning. I longed for teachers’ customized feedback so that I would be able to understand my personal strengths and weaknesses. It is fair to say that the process of assessment had a significant impact on my learning and possibly on my future.

Decades later, now that I am a lecturer in a teacher training program, I notice that assessment also plays a crucial role in my students’ learning. They tend to direct their focus towards assessment-related activities, such as how the particular course is assessed or why they have a particular score. Observing this, I completely agree that ‘assessment is the engine that drives learning’ (Cowan, 1998, p. 12). It is a powerful means of directing students’ learning. However, as with other powerful tools, it needs to be used with care, or it can become counterproductive and even harmful. As Boud (1995, p. 35) argued, ‘students can escape bad teaching, but they cannot escape bad assessment’.

Essentially, assessment in education is about assessing students’ competence as reflected by their performance, for example in tasks or tests. In much of the literature, assessment is
divided into two streams: formative assessment and summative assessment. Commonly, both terms are differentiated according to purpose and time. Formative assessment is assessment that aims to improve learning, and it is conducted during learning. Summative assessment, by comparison, refers to assessment that aims to identify what students have learned and is conducted at the end of a period of learning. Although formative and summative assessments are equally important, teachers should focus on formative assessment more than summative assessment because it provides valuable information for instructional decision-making on a daily-basis (Popham, 2006; Stiggins, 2005). Black and Wiliam (1998a, 1998b) in their ground-breaking work, also proposed that teachers should focus on formative assessment because it provides the crucial information that teachers need to enhance learning such as students’ current abilities. They argue that formative assessment has the potential to improve learning, and many subsequent studies support this finding (for example, Ash & Levitt, 2003; Buck & Trauth-Nare, 2009; Cauley & McMillan, 2010; Chappuis & Chappuis, 2008; Curry, Mwavita, Holter, & Harris, 2015; Marzano, Pickering, & Pollock, 2001). Nevertheless, the notion of formative assessment has not gone uncontested; concerns are frequently raised about the underpinning theory of formative assessment and the practice entailed (Bennett, 2011; Crossouard & Pryor, 2012; Dunn & Mulvenon, 2009; Sebatane, 1998).

In their seminal work, Black and Wiliam (1998b, p. 2) defined formative assessment as

... all those activities undertaken by teachers - and by their students in assessing themselves - that provide information to be used as feedback to modify teaching and learning activities. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching to meet student needs.

Since this definition does not provide strict boundaries around what formative assessment is, a wide array of definitions with different perspectives and emphases have arisen which claim to clarify the term: Assessment for Learning (AfL); assessment and classroom teaching and learning; learning-oriented assessment; and classroom-based assessment. Nonetheless, there
has not yet been any clear agreement on how formative assessment should be defined. Further, the multiple interpretations of formative assessment have caused several misconceptions among teachers. For example, Wiliam (2011) reported that sometimes formative assessment has been misinterpreted as frequent testing; while Pinchock & Brandt (2009) said that often formative assessment is considered to be simply a set of assessment practices. In addition, Davison and Leung (2009) identify two common misinterpretations of formative assessment made by teachers coming from traditional backgrounds: first, that any continuous assessment is formative and, second, that any assessment instrument recommended by the notion of alternative assessment is also formative.

To a large extent, what allows an assessment to be categorised as formative is based on how the data is used, rather than on the practice itself (Black, 2015; Black & Wiliam, 1998b, 1998a, 2009; Greenstein, 2010; Pinchock & Brandt, 2009; Rea-Dickins, 2006). Further, Stobart (2008) argued that formative assessment should be seen as an approach, rather than a set of prescriptive techniques. Black (2015, p. 170) confirmed that ‘it is the use of evidence that distinguishes the formative and summative, although of course the methods used and the outcomes may differ according to the uses envisaged’. One can see that reaching a consensus of what formative assessment is and prescribing formative assessment techniques may not be achievable. Drawing on the numerous definitions available, in this study I argue that formative assessment can be identified by several consistent elements:

- **Process**
  
  Formative assessment is not a one-off activity, rather it is a continuing process.

- **Feedback**
  
  Teachers’ assessment should involve formative feedback which is evidence-based so that follow-up action can be taken.
• **Students’ involvement**

Assessment is not solely the teachers’ responsibility: the students themselves and their peers should actively participate in the process.

• **Teaching adjustment**

Evidence gathered from assessment should be used by teachers to adjust teaching; in fact, this is the core element that differentiates formative assessment from summative assessment.

I, therefore, define formative assessment as *a continuous activity of assessing students’ developing competencies that includes feedback and students’ involvement, and subsequent adjustment to teaching*. The elements of formative assessment which are mentioned above will be discussed further in Chapter 2.

Fortunately, although the literature lacks agreement on a definition of formative assessment and its techniques, there are several formative assessment strategies that scholars have agreed on (for example, Black, Harrison, Lee, Marshall, & Wiliam, 2003; Heritage, 2007; Leahy, Lyon, Thompson, & Wiliam, 2005; McManus, 2008; Stiggins, Arter, Chappuis, & Chappuis, 2004; Thompson & Wiliam, 2008; Wiliam & Leahy, 2015). Wiliam and Leahy (2015) for example recommend the following assessment strategies: clarifying, sharing, and understanding learning intentions and criteria for success; engineering effective classroom discussion, questions, and learning tasks; providing feedback that moves learners forward; activating students as owners of their learning; and activating students as instructional resources for one another. Nevertheless, a smorgasbord of strategies may be insufficient to support teachers in implementing formative assessment because teachers want clear, practical and hands on activities instead of a set of application guidelines (Edelenbos & Kubanek-German, 2004; Fulcher, 2012). Furthermore, the diverse definitions may also lead to
misinterpretation. For this reason, teachers need to be supported with contextualised descriptions of actual classroom practice. If the aim of this study of formative assessment is to encourage teachers to implement the practice in their setting, it is critical to provide them with examples of how formative assessment strategies can be carried out in authentic contexts. Describing and analysing such examples may provide inspiration for other teachers to adopt or adapt the practice.

1.1 Background of the study

Indonesia is an archipelago of 17,508 islands; it is the largest country by area in South-East Asia, and the fourth largest in the world (Badan Informasi Geospasial, 2014; OECD, 2015). The country is divided into 34 provinces and around 500 municipalities and/or districts, stretching to a total area of 1.9 million square kilometres (Asian Development Bank, 2014). With 261,115 million people, Indonesia is the fifth most populous country in the world (Unesco, 2018). Its diverse population consists of 600 ethnic groups, the largest of which are Javanese, Sundanese, Batak, Aceh, Malay, Bugis, and Dayak (Zuhdi, 2015). Undoubtedly, Indonesia is large and diverse.

Most Indonesians learn at least three languages. The first is the local language or indigenous language; there are 700 languages and dialects in Indonesia which can be identified based on the region. For example, Sundanese is spoken by people who live in West Java, and Javanese is spoken by people who live in East Java, Central Java, and Yogyakarta. The local language is included in the curriculum as ‘local content’. The local government manages local content in terms of policy, human resources, and curriculum development based on the Ministry of Education and Culture (hereafter MoEC) Decree on the Local Content Kurikulum 2013 (Peraturan Menteri Pendidikan dan Kebudayaan Nomor 79, 2014).
Basically, the education system in Indonesia is managed by the government under two ministries: MoEC and the Ministry of Religious Affairs (hereafter, MoRA) (MoEC, 2013). Essentially, these two ministries have different responsibilities in managing education in Indonesia; MoEC is responsible for managing public and private schools, while MoRA manages Islamic religious-based education. In practice, schools under the two ministries differ although they are similar in several aspects. First, in terms of educational stream, both ministries manage formal and non-formal education. Playgroups and language courses are two examples of non-formal education managed by MoEC, while MoRA takes responsibility for religious-based non-formal education such as the pesantren (religious boarding schools) and Raudatul Athfal (playgroup and kindergarten). In formal education, both ministries manage public and private schools. Although these schools have different names, the schooling level is the same, as illustrated in the table below:

### Table 1.1 Schooling systems in Indonesia

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Year (school grade)</th>
<th>MoEC</th>
<th>MoRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-6</td>
<td>TK (Taman Kanak-kanak/ kindergarten)</td>
<td>RA (Raudatul Athfal/ Kindergarten)</td>
<td></td>
</tr>
<tr>
<td>7-12</td>
<td>1-6</td>
<td>SD (Sekolah Dasar/ Primary School)</td>
<td>MI (Madrasah Ibtidaiyah/ Primary School)</td>
</tr>
<tr>
<td>13-15</td>
<td>7-9</td>
<td>SMP (Sekolah Menengah Pertama/ Junior Secondary School)</td>
<td>MTs (Madrasah Tsanawiyah/ Junior secondary school)</td>
</tr>
<tr>
<td>16-18</td>
<td>10-12</td>
<td>SMA (Sekolah Menengah Atas/ General Senior secondary school)</td>
<td>MA (Madrasah Aliyah/ General Senior Secondary School)</td>
</tr>
</tbody>
</table>

Source: MoEC (2013)

In addition, these schools use the same curriculum, the national curriculum, which is issued by MoEC. Students under MoRA administration are required to study religious-based subjects besides the subjects outlined in the national curriculum; in other words, schools
managed by MoRA include more subjects in their curriculum than do MoEC schools. In terms of teacher recruitment, each ministry may recruit teachers based on individual needs. MoEC and MoRA have different teacher requirements, but the policy of teacher qualification standard applies to all teachers in both ministries.

The main language used throughout Indonesia is Bahasa Indonesia. It was at the Youth Pledge Congress, 28 October 1928, that the language was first mentioned as a national language, and it was then formalized by the 1945 Constitution Article 36. Bahasa Indonesia has two main roles: the first is as a national language that symbolizes nationhood and national identity; and the second is as a means of communication that unites ethnic diversity in Indonesia (Ricklefs, 2001). Bahasa Indonesia also plays the role of an official language which has the following functions:

- official state language
- official medium of instruction
- official language for communication at the national level in social and government affairs, and
- official language for development of culture and the use of science and technology (Hamied, 2015, p. 32)

Bahasa Indonesia is a compulsory subject which is taught to all students, from primary to tertiary level.

In addition to the local and national languages, Indonesia recognizes the use of a foreign language which is also included in the education system, that is, English. Candraningrum (2008) points out that several foreign languages such as Dutch, German, French, and English were part of the Indonesian education system, prior to independence in 1945. Unlike neighbouring countries which adopted the colonists’ language as their foreign or second language, Indonesia decided to use English as the first foreign language, and it has been mandated in the education system as a compulsory subject since 27 December 1949 (Yulia,
English was chosen for two obvious reasons: Dutch was the language of the colonists, whilst English was becoming increasingly prevalent in the region and world-wide (Lauder, 2008; B. D. Smith, 1991).

Today, English is still the first foreign language taught in Indonesia. It is included in the Indonesian education system as a compulsory subject for students in lower secondary and higher secondary schools due to its important role in international communication, information and technology advancement, and education (Dardjowidjojo, 2001; Hamied, 2015; Lauder, 2008; Renandya, 2000). According to MoEC Decree Number 21 Year 2016 (Peraturan Menteri Pendidikan dan Kebudayaan Nomor 21, 2016), ELT in secondary schools in Indonesia is undertaken with the following aims: to develop communicative competence in oral and written form; to raise students’ awareness of the crucial role of English in learning; and to develop cross-cultural understanding. English is also taught to primary school students at the fourth-grade level in some provinces. In higher education, English is taught in a variety of ways, depending on university policy. Universities have the autonomy to develop and manage their own curriculum with the government only setting the minimum compulsory content.

Generally, there are three parties involved in the process of assessment in the Indonesian education system: the teacher, the school, and the government. Since English is one of the compulsory subjects in secondary high school, its assessment practice is ruled by the policy on curriculum and is enacted by law or ministry decree. With regard to formative assessment, the notion was documented in the policy on teacher qualification, and enacted through MoEC Decree number 16 Year 2007 (Peraturan Menteri Pendidikan Nasional Nomor 16, 2007). The policy stipulates that there are two assessment competencies that teachers need to demonstrate: the ability to implement and evaluate the process of learning outcomes,
and the ability to make teaching adjustments based on the assessment result. Furthermore, the formative assessment concept was implemented in the Kurikulum Berbasis Kompetensi (classroom-based competency, hereafter, KBK) 2004, and strengthened by Kurikulum 2013 which explicitly stated students’ active involvement in the assessment process.

Nevertheless, teachers directed criticism at Kurikulum 2013, because they found the assessment systems to be complicated and burdensome (Astuti, 2015; Darsih, 2015; Khasanah, 2015). This may be due to teachers’ conceptions which are still dominated by summative assessment (Azis, 2015); hence, these teachers considered formative assessment to be challenging to implement. As Hamied (2015) points out, language assessment is one of the burning issues in ELT in Indonesia; many teachers still consider assessment to be complicated and difficult to implement. Although it is important to note that the criticism was directed at the overall curriculum, a particular target was the assessment policy. Due to the sizeable protests (Fajerial & Tarigan, 2014; Ferdinandus, 2014; Gunawan & Suhari, 2014), the government decided to postpone the implementation of the curriculum, and only pilot schools were expected to implement Kurikulum 2013 (MoEC Letter No.179342/MPK/KR/2014 on 5 December 2014). The curriculum was then re-issued in 2016 with significant modifications made to the assessment policy, for example several terms such as authentic assessment, affective assessment, or self-reflection were no longer used. From 2017, all schools in Indonesia were obliged to apply Kurikulum 2013.

Teachers’ criticisms about the assessment practice as stipulated in Kurikulum 2013, may also imply that they had encountered challenges in implementing formative assessment. These may derive from two factors: the national high-stakes tests and teachers’ limited knowledge as to how formative assessment should be carried out. For decades, the administration of high-stakes-testing has hampered the implementation of formative assessment. A number of
studies document that the Ujian Nasional (National Examination, henceforth UN) has created various negative backwash effects such as anxiety, cheating, narrowing of the curriculum, and placing an excessive burden on students (Furaidah, Saukah, & Widiati, 2015; Hamied, 2015; OECD, 2015; Sukyadi & Mardiani, 2011; Zuhdi, 2015; Zulfikar, 2009). Another substantial negative effect of the UN is what Popham (2001) referred to as ‘teaching to the test’. This may have weakened the teachers’ competency to perform formative assessment because it demands teachers apply suitable assessment instruments to measure students’ ability, yet teachers continue to rely on tests due to their responsibility to prepare students to pass the UN. It is important to note that the government has changed its policy about the UN and it is no longer the determining factor for students’ graduation. Unfortunately, high-stakes testing is still taken into consideration when students are progressing to the next level (Peraturan Menteri Pendidikan nomor 23, 2016).

There have been few studies that focus on formative assessment in the field of English language teaching (hereafter, ELT). Language assessment has undergone a considerable change from standardised tests to an assessment culture (Brindley, 2008; Stoynoff, 2012). Several factors are considered to have contributed to this change: the shifting sands of learning theory; the expanded use of English internationally; and the emergence of communicative language teaching (hereafter, CLT) (H. D. Brown & Abeywickrama, 2010; Renandya, Lee, Kai Wah, & Jacobs, 1999; Richards, 2017; Spolsky, 1978). ELT reform may have ended the hegemony of standardised testing; however, the testing culture in ELT persists, although its focus has shifted from discrete-point testing of grammar and vocabulary to communicative language testing. Another phenomenon of language assessment reform is the emergence of various assessment types: alternative assessment, performance-based assessment, authentic assessment, classroom-based assessment, teacher-based assessment, and dynamic assessment.
Seemingly, formative assessment has not yet become a clear focus in ELT. This can be seen from the scant number of studies in this area. In fact, there are two categories of study: studies which implicitly relate to formative assessment (Rea-Dickins, 2001, 2006; Rea-Dickins & Gardner, 2000) and studies that specifically relate to formative assessment. Studies in the second category have mostly focused on the effectiveness of language assessment and the skill of writing and have involved the researcher’s intervention, which may not be feasible if applied in other contexts. Although such studies have given insights into formative assessment practice in ELT, the nature of how teachers implement formative assessment in actual classrooms remains unclear. This information is crucial because it can support and encourage teachers to practice formative assessment in their classrooms, and ultimately reap the benefits of it.

Teachers need support to conduct successful formative assessment; as Leung and Scott (2009) have said, studies on implementing the change are just as important as studies providing evidence of the benefits of the new assessment approach. Similarly, Edelenbos and Kubanek-German (2004, p. 281) stated that ‘User-friendliness and acceptance by teachers should be prerequisites for any new assessment schemes or tests if assessment is to be truly formative in language teaching and learning’. Therefore, it is vital to provide examples or models of good practice in formative assessment that derive from teachers’ own practice, rather than researcher interventions. This would inspire teachers to learn from their peers, which has been shown to be an effective means of professional learning (Cole, 2012; Eraut, 2004; Meirink, Meijer, & Verloop, 2007; Meirink, Meijer, Verloop, & Bergen, 2009).

1.2 Statement of the research problem
As will be made clear in Chapter 2, the potential of formative assessment to improve learning has been extensively studied in many disciplines (for example, Ash & Levitt, 2003; Black & Wiliam, 1998a, 1998b; Buck & Trauth-Nare, 2009; Cauley & McMillan, 2010; Chappuis &
Chappuis, 2008; Curry et al., 2015; Marzano et al., 2001). In addition, studies of formative assessment strategies have also supported the central role of formative assessment in enhancing learning (for example, Heritage & Heritage, 2013; Ozuem & Lancaster, 2015; Sadler & Good, 2006; Yang, Badger, & Yu, 2006). So far, however, there are only a few examples of how formative assessment can be conducted in the classroom; for instance, Wyle (2008) categorises classroom activities as formative assessment along with its counterexamples. Other examples of how formative assessment practice can be carried out are documented in studies of teacher professional development programs that focus on formative assessment (for example, Andersson & Palm, 2017; Black et al., 2003; DiBiase, 2014; McGatha, Bush, & Rakes, 2009; Randel, Apthorp, Beesley, Clark, & Wang, 2016; Thum, Tarasawa, Hegedus, Yun, & Bowe, 2015). Although these studies have shed light on how formative assessment can be carried out, they do not take into account the pre-existing methods of assessment used by teachers in their own context; this may be counterproductive for the improvement of formative assessment application as teachers are reluctant to change their practice.

In addition, most studies of formative assessment have been conducted in the areas of mathematics and science teaching (for example, Brookhart & Durkin, 2003; Deluca, Lapointe-mcewan, & Luhanga, 2016; Opre, 2015; Ruiz-Primo & Furtak, 2004; Taras, 2010). Further, several studies that do use formative assessment as their main focus address only one aspect of the topic such as the effectiveness of formative assessment, teachers’ understanding of formative assessment or formative assessment in writing skills. Although these studies shed light on how formative assessment can be used in ELT, they are not comprehensive. Firstly, they focus only on one macro skill; however, it is generally accepted that there are four skills integrated into modern language teaching. Secondly, they concentrate on researcher intervention, while the teachers’ actual practice remains unknown. As Popham
(2010) noted, teachers need a ‘starter kit’ applicable to their own setting and which is not merely an abstract concept. To encourage teachers to embrace the notion of formative assessment, it is important that they see how their daily practice can be permeated with formative assessment.

In the Indonesian context, formative assessment does not appear as a prominent field of research. This is unfortunate, considering that a large number of formative assessment notions were recommended more than a decade ago, as documented in MoEC Decree Number 20 Year 2007 and MoEC decree number 23 Year 2016. A number of Indonesian studies related to formative assessment have been conducted in the fields of chemistry (Nahadi, Firman, & Farina, 2015), mathematics (Purnomo, 2013), biology (Listiani, 2014), and physics (Edyanto, 2014). In contrast, studies focusing on formative assessment in the ELT area are relatively scant. To date, there are three studies that concern formative assessment (Arrafii & Suhaili, 2015; Arrafii & Sumarni, 2018; Widiastuti & Saukah, 2017), none of which provide examples of how formative assessment can be used to inspire other teachers, in the same context, to follow. For example, Widiastuti and Saukah (2017) mainly concerns teachers’ difficulties in conducting formative assessment. A study conducted by Arrafii and Suhaili (2015) focused on evaluating teachers’ formative assessment literacy and practice. More recently, a study focusing on teachers’ understanding of formative assessment in ELT in Indonesia was conducted by Arrafii and Sumarni (2018). These studies have provided valuable information necessary for English teachers in Indonesia to improve their practice of formative assessment. However, they have not provided examples of how formative assessment can be applied in an authentic classroom setting, to help teachers to improve their practice.
In light of the great value of formative assessment in enhancing learning, it is crucial to support teachers with examples of how formative assessment can be carried out by other teachers in similar contexts. This type of research is crucial because it can inspire other teachers to follow. If the practice is conducted by peer-teachers, teachers are more likely to believe that it will be viable in their context also; thus, relieving the image that formative assessment is complicated and burdensome.

1.2.1 Research aims and research questions

This prospective study was designed to provide examples of how formative assessment can be carried out by English as a foreign language (hereafter, EFL) teachers in Indonesia. To achieve this purpose, the study explores teachers’ understanding of formative assessment, and their practice in implementing formative strategies. In addition, this study also investigates challenges encountered by teachers in applying formative assessment, and teachers’ solutions for overcoming these challenges. The specific objective of this study was to seek answers to the following research questions:

1. How do EFL senior high school teachers in Indonesia perceive formative assessment?
   a. What do they understand by the term formative assessment?
   b. What do they believe are its advantages?
   c. Ideally, how do they think formative assessment should be carried out?

2. How are teachers’ perceptions of formative assessment manifested in their practice of it? What instruments and strategies do teachers use to implement formative assessment?

3. What challenges are encountered by teachers when implementing formative assessment? How do they cope with these challenges?
1.2.2 Significance of the study

There are several important areas where this study makes an original contribution to the knowledge base of formative assessment. Firstly, it fills a literature gap as formative assessment studies in ELT are relatively scant and underexplored. Conducting an in depth study on formative assessment in ELT will enhance our understanding of formative assessment, and this can also be adapted and applied in other language learning contexts, or other learning fields in general. In addition, the findings from the Indonesian context may be applicable to other countries, especially in south-east Asia, that have encountered similar challenges. These findings include teacher resistance and reticence towards implementing a policy of formative assessment.

Providing examples from actual classroom practice can make a significant contribution to teachers’ knowledge, as formative assessment is not confined to particular assessment instruments or techniques, and teachers are only equipped with general formative assessment strategies. Examples and analysis will have the potential to inspire teachers to adopt or adapt the practice in their own classrooms; particularly, if the examples originate from their peer-teachers in a similar context and not merely from researchers’ idealised interventions. In addition, supporting teachers with practical examples of how formative assessment can be carried out would minimize misinterpretation of formative assessment practice.

Furthermore, this study provides an important opportunity to advance the understanding of teachers’ formative assessment practice. Teachers’ understandings of formative assessment are also explored in this study and will serve as the basis for teachers’ professional development programs and for courses on assessment in teacher education programs.
1.3 Research methodology

The study aims to give examples of how formative assessment can be carried out in EFL settings in Indonesia. In relation to the purpose of the research, to identify teachers’ formative assessment practices and the beliefs that underlie their practice, the thesis uses multiple case studies in order to compare and contrast results (Creswell, 2005; Merriam, 1998; R. K. Yin, 2003). This study involved eight EFL teachers: four of whom are teaching under MoEC and four under MoRA. These teachers were selected based on the following criteria: they represented the diverse contexts present in Indonesia, especially in West Java; they had recognised teaching competency and experience in implementing formative assessment; and they were recommended by peers.

Three sources of data were collected for this study: interviews, observations, and formative assessment artefacts. The semi-structured interviews were conducted twice; the first interview aimed to identify teachers’ understanding of formative assessment, and it was conducted prior to classroom observation. The second interview was conducted post-classroom observation; the purpose of which was to clarify some unclear information from the first interview and to confirm and explore teachers’ actions during the classroom observation. Formative assessment observation instruments developed by Oswalt (2013) were adapted and used during the observations. Each participant was observed in the period of formative assessment medium cycle (Thompson & Wiliam, 2008). Observations were documented by audio-recording and the taking of field notes. Assessment instruments that teachers applied during classroom observations were used as another source of data and are referred to as formative assessment artefacts. All sources of data were collected after approval from the University of Canberra Human Research Ethics Committee was secured (Appendix 1). The data analysis methodology for this study was adapted from Creswell (2005).
1.4 Structure of the thesis

This thesis consists of six chapters: Chapter 2 reviews the extant literature and research that underpins the research questions formulated for this thesis. There are three main issues in this chapter: formative assessment; formative assessment in ELT; formative assessment in EFL study in Indonesia; and challenges and solutions in applying formative assessment.

Controversies and major findings from empirical studies of formative assessment are also reviewed here. Gaps in the literature are identified for further investigation.

Chapter 3 outlines the methodological approach adopted in the study. Relevant to the nature of the research purpose, a multiple case studies was adopted for this study. I provide justification for the approach and participant selection. The major research instruments such as semi-structured interviews, classroom observations, and formative assessment artefacts are identified and the procedures followed in collecting and analysing data are described.

Key findings from an analysis of the research data are presented in Chapter 4. These include results based on semi-structured interviews, classroom observations, and formative assessment artefacts; as well as the difficulties encountered by teachers, and solutions that teachers applied to overcome their difficulties.

In Chapter 5, a detailed account is given and the findings of the study discussed in relation to previous research.

The last chapter, Chapter 6, summarizes the study findings, and focuses on both the pedagogical and practical implications of the results for teachers, curriculum designers, and policy makers. The chapter ends with limitations of the study and recommendations for future research.
Chapter 2
Literature Review

The notion of formative assessment came to prominence when Black and Wiliam (1998a, 1998b) published their seminal work and claimed that the potential role of formative assessment is to improve learning. Nevertheless, the notion of formative assessment has not gone uncontested; concerns are frequently raised about the definition of formative assessment and what constitutes its practice. Numerous definitions have been proposed, yet they have not been agreed to by all scholars. Formative assessment can be complicated to define because the term embodies a multitude of teaching concepts. Nonetheless, a great deal of literature has emphasised that formative assessment refers to assessment which aims to improve learning; in addition, a large number of formative assessment definitions include the following elements:

- **Process**
  Formative assessment is not a one-off activity; rather it is a continuing process.

- **Feedback**
  Teachers’ assessment should involve formative feedback which is evidence-based so that follow-up action can be taken.

- **Students’ involvement**
  Assessment is not solely the teachers’ responsibility. The students themselves and peers should actively participate in the process.

- **Teaching adjustment**
  Evidence gathered from assessment should be used by teachers to adjust their teaching; in fact, this is the core element that differentiates formative assessment from summative assessment.

Therefore, in this study I define formative assessment as a continuous activity of assessing students’ developing competencies that includes feedback and students’ involvement, and
subsequent adjustment to teaching. Furthermore, formative assessment does not relate to a specific set of assessment instruments; however, teachers can incorporate formative assessment practice based on strategies such as: clarifying, sharing, and understanding learning intentions and success criteria; engineering effective classroom discussion, questions, and learning tasks; providing effective feedback that moves learning forward; activating students as instructional resources for one another; and activating students as owners of their own learning (Wiliam & Leahy, 2015).

Nevertheless, translating formative assessment strategies to the classroom setting can be challenging for teachers, and teachers need to be supported with examples of how these strategies can be implemented in their own teaching contexts; this includes how to overcome the prevailing challenges during formative assessment implementation. To date, most of the research on formative assessment has been limited to giving recommendations and experimenting on how formative assessment strategies can be implemented. Moreover, studies in formative assessment in ELT are relatively scarce and underexplored. In Indonesia specifically, to date, only three studies have been published on formative assessment in ELT, but none of these studies focused on how formative assessment in ELT can be carried out. This chapter reviews the theories underpinning formative assessment. It begins by outlining assessment in general so as to lay a firm foundation for the shifting paradigm from testing culture to assessment culture. The following section will discuss the notion of formative assessment, and formative assessment strategies. In addition, challenges of and solutions for implementing formative assessment are also discussed in this chapter. I discuss the application of formative assessment to ELT, and formative assessment in Indonesia. This chapter ends with a summary of the main points discussed throughout the chapter and highlights the research gap that exists in the literature.
2.1. Outlining assessment

Assessment can broadly be defined as a systematic process of collecting evidence about students’ competence and performance using various methods and techniques (Butt, 2010; Gipps, 2011; Katz, 2014; Oosterhoof, 2009; Shepard, 2000; Stobart, 2008). Assessment has often mistakenly been regarded as being interchangeable with tests (H. D. Brown, & Abeywickrama, 2010; Stobart, 2008). A large number of studies has tried to clarify this issue, for example H. D. Brown and Abeywickrama (2010) illustrate differences between assessment and tests in the figure below:

![Figure 2.1 Test, assessment, and teaching and learning](image)

Adapted from H. D. Brown and Abeywickrama (2010, p. 5)

As illustrated by Figure 2.1, H. D. Brown and Abeywickrama (2010, p. 4) defined tests as ‘… prepared administrative procedures that occur at identifiable times in a curriculum when learners master all their faculties to offer peak performance, knowing that their responses are being measured and evaluated’. They elaborate by describing three elements of tests: firstly,
a test may use multiple-choice, fill-in the blanks, writing prompts, or oral interviews; secondly, a test should give a specific result in the form of a measurement such as a score, praise, or pass/fail; finally, tests aim to measure an individual’s ability, knowledge, or performance. Similarly, two key elements of assessment are systematic processes and tools. However, unlike testing, assessment is not a one-off activity that ends with a result; instead, it involves an integrated process that may include the whole cycle of assessment such as designing, implementing, and interpreting as well as involving an ongoing sequence of assessment activities. In the same vein, Stobart (2008) explains that assessment may include an array of methods or techniques used to attain evidence of students’ learning, one of which may be tests. In other words, tests can be considered to be a subset of assessment, while other methods or techniques of assessment may include portfolio, learning logs, dialogue journals, self-assessment, or peer assessment, for example.

Presumably, the ‘shifting sands’ of the dominant learning paradigm have influenced how assessment and tests are perceived (H. D Brown & Abeywickrama, 2010; Davison & Leung, 2009; Gipps, 2011). Gipps (2011, p. 2) who defined a paradigm as ‘a set of interrelated concepts which provide the framework within which we see and understand a particular problem or activity’, contends that assessment has undertaken a paradigm shift from testing culture to assessment culture. Testing culture is often tightly related to high-stakes standardised tests because it is the need to pass the tests which has shaped the culture. Specifically, Birenbaum (1996) characterized testing culture as a learning condition with the following features: there is a distinction between instruction and assessment; students have a passive role in assessment; the tests are decontextualized and discrete, and commonly employ a multiple choice format. On the other hand, assessment culture is characterized by: students’ active participation; multiple forms of assessment; and students’ performance is determined as a profile and not as a single score (Birenbaum, 1996). In the same vein, Hamp-Lyons
(2007, p. 488) described a testing culture as when ‘classroom assessment is seen as simply preparation for an externally set and assessed exam’; and assessment culture in the classroom is seen as ‘considerations of learning and teaching’ (Hamp-Lyons, 2007, p. 488). In a recent study, Birenbaum (2016) provides more comprehensive information about the distinction between testing culture and assessment culture. In her multiple case studies involving eight primary and middle schools in Israel, Birenbaum (2016) contrasts two school cultures, testing and assessment. They are based on several factors: mindset, classroom assessment, teachers’ professional development, leadership, and impacts of external assessment. Drawing on data from semi-structured interviews and focus group discussion, the findings from this study are summarised in the table below:

Table 2.1 Testing culture versus assessment culture

<table>
<thead>
<tr>
<th>Category</th>
<th>Testing culture</th>
<th>Assessment culture</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Mindset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The purpose of assessment</td>
<td>Making grades for reports</td>
<td>Aiming for learning</td>
</tr>
<tr>
<td>The function of assessment</td>
<td>Accountability</td>
<td>Direction for further learning</td>
</tr>
<tr>
<td>The methods of assessment</td>
<td>Prefers standardised tests</td>
<td>Focuses on establishing dialogue (interaction) with learner</td>
</tr>
<tr>
<td>Power relation in assessment</td>
<td>Controlled by assessor</td>
<td>Shared assessment power</td>
</tr>
<tr>
<td>Attitudes towards diversity</td>
<td>One instrument will suit all students</td>
<td>Acknowledges students’ diversity</td>
</tr>
<tr>
<td>Expectation about learning</td>
<td>Teachers believe that students’ capacities are fixed</td>
<td>Teachers believe that students have distinct capacity which can be used to move learning forward</td>
</tr>
<tr>
<td>The fidelity of assessment</td>
<td>Tests can measure students’ ability accurately</td>
<td>Tests may not depict students’ overall ability</td>
</tr>
<tr>
<td><strong>2. Classroom assessment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formal assessment</td>
<td>Assessment for learning is interpreted as frequent testing, and formative assessment strategies are applied superficially</td>
<td>Emphasises and applies formative assessment</td>
</tr>
<tr>
<td>Classroom culture</td>
<td>Competitive and score-oriented</td>
<td>Applies the notions of constructivist learning theory such as collaborative learning, higher-order thinking, student agency</td>
</tr>
</tbody>
</table>
Using the development of learning theories as a framework, James (2006) divides the development into three generations and explains how the changes affect assessment practice. In the first generation of learning theories, assessment covers what has been taught. This generation includes ‘folk’ views of learning and behaviourist views of learning. The assumption in folk views of learning is evidence- and theory- based; learning is considered to have occurred when knowledge has been retained. Consequently, the teachers’ role is to be the source of knowledge; they transfer knowledge to students who are required to memorize what is being taught. Besides the folk views, this generation also includes the behaviourist view. Although behaviourism theory is derived from research-based theory, this view of learning may be similar to folk views of learning. Behaviourism suggests that learning should be sequential and hierarchical (Gipps, 2011); thus, breaking down learning objectives is one of the key features of behaviourism. It is believed that complex understanding will be attained as the elements of the required body of knowledge are mastered. Learning is considered to have occurred when learners can recall learning materials taught by teachers; therefore, tests hold a prominent role in assessment in order to ensure learners’ step by step mastery of the prerequisite knowledge (Shepard, 2000).
The second generation of assessment practice is assessing learning as individual sense-making (James, 2006). This generation of assessment is based on cognitive-constructivist views of learning which consider learning to be a process of acquiring knowledge and skills; however, learners have an active role in the process, connecting and interpreting what they have learned. The teacher must discover students’ prior learning in order to scaffold students’ knowledge. Students are considered to be novices and the teacher’s role is to engage learners and facilitate the process from novice to ‘expert’. The aim of learning is for students to construct an understanding of particular knowledge. Learning is considered to have occurred when students are able to apply the knowledge they have learned in other relevant contexts. One of the significant differences between second generation and first generation views of learning is the reliance on tests; cognitive-constructivism does not rely on tests as much as behaviourism does. It places more emphasis on the students’ depth of knowledge and understanding, and their ability to solve problems.

The third generation of assessment practice views learning as building knowledge while doing things with others; this generation underpins the sociocultural view of learning and is also known as situated learning (James, 2006). The significant difference between sociocultural and other views of learning is that learning is considered to be part of social practice. Learning is enhanced by what Vygotsky referred to as ‘joint productive activity’ in a social setting. For example, in a classroom setting this is translated into learners and teachers who collaborate in achieving learning objectives; thus, learners’ active participation is one of the crucial elements in the sociocultural view of learning. Students learn from the social groups that they belong to by sharing a common language and understanding. Knowledge is understood to be socially constructed, hence learners are not only learning from teachers but also from their peers. Other concepts from this learning theory are learners’ autonomy, move from lock-step learning, and high challenge, high support. Socio-
Constructivism has influenced assessment in several ways: feedback is important because it facilitates the desired interaction; assessment can include self, peer, and teacher assessment; forms of assessment should replicate the authentic contextual problems encountered in learners’ surroundings (James, 2006). Similarly, Shepard (2000) elaborated on some characteristics of classroom assessment from a sociocultural constructivist perspective:

- Challenging tasks to elicit higher order thinking
- Addressing learning processes as well as learning outcomes
- An ongoing process, integrated with instruction
- Used formatively in support of student learning
- Expectations visible to students
- Students active in evaluating their own work
- Used to evaluate teaching as well as student learning (Shepard, 2000, p. 8)

In retrospect, this section reviews the general concept of assessment and how it differs from tests. Taken together, these studies support the notion that the shifting paradigm of learning theories has had a substantial impact on assessment practice: from testing culture to assessment culture. Today, the crucial role of assessment is widely recognized; it serves to inform teachers, students, and parents about learning; directing curriculum changes; measuring accountability; and acting as a form of certification (Black & Wiliam, 2009; Gipps, 2011).

2.2 Understanding formative assessment

In the academic literature, assessment is divided into two streams, formative and summative. Much of the literature maintains that both assessment streams are equally important and have the potential for improving learning (for example, Black, 2012; Broadbent, Panadero, & Boud, 2017; Houston & Thompson, 2017; Lau, 2016; Wei, 2015). Although formative and summative assessments are equally important, it is recommended that teachers focus on formative assessment because it provides valuable information about students’ learning process on a daily basis.
It is generally accepted that the term formative assessment, first introduced as ‘formative evaluation’, was established by Michael Scriven in 1967 (Black & Wiliam, 1998a, 1998b; Earl, 2013; Popham, 2006; Stiggins, 2005). The term was in common use when Bloom, Madaus and Hastings (1981) defined it as testing for learning while summative assessment was defined as testing for final performance (Guskey, 2003). The concept of formative assessment revived when Black and Wiliam published their seminal work in 1998. They based it on 250 publications about classroom assessment dating from 1987 to 1997, and they concluded that ‘There is a body of firm evidence that formative assessment is an essential component of classroom work and that its development can raise standards of achievement.’ (Black & Wiliam, 1998b, p. 12). In addition, the King’s Medway-Oxfordshire Formative Assessment Project (hereafter, KMOFAP) has also contributed to the prominence of the formative assessment concept established by Wiliam, Lee, Harrison & Black (2004). The project, which involved 24 teachers and was conducted over a two-year period, aimed to disseminate the notion of formative assessment which focuses on the following strategies: questioning, comment only without marking, peer and self-assessment, and formative use of summative tests (Wiliam, Lee, Harrison & Black, 2004). These researchers claimed that ‘the results presented here provide firm evidence that improving formative assessment does produce tangible benefits in terms of externally mandated assessment’ (Wiliam, Lee, Harrison, & Black, 2004, p. 63).

Subsequently, over the past decades, most research in formative assessment has emphasised the potential benefits of formative assessment in improving learning. This large and growing body of literature highlighted a number of benefits such as motivating students, informing them about learning areas that need improving, and enhancing their achievement in an external accountability test (Andersson & Palm, 2017; Black, Harrison, Lee, Marshall, &
Wiliam, 2003; Cauley & McMillan, 2010; Marzano, Pickering, & Pollock, 2001). Cauley and McMillan (2010, p. 2) argue that there are four reasons why formative assessment can improve students’ learning:

1. Frequent, ongoing assessment allows for fine-tuning of instruction and student focus on progress.
2. Immediate assessment helps ensure meaningful feedback.
3. Specific, rather than global, assessments allow students to see concretely how they can improve.

Furthermore, Ash and Levitt (2003) mention that formative assessment can also benefit teachers. In their case studies, Ash and Levitt (2003) found that formative assessment enables teachers to improve alongside students because it enforces reflection on their teaching style. They concluded their study below:

We have provided evidence in the form of two case studies that formative assessment provides a vehicle for growth for not only the learner, but for the teacher. In each of the projects teachers’ use of formative assessment propelled their professional transformation in tandem with the development of the learners as they self-monitored their practices through the lens of assessing the understanding of the learner (Ash & Levitt, 2003, p. 45)

Despite these potential benefits to both the teacher and the student, the notion of formative assessment has not gone uncontested. One criticism is that a definition of what constitutes formative assessment has not been agreed upon (Bennett, 2011; Dunn & Mulvenon, 2009; Kingston & Nash, 2011). As Bennett (2011, p. 5) points out, ‘the term “formative assessment,” does not yet represent a well-defined set of artefacts or practices’. In their ground-breaking work Black and Wiliam (1998b, p. 2) defined formative assessment as:

… all those activities undertaken by teachers and —— by their students in assessing themselves—that provide information to be used as feedback to modify teaching and learning activities. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching to meet student needs.(Black & Wiliam, 1998b, p. 2)
This definition poses a problem because it is considered too broad and does not adequately inform practitioners about formative assessment; thus, to a certain extent, it has created ‘definitional fuzziness’ (Yorke, 2003). Numerous studies have attempted to clarify the term by substituting it under different labels; one of the most widely used is ‘Assessment for Learning’ (AfL). AfL has been used interchangeably with formative assessment in a large volume of published studies (DeLuca, Luu, Sun, & Klinger, 2012; Harlen & Winter, 2004; Harrison, 2005; Heritage, 2018; I. Lee, 2011b; Remesal, 2011a; Rosemartin, 2013; Taras, 2010b; E. White, 2009; Wiliam et al., 2004). Nevertheless, it is important to note that Stiggins (2005) and Chappuis and Chappuis (2008) argue that AfL is slightly different from formative assessment; when data gathered from formative assessment is truly used for learning, then it becomes AfL. Furthermore, other studies maintain that the term AfL has not contributed to the clarification of a definition of formative assessment because the purpose of assessment is to improve learning (Bennett, 2011; Newton, 2007; Taras, 2010b).

Several scholars agree on the importance of finding an acceptable definition of formative assessment. For example, Bennett (2011, p. 5) argues that ‘To realise maximum benefit from formative assessment, new development should focus on conceptualising well-specified approaches built around process and methodology rooted within specific content domains’ (Bennett, 2011, p.5). Similarly, Crossuard and Pryor (2012) argued that the practice of formative assessment required a sound theory to justify its practice.

One question that needs to be asked, however, is whether a consensus on the definition of formative assessment is worth finding. The need to formulate an agreed definition of formative assessment can be contested by the following rationales. First, the central proposition of formative assessment is that it enables teachers to help students’ learning on a daily basis (Black & Wiliam, 1998b). It is a delicate matter to determine what formative assessment is because it very much depends upon the context. Second, a serious weakness in
these claims is that they fail to address how a consensus of formative definition can advantage teachers. As McEntarffer (2012, p. 156) emphasises ‘simply sharing the definition and supporting research behind formative assessment may not help teachers to think and discuss formative assessment in a useful way’. Third, it is now well established from a variety of studies that formative assessment permeates teaching and creates an interpenetrated connection between teaching and assessment (Andrade & Valtcheva, 2009; Black, 2012; Hickey, 2015; Jin, 2009; Leahy, Lyon, Thompson, & Wiliam, 2005; Ruiz-Primo & Furtak, 2004; Thompson & Wiliam, 2008); thus, determining what formative assessment is can be complicated as it entails a wide array of teaching concepts. Furthermore, Rea-Dickins (2001, 2006) postulates that formative assessment opportunities occur during learning and can be seen as a continuum from formal to informal practice. According to Shavelson et al. (2008, p. 300), this continuum is influenced by several factors: ‘amount of planning involved, its formality, the nature and quality of the data sought, and the nature of feedback given to students by the teacher’. Ruiz-Primo and Furtak (2004) illustrate the representation of formal and informal formative assessment as follows:

![Graphical representation of formal and informal formative assessment](image)

**Figure 2.2 Graphical representation of formal and informal formative assessment**

Source: Ruiz-Primo and Furtak (2004, p.4)
Further, the formality of formative assessment practice can be divided into the following three categories: on-the-fly formative assessment; planned-for interaction; and formal and embedded in curriculum (Heritage, 2007; Shavelson et al., 2008). These categories are illustrated by Shavelson (2008, p.300) in Figure 2.3.

**Figure 2.3 Variation in formative assessment practices**
Source: Shavelson et al. (2008, p. 300)

Further, using a slightly different perspective, Cauley and McMillan (2010) proposed that formative assessment is a continuum which they refer to as ‘variations of formative assessment characteristics’ that can be categorised as low to high level formative assessment practice. The continuum is outlined as follows:
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Low-level formative</th>
<th>High-level formative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of the evidence</td>
<td>Mostly objective, standardized</td>
<td>Varied assessment, including objective, constructed response, and anecdotal</td>
</tr>
<tr>
<td>Structure</td>
<td>Mostly formal, planned, anticipated</td>
<td>Informal, spontaneous, at the moment</td>
</tr>
<tr>
<td>Participants</td>
<td>Teachers</td>
<td>Teachers and students</td>
</tr>
<tr>
<td>Feedback</td>
<td>Mostly delayed (e.g. give quiz and give students feedback the next day) and general</td>
<td>Mostly immediate and specific for low-achieving students, delayed for high achieving students</td>
</tr>
<tr>
<td>When conducted</td>
<td>Mostly after instruction and assessment (e.g. after a unit)</td>
<td>Mostly during instruction</td>
</tr>
<tr>
<td>Instructional adjustment</td>
<td>Mostly prescriptive, planned (e.g. pacing according to instructional plan)</td>
<td>Mostly flexible, unplanned</td>
</tr>
<tr>
<td>Choice of instructional tasks</td>
<td>Mostly teacher determined</td>
<td>Teacher and student determined</td>
</tr>
<tr>
<td>Nature of teacher-student interaction</td>
<td>Most interactions based primarily on formal roles</td>
<td>Extensive, informal, trusting, and honest interaction</td>
</tr>
<tr>
<td>Role of student self-evaluation</td>
<td>Little or none</td>
<td>Integral</td>
</tr>
<tr>
<td>Type of motivation</td>
<td>Mostly extrinsic (e.g. passing a competency test)</td>
<td>Mostly intrinsic</td>
</tr>
<tr>
<td>Attribution for success</td>
<td>External factor (teacher; luck)</td>
<td>Internal, unstable factor (e.g. moderate student effort)</td>
</tr>
</tbody>
</table>

**Figure 2.4 Variations of formative assessment characteristics**

Source: Cauley & McMillan (2010, p. 2)

Fourth, differing definitions of formative assessment may be due to seeing it from different perspectives; however, these definitions share similar central elements (Hamm & Adams, 2009; Hondrich, Hertel, Adl-Amini, & Klieme, 2015; Wylie et al., 2012). Wylie et al. (2012) point out that the term formative assessment is commonly related to ‘in the moment’ and ‘adjust, adapt, or use’; thus ‘collected evidence that is not then used is of no formative value’ (Wylie et al., 2012, p. 16-17).
Drawing on an extensive range of publications on the definitions of formative assessment (for example, Black & Wiliam, 1998a, 1998b; Colby-Kelly & Turner, 2007; Gattullo, 2000; Harlen & James, 1997; Heritage, 2010; Hondrich et al., 2015; Popham, 2006; Sheppard, 2005; Wiliam & Leahy, 2015; Yin & Buck, 2015), this thesis identifies four important elements of formative assessment, namely process, feedback, students’ involvement, and teaching adjustment as the core element which differentiates formative assessment from other notions of assessment (Black, 2015; Black & Wiliam, 1998a, 1998b, 2009; Chappius & Stiggins, 2002; OECD, 2005; Popham, 2006). As illustrated by Harlen and James (1997), a test which is conducted at the end of a lesson can be categorised as formative assessment if the data taken from the test is used to identify which parts of learning need further teaching. However, if teachers administer tests, and the data is used simply to report students’ progress to parents at the end of the semester, then the assessment cannot be categorised as formative assessment. In a similar vein, Black (2015) contends that the use of assessment results, which he referred to as ‘learning evidence’, determines whether an assessment practice is formative or not. Consequently, identifying formative assessment can be carried out by analysing first how the evidence of learning is used, and second by examining the presence of the elements of formative assessment as mentioned above.

Having defined what is meant by formative assessment, it is now necessary to identify what constitutes its practice. A number of studies have agreed that formative assessment is not limited to a particular practice (for example, Pinchock & Brandt, 2009; Stobart, 2008; Wiliam, 2011); indeed, Stobart (2008) has argued that formative assessment should be perceived as an approach rather than a prescriptive assessment technique. According to Pinchock and Brandt (2009, p. 2), ‘Any form of assessment from performance-based to multiple-choice items can be used in formative assessment practice’. This concept may be too broad and cause misinterpretation among some teachers. As reported by several studies,
Formative assessment is sometimes erroneously defined as frequent testing, a set of instruments, any continuous assessment, and any instruments other than tests (Davison & Leung, 2009; Pinchock & Brandt, 2009; Wiliam, 2011).

Nevertheless, a number of studies document several strategies that can help teachers implement formative assessment practice. For example, in the KMOFAP, Black et al. (2003) propose four strategies: oral feedback in classroom dialogue, feedback through marking, peer and self-assessment, and formative use of summative tests. While these strategies provide an insight into the practice of formative assessment, Black et al. (2003) fail to acknowledge one of the elements behind teachers’ practice, that is, the learning goals that direct learning.

Another formative assessment strategy, the FAST SCASS (Formative Assessment for Students and Teachers State Collaboration on Assessment and Students Standards), outlines five attributes of formative assessment:

1. **Learning Progressions:**
   Learning progressions should clearly articulate the sub-goals of the ultimate learning goal.

2. **Learning Goals and Criteria for Success:**
   Learning goals and criteria for success should be clearly identified and communicated to students.

3. **Descriptive Feedback:**
   Students should be provided with evidence-based feedback that is linked to the intended instructional outcomes and criteria for success.

4. **Self- and Peer-Assessment:**
   Both self- and peer assessment are important for providing students with an opportunity to think meta-cognitively about their learning.
5. **Collaboration:**

A classroom culture in which teachers and students are partners in learning should be established (McManus, 2008, p. 4-5).

Although these attributes have addressed the importance of learning goals, the study does not take into account the central role of teachers in designing assessment instruments which can be used to adjust learning.

In comparison, Wiliam and Leahy (2015) formulated five formative assessment strategies which are more comprehensive and well accepted because they include learning goals as the central point, as well as the teacher’s prominent role in designing assessment instruments. Specifically, these strategies are as follows:

1. Clarifying, sharing and understanding learning intentions and successful criteria;
2. Engineering effective classroom discussion, questions, and learning tasks;
3. Providing feedback that moves learners forward;
4. Activating students as the owners of their own learning;
5. Activating students as instructional resources for one another. (Wiliam & Leahy, 2015)

The present study applies formative assessment strategies as a framework based on the following rationales: they encompass the formative assessment attributes and formative assessment strategies applied successfully in KMOFAP; and they have been applied in previous studies (Leahy, Lyon, Thompson, & Wiliam, 2005; Thompson & Wiliam, 2008). In addition, Wiliam and Leahy (2015) specifically mention that the strategies are appropriate for secondary students, which is relevant to the current study.

In summary, this section has attempted to provide an overview of the literature relating to formative assessment. Based on a considerable amount of literature, this thesis formulated
formative assessment as an assessment that aims to improve learning, and it can be ‘recognized’ according to the presence of several key elements, and by whether teachers apply formative assessment strategies. A more detailed account of formative assessment strategies is given in the following section.

2.3 Implementing formative assessment

Previous research has established that formative assessment resides in the sociocultural constructivist learning theory (Ash & Levitt, 2003; Black & Wiliam, 2009; Heritage, 2010; Shepard, 2000, 2005; Trumbull & Lash, 2013). In fact, according to Shepard (2005), formative assessment and instructional scaffolding within the zone of proximal development (ZPD) are essentially the same thing. Vygotsky’s ZPD is the space between ‘the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers’ (Vygotsky, 1978, p. 90). These two terms are considered to be ‘in sync’ (Shepard, 2005) because both terms define strategies that teachers implement during instruction with the aim of moving learning forward in the ZPD. Consequently, many formative assessment strategies reflect sociocultural learning theory, such as feedback that moves learning forward and students’ active involvement.

This section will thoroughly discuss formative assessment strategies, as outlined by Wiliam and Leahy (2015), they are: clarifying, sharing, and understanding learning intentions and success criteria; engineering effective classroom discussion, questions, and learning tasks; providing effective feedback that moves learning forward; activating students as instructional resources for one another; and activating students as owners of their own learning.
2.3.1 Clarifying, sharing, and understanding learning intentions and successful criteria

The first strategy relates to learning goals, variously called learning targets, benchmarks, content standards, goals and outcomes (Stiggins et al., 2004). Simply put, this strategy emphasises mapping learning goals by communicating them to students in order to give clear direction to learning objectives. The literature of formative assessment has highlighted the crucial role of implementing this strategy (Black & Wiliam, 1998a; Entwistle & Smith, 2002; Leahy et al., 2005; McManus, 2008; Pinchock & Brandt, 2009); further, Black and Wiliam (1998a) highlight the fact that students’ lack of understanding about what is expected from their learning is one of their causes of low achievement. Therefore, the aim of this strategy is to build a mutual understanding between teacher and students about what is successful learning. Further, Sadler (1989) and James (1997) argue that a mutual understanding of the expected learning goals would improve the effectiveness of feedback; however, students who are not grasping the concepts may not improve much if they are unaware of what is expected from them, despite a large amount of good feedback provided by the teacher. In addition, Wiliam et al. (2004) point out that communicating learning goals encourages autonomous learning provided that students are aware of their strengths and weaknesses, and they work toward the learning objective.

In a classroom setting, this strategy can be interpreted to be providing assessment criteria and/or rubrics for learning activity (Chappius & Stiggins, 2002; Greenstein, 2010; L. Harris, 2007; Wiliam & Leahy, 2015). Rubrics are a set of rules or guidelines that are used to align standards with the lesson and to determine students’ level of competence (Brookhart, 2013; Burke, 2006; O’Malley & Valdez Pierce, 1996). Several lines of evidence suggest that assessment criteria and/or rubrics affect learning positively. For example, McDonald and Boud (2003) reported that standards which are translated as assessment criteria for students
can improve students’ achievement. This quantitative study was conducted in ten senior high schools in Barbados. The experimental groups were formally trained in self-assessment skills, with emphasis on criteria and standards for three academic years. The findings showed that significant differences occurred between the experimental and control groups, with the experimental group benefiting more in the external examinations. Another study conducted by Panedro and Johnsson (2013) revealed similar findings: in their meta-analysis study, they examined the impacts of applying rubrics for formative assessment and confirmed that rubrics affect learning positively; however, several factors such as educational level, time, and gender may moderate its impacts.

Besides establishing assessment criteria and/or rubrics, another practice that can clarify learning intentions is giving examples of successful learning. This will make learning intentions more visible (Black et al., 2003; Tuttle, 2008; Wiliam & Leahy, 2015). In fact, L. Harris (2007) suggests that examples should be given prior to explaining assessment criteria or the rubric itself. Specifically, L. Harris (2007) explains that a lesson should begin with an overview followed by the introduction of learning objectives which should be discussed and clarified; afterwards, teachers should give examples. Once students begin with the assigned task, the assessment criteria should be explained. Similarly, Chappius and Stiggins (2002, p. 41) have elaborated several strategies that can be used to engage students with the learning targets: (1) determine the attributes of good performance; (2) use scoring guides to evaluate real work examples; (3) revise anonymous work samples; (4) create practice tests or test items based on their understanding of the learning targets and the essential concepts in the class materials; (5) communicate with others about their growth and determine when they are nearing success.
To sum up, this section has provided a brief summary of one of the strategies used to implement formative assessment; that is achieving a mutual understanding, between teacher and students, of successful learning. The establishment of assessment criteria and/or rubrics and the provision of examples are considered two of the best examples of how this strategy can be implemented. However, taken together, these studies have not yet provided an insight into how teachers actually implement the strategy in a real classroom setting, and how teachers make use of this strategy to improve learning.

2.3.2 Engineering effective classroom discussion, questions, and learning tasks

The second strategy concerns selecting a suitable assessment instrument in order to acquire evidence of learning which will be used later to enhance learning. Implementing this strategy can be challenging for some teachers due to a wide array of assessment instruments which can be applied in classrooms. As mentioned in Section 2.1, formative assessment is not confined to a set of practices, any assessment instrument may have the potential to be used formatively; consequently, teachers need to select and incorporate assessment instruments in line with their teaching and learning context (Black & Wiliam, 2009; Heritage, 2010; Pinchock & Brandt, 2009).

Interestingly, most researchers investigating formative assessment have utilised questioning as one of the instruments (for example, Black & Wiliam, 1998a, 1998b; Chappuis & Chappuis, 2008; Heritage, 2010; Stiggins, Arter, Chappuis, & Chappuis, 2006; Wong, 2010); hence, a more detailed account of questioning is given in the following subsection.

2.3.2.1 Questioning

It is now well established from a variety of studies that questioning suits the concept of formative assessment because it has the flexibility to be implemented at any stage of
teaching. As Al-Darwish (2012) points out, teachers believe questioning to be an integral part of teaching, and not a teaching method or assessment instrument in itself. This is in agreement with formative assessment which can be tailored seamlessly into teachers’ teaching practice (Greenstein, 2010; Hickey, 2015; Poehner & Lantolf, 2005). Questioning provides immediate evidence of learning; this information is crucial because it gives teachers the opportunity to adjust their teaching while it is still in progress. As Heritage and Heritage (2013, p. 176) conclude in their study, ‘open and respectful pedagogical questioning is a key resource in eliciting students’ current status, and for making decisions about next steps in student learning’; thus, the researchers claimed that ‘questioning lies at the epicenter of formative assessment’ (Heritage & Heritage, 2013, p. 178).

A large volume of published studies have investigated the use of questioning in the classroom setting; however, much of the research up to now has been limited to two topics: the use of questioning in the classroom; and the classification of teachers’ questions according to a previously established questioning framework, such as Bloom’s taxonomy (Heritage & Heritage, 2013; İnan & Fidan, 2013; Ozuem & Lancaster, 2015; Shen & Yodkhumlue, 2010; Tan, 2007; Wong, 2010). Although these studies have shed light on the practice of classroom questioning, they have failed to address how it can be applied to improve students’ achievement and overall learning.

To date, only a few studies have attempted to fill the gap on the needs to provide information about how questioning can be implemented. For example, Barnette et al. (1995) developed a project entitled QUILT (Questioning and Understanding to Improve Learning and Thinking) which focused on improving teachers’ questioning techniques when stimulating students’ thinking. QUILT is divided into the following stages: (1) preparing questions, which consists of formulating types of questions, their wording and context; (2) presenting the questions,
which includes indicating the expected answer; (3) prompting students’ responses, which includes providing wait time after presenting the questions and encouraging nonparticipating students; (4) processing students’ responses, providing good feedback and probe questions; and (5) critiquing the questioning episode, reflection of the question-answer session, and figuring out ways to improve the technique.

Another strategy to improve teachers’ questioning strategy is highly effective questioning (HEQ) developed by Hannel (2009) who demonstrated that there are three prerequisites for effective questioning. First, teachers should be able to engage students in the questioning culture. He argued that non-participating students may not have a problem with cognition, but rather may not be accustomed to the questioning culture. A second requirement is that teachers must develop expert patterns. Hannel (2009, p. 66) defined expert patterns as ‘a series of patterns that an expert uses to understand’. Third, teachers should be able to recognize students’ cognitive roadblocks and know how to overcome them.

Further, Slack (1998, cited in Afflerbach 2010) provides a self-assessment checklist for teachers to identify their own questioning practice. This checklist is illustrated below:
A Checklist for Asking Appropriate Questions

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>___</td>
<td>I ask questions that are appropriately phrased and understood by students</td>
</tr>
<tr>
<td>___</td>
<td>I ask questions that are at an appropriate level for the material being covered</td>
</tr>
<tr>
<td>___</td>
<td>I ask questions that require students to think at various intellectual levels</td>
</tr>
<tr>
<td>___</td>
<td>My questions follow a logical sequence</td>
</tr>
<tr>
<td>___</td>
<td>Student responses are used to guide my next questions</td>
</tr>
<tr>
<td>___</td>
<td>I ask questions that assess student understanding</td>
</tr>
<tr>
<td>___</td>
<td>I ask processing questions if a student’s answer is incomplete or superficial</td>
</tr>
<tr>
<td>___</td>
<td>I encourage students to answer difficult questions by providing cues or rephrasing</td>
</tr>
<tr>
<td>___</td>
<td>I avoid closed questions that restrict students’ demonstration of learning</td>
</tr>
</tbody>
</table>

**Figure 2.5 Checklist for asking appropriate questions**

Studies into the use of questioning for formative assessment have been extensively explored; however, these studies mostly focused on identifying the use of questioning and classifying teachers’ questions in classrooms. Although these studies have provided an insight into how teachers apply questioning, they have not addressed the crucial issue of how questioning, as an instrument of formative assessment, can be carried out to improve learning. To date, there are only a few studies that focus on this area and most of them are limited to suggestions. These studies would have been more useful if they had included empirical data on how teachers actually implement those strategies.

**2.3.2.2 Tests**

Broadly defined, testing as an instrument aims to measure students’ competence against specific assessment criteria (H. D. Brown & Abeywickrama, 2010; Popham, 2003; Stobart, 2008). There are multiple formats of tests such as yes/no questions, true/false questions, multiple-choice questions, gap-filling completion, cloze test, matching questions, essay questions, and so on (H. D. Brown & Abeywickrama, 2010; Hughes, 2013; McNamara, 2000). Tests have often been considered as exclusively summative assessment; however, as
stated previously, any assessment instrument can be applied for formative assessment. In other words, tests can also be used for formative purposes. Although it is important to note there have been several misconceptions regarding testing as an instrument for formative assessment, for example, that formative assessment relates to frequent tests; that tests are not part of formative assessment; and that alternative assessment does not include tests (H. D. Brown & Abeywickrama, 2010; Chris Davison & Leung, 2009; Wiliam, 2011).

In fact, summative testing itself can be used for formative assessment. This idea was first established by Black et al. (2003) as one of the strategies in the KMOFAP project (see Section 2.2 for details about the project). This idea is known as formative use of summative tests (henceforth, FUST). There were three FUST strategies recommended in the project:

1. Students should be engaged in a reflective review of the work they have done to enable them to plan their revision effectively.
2. Students should be encouraged to set questions and mark answers to help them understand the assessment process and to focus their efforts towards improvement.
3. Students should be encouraged through peer- and self-assessment to apply criteria to help them understand how their work might be improved. This may include providing opportunities for them to re-work examination answers in the class. (Black et al., 2003, p. 56)

Further, Black et al. (2003) point out that summative testing can be advantageous as long as it is used to improve learning rather than direct learning. In other words, summative tests have the potential to enhance learning only if teachers use the test results to adjust instruction and not merely for summing up learning. Specifically, Carless (2012, p. 118-119) elaborates several advantages of FUST:

… acknowledging the dominance of summative assessment, while bringing a formative orientation to testing; having potential for the development of productive synergies between summative and formative assessment; and as a ‘restricted’ form of formative assessment that could serve as an entry point to learning-oriented assessment practices.

In addition, there are other benefits of FUST as reported by Lam (2013), for example, it can improve students’ performance and promote modest self-regulation. Unfortunately, studies
on FUST are relatively scarce. Most studies in this area are focused on how to align formative and summative assessment instead of making use of tests for formative assessment (for example, Broadbent et al., 2017; Dixson & Worrell, 2016; Gulikers, Biemans, Wesselink, & van der Wel, 2013; Houston & Thompson, 2017; Lau, 2016; Taras, 2008; Wei, 2015). Two of the current studies concerned with FUST were conducted by Lam (2013) and Xiao (2017). Drawing on the work of Black et al. (2003) and Carless (2012), Lam (2013) conducted research that focused on students’ perspectives on the use of preparation tests to improve students’ performance in internal school tests. The study which was conducted in Hong Kong primary schools involved two teachers and 14 students; it used data from interviews and classroom observation. Lam (2013, p. 69) reported that test preparation ‘could enhance students’ performance and promote modest self-regulated learning’. A more current study on FUST application was conducted by Xiao (2017). This multiple case study, which was carried out in the test-dominated context of China, involved three classes in a secondary school. Drawing on data from interviews and classroom observations, the researcher concluded that summative tests have the potential to improve learning to a certain extent, especially as a follow-up strategy.

2.3.3 Providing effective feedback that moves learning forward

The third strategy for implementing formative assessment relates to provision of feedback. Feedback can broadly be defined as information provided to students that focuses on the description of students’ achievement after completing a learning task. Meanwhile, Hattie and Timperley (2007, p.81) simply defined feedback as ‘a consequence of learning’. Feedback can be delivered in written or oral form, or even in the form of a gesture, as long as it includes constructive criticism, confirmation/negation, and advice as its focus (Moss & Brookhart, 2009). Traditionally, teachers are the sole providers of feedback; however, recent
studies have mentioned that feedback may also derive from sources such as peers, books, or parents (Brookhart, Moss, & Long, 2009; Colby-Kelly & Turner, 2007).

A number of studies have postulated that the connection between feedback and formative assessment is inevitable (for example, Black & Wiliam, 1998b, 2003; Heritage, 2010; OECD, 2005; Stiggins et al., 2006). The rationale behind this is that feedback functions as a means to improve learning, which is the most important concept of formative assessment; hence, Black and William (1998b, p. 47) point out ‘the two concepts of formative assessment and of feedback overlap strongly’, while Brookhart, Moss, and Long (2010) refer to feedback as an integral part of formative assessment. Specifically, feedback has the potential to provide the following benefits for learning: it motivates students to enhance their learning as they can recognize their strengths and weaknesses; it provides information about a learning area that needs to be improved; and it facilitates self-regulated learning (Black & Wiliam, 1998b, 2006; Hattie & Timperley, 2007; Kumar & Stracke, 2017; Shepard, 2005; Shute, 2008). In addition, in the field of language learning, feedback enhances language development because teachers use language that is slightly beyond the students’ level (Trumbull & Lash, 2013).

Nevertheless, Kluger and DeNisi (1996) point out that providing feedback is not automatically directed to learning improvement. In their meta-analysis study, Kluger and DeNisi (1996) showed that when students receive feedback, they follow-up with one of the following responses: change behaviour, change goal, abandon goal, or reject feedback. Likewise Hattie and Timperley (2007) assert that students may accept, modify, or reject feedback. This suggests that there are some enforcing factors that may increase the effectiveness of feedback. The literature on formative assessment feedback has highlighted several requirements that make feedback effective. For example, D. R. Sadler (1989) lists three conditions that need to be fulfilled for effective feedback:
a. possess a concept of the standard (or goal, or reference level) being aimed for
b. compare the actual (or current) level of performance needs to be compared with the standard
c. engage in appropriate action which leads to some closure of the gap (D. R. Sadler, 1989, p. 121)

D. R. Sadler (1989) and Hattie and Timperley (2007) suggest three similar points on how to make feedback effective. First, feedback concerns students’ understanding of the established goals because they contribute to good feedback; as Hattie (2012) pointed out, goals should be well-defined in order to close the learning gap while the power of feedback lies in closing the discrepancy between learning goals and students’ current performance.

Second, feedback should provide a snapshot of students’ current learning. Students should receive a vivid description of their position in relation to the learning goals to help them navigate their learning. Last, feedback should provide information on how to close the gap. It is at this point that feedback becomes formative feedback. When students are only given a summary of their performance and no information on how to close the learning gap, then it is not formative feedback.

**2.3.3.1 Formative feedback**

Much of the current literature on feedback pays particular attention to formative feedback which serves two important functions, that is to improve learning and to alter learning (Irons, 2008; Kumar & Stracke, 2017; Ruiz-Primo & Li, 2012; Shute, 2008; Wiliam, 2012). Ruiz-Primo and Li (2012) maintain that formative feedback needs to fulfil the following requirements: first, it involves students, process, and scaffolding to improve learning outcomes; second, it should ensure that it is useful, accessible, practical, and varied; lastly, it should relate to the learning trajectory of the lessons.

The term formative feedback has often been used to differentiate it from feedback in summative assessment. Summative assessment feedback has the following features: isolated
comments about the quality of the student’s work; connected mostly with evaluative
judgements; and applied numbers or a grade that informs students of their achievement, in
terms of the learning objective (Schneider & Randel, 2010). The distinction between
formative feedback and summative feedback is further illustrated by the following diagram:

![Diagram illustrating the formative feedback loop]

**Figure 2.6 The formative feedback loop**
Adapted from Irons (2008), Kumar and Stracke (2017), Randall and Mirador (2003), Ruiz-
Primo and Li (2012) and Shute (2008)

Besides providing formative feedback, the literature of feedback has highlighted several
requirements in order to make feedback enhance learning. First, both teacher and students
should have the same perceptions of what constitutes good feedback. As revealed by
Carless’ (2006) study there have been differing perceptions between students and teachers on
how feedback benefits learning. This exploratory study was conducted in higher education in Hong Kong involving 460 staff and 1,740 students. Data were collected from questionnaires, focus group discussions and interviews. Findings of the study showed that teachers and students have different perceptions about the extent of feedback detail, usefulness of feedback, appropriate forms of feedback, and marking criteria. Thus far, several studies have suggested that there should be an agreement between teachers and students on how feedback will be provided (Higgins, Hartley, & Skelton, 2002; Nicol & Macfarlane-Dick, 2006; Safari & Pourhashemi, 2015).

The second requirement is to provide quality feedback. Nicol and Macfarlene-Dick (2004, p. 11) defined quality feedback as ‘information that helps students trouble-shoot their own performance and take action to close the gap between intent and effect’. Studies indicate that the features of quality feedback are: that it should be relevant to students’ needs, and that it should be comprehensible, timely, and clear. In addition, quality feedback should stimulate thought and open dialogue (Hattie & Timperley, 2007; Kumar & Stracke, 2017; Safari & Pourhashemi, 2015; Shute, 2008; Wiliam, 2012).

Moss and Brookhart (2009) emphasise that teachers’ quality feedback is influenced by the methods that they employ when giving feedback. They elaborate four factors that affect the effectiveness of feedback:

a. **The amount of feedback.** Shute (2008, p.157) refers to this as feedback complexity which she defined as ‘how much and what information should be included in the feedback messages’. The amount of feedback should be determined by students’ characteristics and the nature and materials of the task (Brookhart et al., 2009; Shute, 2008).
b. **Timelines.** Feedback can either be immediate or delayed. Different theories exist in the literature regarding which is the best time to give feedback. Although several studies suggest giving immediate feedback (Irons, 2008; Metcalfe, Kornell, & Finn, 2009), Shute (2008) concludes that immediate and delayed feedback can be implemented by teachers because each has its own strengths and weaknesses.

c. **Mode** relates to the form in which the feedback is given to students. Although a number of articles mention that scores as a form of feedback are ineffective at improving students’ learning (Black et al., 2003; Hattie & Timperley, 2007; Safari & Pourhashemi, 2015), to date, they are still the most prominent forms of feedback (Brookhart, 2008; Marzano et al., 2001; Oosterhoof, 2009). Black et al. (2003) suggest giving ‘comments only feedback’ as this type of feedback helps parents to focus and support students’ learning. However, they also note that it takes time for teachers to develop the skills to give effective comments as feedback.

d. **Audience.** Feedback may be given to individuals or to a collective. The literature on feedback has highlighted the importance of individual feedback (Broadbent et al., 2017; Brookhart et al., 2009; OECD, 2005; Quyen & Khairani, 2016; X. Yin & Buck, 2015).

Besides having the same perceptions of good feedback between teacher and students and providing quality feedback, another requirement for feedback to be effective is good instruction. Hattie (2012, p. 43) points out that ‘good instruction produces good feedback’; hence, feedback is influenced by how learning goals are formulated and the learning activities that teachers create to gain evidence of student learning. Wiliam (2012) describes predictability as being another element of good teaching. If learning is predictable, effective instruction can take place, and quality feedback can be given by teachers. In addition,
students’ active role also contributes to good instruction because it is needed to close the learning gap (Kumar & Stracke, 2017; Ruiz-Primo & Li, 2012; D. R. Sadler 1989).

To conclude, feedback has a crucial role in improving learning; nevertheless, if the purpose of feedback is only to give information about the gap between learning objectives and students’ current performance, then it is unlikely to enhance learning. Feedback should include information on how to close the learning gap, and only formative feedback can fulfil this requirement. Other requirements that make feedback effective are: teacher and student having the same perceptions of what constitutes good feedback; providing good quality feedback; and giving good instruction. Studies on formative assessment as explained in this section have provided lucid information; however, much of the research up to now has been descriptive in nature. The research to date has tended to focus on suggestions on how to give feedback rather than providing information on what actually works in the classroom.

2.3.4 Activating students as instructional resources for one another

As was pointed out in Section 2.2, formative assessment is situated under socio-constructivist learning theory which embraces the notion of students’ active involvement in learning. In the previous section, I discussed the pivotal role that students’ active involvement has in closing the feedback loop so as to address the learning gap. In this section, I discuss the fourth strategy of formative assessment implementation (Wiliam & Leahy, 2015), which focuses on how students can act as learning resources for one another. This strategy is in agreement with Vygotsky’s developmental theory that supports the role of ‘others’ in the ZPD. Hence, the socio-constructivist learning theory has changed peers’ roles from previously believed as competitor to collaborator.

There is a large volume of published studies describing the crucial role of peers in learning, or ‘peer learning’ (Topping, 2005, p. 631), for students’ improvement (Carless, 2011;
Panadero, Brown, & Strijbos, 2016; Topping, 1998, 2009). Peer learning has been variously referred to as collaborative learning, cooperative learning, peer monitoring, peer cooperation, peer tutoring, and peer counselling (Carless, 2011; Topping, 2005). Among these concepts of peer learning, much of the current literature pays particular attention to peer assessment (Panadero et al., 2016; Topping, 2005; van Gennip, Segers, & Tillema, 2010). The term peer assessment is generally understood to mean an arrangement of peers in pairs or groups to assess each other’s learning product, such as an essay, portfolio, oral presentation, and so on. This activity can be carried out in reciprocal or one way, and can be conducted inside or outside classroom settings (Carless, 2011; Sebba et al., 2008; Topping, 2009). Sluijsman, Brand-Gruwel, van Merriënboer, and Martens (2004) demonstrated that when students are involved in peer assessment, there are three constituent skills required to conduct the activity as illustrated by the figure below:
Figure 2.7 Peer assessment model
Source: Sluijsmans et al. (2004, p. 62)
The literature on peer assessment as one of the instruments of formative assessment describes its advantages as enhancing students’ motivation, improving learning engagement and interaction, and initiating self-regulated learning (Black et al., 2003; Bransford, Brown, & Cocking, 2000; Carless, 2011). Furthermore, peer assessment provides additional feedback from teachers, which is considered to be a potential learning resource for the following reasons. First, peers share the same ‘language’; thus, learning points can be more comprehensible. Teachers’ explanations may be difficult to understand because of their different style of discourse. Second, peers provide better learning opportunities because students can ask, interrupt, clarify, and discuss learning points. These opportunities are not common occurrences with teachers, because students tend to accept what teachers say. As Carless (2011, p. 154) said ‘with peers, students are less likely to fake comprehension than when they are interacting with teachers; they are more likely to interrupt and ask for clarification; peers use different language to explain and sometimes this can be more readily understood’. Consequently, peer assessment is deemed to be part of formative assessment practice due to its potential role in giving abundant feedback in order to improve learning (Carless, 2011; Panadero, Brown, et al., 2016; Topping, 2009).

Nevertheless, the practice of formative assessment has not gone unchallenged; more recent attention has focused on the critiques of feedback provision from peers. A study conducted in secondary schools in Belgium found that there is a tendency for students to give less value to feedback from peers (Gielen, Peeters, Dochy, Onghena, & Struyven, 2010) and similarly for a study conducted in higher education in China (Yang et al., 2006). In addition, a study conducted by Mok (2011) in a secondary school in Hong Kong revealed that students encountered anxiety in the process of peer assessment because they felt they weren’t qualified to assess their peers’ work. Another issue raised by several studies concerns the validity and reliability of peer assessment (Dochy, Segers, & Sluijsmans, 1999; Miller, 2003;
Nelson & Murphy, 1993). Moreover, a review from six peer assessment studies conducted by Strijbos and Sluijsman (2010) revealed that, to a certain extent, peer assessment can fall into the category of summative assessment. This relates to the practice of peer assessment which assigns students to score their peers; thus negating the concept of peer assessment which aims to improve learning and not to sum up the assessee’s performance.

However, much of the current literature revealed the opposite findings: feedback from peers can be reliable and valid (Cho, Schunn, & Wilson, 2006; Falchikov, 2005; Strijbos & Sluijsmans, 2010). In addition, it is important to note that although Gielen et al. (2010) found that students gave less value to their peers’ feedback, their experimental study showed that peer assessment did improve learning. Further, studies suggest that in order to keep peer assessment formative, scores assigned by peers should not be more than 10-15% of the total score (Ballantyne, Hughes, & Mylonas, 2002; Race, Brown, & Smith, 2005).

Topping (1998, p.254) highlights that peer assessment is primarily ‘learning by assessing’. In the same vein, Black et al (2003, p. 62) point out that ‘the real purpose (of peer assessment and self-assessment is) the identification of learning needs and the means of improvements’; in other words, peer assessment should focus on students’ learning enhancement (Black et al., 2003; P. M. Sadler & Good, 2006; Strijbos & Sluijsmans, 2010). In the same vein, van Zundert, Sluijsmans, and van Merrinboer (2010) found in their meta-analysis study that training and experience in peer assessment enhanced psychometric qualities; increased students’ positive attitude towards peer assessment; and developed students’ thinking styles and academic achievement.

A more comprehensive guide on how to conduct peer assessment for learning is outlined by Panadero et al. (2016) who maintain that peer assessment (PA) should include the following stages:
a. Clarify the purpose of PA, its rationale and expectations to the students;
b. Involve students in developing and clarifying assessment criteria;
c. Match participants (e.g. individual, groups) in a way that fosters productive PA;
d. Determine the PA format (e.g. rating with or without comments) and mode of PA interaction (e.g. face-to-face or online);
e. Provide quality PA training, examples and practice (including feedback about PA);
f. Provides rubrics, scripts, checklists, or other tangible scaffolding for PA;
g. Specify PA activities and timescale;
h. Monitor the PA process and coach students (Panadero, Jonsson, & Strijbos, 2016).

To date, there has not been any follow-up study that provides further information as to whether the guideline above has actually improved peer assessment practice and overall learning.

To sum up, this section has presented the crucial role of peer learning; specifically, peer assessment, as the fourth strategy in implementing formative assessment. As mentioned in this section, peer assessment has the potential to improve learning because of its role in giving additional feedback to students. Nevertheless, the use of peer assessment has been contested; several issues have been raised including students’ anxiety to assess their peers’ work; feedback validity and reliability; its emphasis on summative assessment; and student’s distrust towards feedback from their peers. Several studies have tried to address the issue and have suggested that ‘learning by assessing’ is one of the most important peer assessment instruments for formative assessment. Unfortunately, studies of how this notion is implemented, in a real classroom situation, remain unexplored.

### 2.3.5 Activating students as the owners of their own learning

The last strategy to be discussed involves students assessing their own learning (Wiliam & Leahy, 2015). This strategy is central to formative assessment because peers and teachers may not always be present during the learning process; consequently, the ability of students...
to assess their own work is pivotal for their learning. One of the assessment instruments that is directly related to this strategy is self-assessment; thus, Black et al. (2003, p. 14) point out, ‘The development of self-assessment by the student might have to be an important feature of any program of formative assessment’. In other words, self-assessment holds a prominent role in formative assessment; although it is important to note that self-assessment becomes formative when the activity involves students reflecting on the quality of their own learning, and using the information for improvement and/or revision if it is necessary (Heidi Andrade & Valtcheva, 2009).

Self-assessment can be broadly defined as an activity in which students appraise their own work on any learning product, such as portfolios, learning logs, diaries or other assessment records, often, based on teacher’s assessment criteria (Heidi Andrade & Valtcheva, 2009; Boud, 1995a; McMillan & Hearn, 2008; Panadero & Romero, 2014). In terms of procedures, self-assessment includes the following stages: recognizing learning performance and determining its quality; formulating self-feedback; and taking relevant follow-up action, such as learning remediation, extension, or correction (G. T. L. Brown, Andrade, & Chen, 2015; McMillan & Hearn, 2008).

A large and growing body of literature has highlighted several benefits of applying self-assessment: it helps students to construct their own learning; assists students in strengthening their ability to make connections; enhances students’ understanding; and contributes to motivation and enhances confidence; develops student thinking; improves metacognitive skills; enhances students’ ability to internalize learning standards; reduces students’ anxiety; and helps students to direct learning (Andrade, 2010; Boud, 1995b; M. Harris, 1997; McMillan & Hearn, 2008; Panadero & Romero, 2014; Shepard, 2005; Wiliam & Leahy, 2015). Further, studies have established that the foremost benefit of self-assessment is self-
Self-regulation is defined by Boekaerts et al. (2005, p. 150) as ‘a multi-component, multi-level, iterative, self-steering process that targets one’s own cognitions, affects (i.e. emotions) and actions, as well as features of the environment, for modulation in the service of one’s goals’. Self-regulation can be developed or practiced through self-assessment because it encourages students to become the producers and consumers of their own feedback; thus, students are trained to self-regulate themselves in learning and to be less reliant on feedback from their peers or teachers (Andrade, 2010; Nicol & Macfarlane-Dick, 2006; D. R. Sadler, 1989).

Previous research has established that self-assessment is not only beneficial, but it also receives positive responses from both teachers and students (Bullock, 2011; L. R. Harris & Brown, 2013; Hood, 2012; Leahy, Lyon, Thompson, & Wiliam, 2005). However, the literature on self-assessment has highlighted several critiques of its practice. The first issue relates to its accuracy, specifically whether students’ self-assessment produces similar results to assessment by their peers or teachers (see Section 2.4 for further details of peer assessment). A recent study involving 3,588 first year students in a post-secondary institution in Singapore compared the results of students’ self-assessment of their progress with their teachers’ assessment (Lew, Alwis, & Schmidt, 2010). Findings revealed that there was a weak correlation between students’ and teachers’ assessment results. The researchers argued that the result can be due to students’ lack of competence. Nevertheless, the study confirms the results of several previous studies which have shown inaccuracy in students’ self-assessment (Boud & Falchikov, 1995; Landrum, 1999; Lejk & Wyvill, 2001).

Further, several lines of evidence from research suggest that there is a distrust of delegating assessment to students. For example, a multiple case study conducted by L. R. Harris and
Brown (2013) at various educational levels in New Zealand revealed that teachers still find it challenging to change their mindset from being the sole provider of feedback. In addition, students themselves still preferred teachers’ feedback above their own feedback; as revealed by a study conducted in secondary and elementary school in New Zealand (L. R. Harris, Brown, & Harnett, 2014). A recent study conducted in a writing course, in a Singaporean university, revealed similar findings (H. H. Lee, Leong, & Song, 2017): students’ attitudes and institutional requirements hamper the implementation of self-assessment.

The existing literature on self-assessment suggests providing a set of learning criteria, commonly known as rubrics, to address the issue (Andrade & Valtcheva, 2009; Panadero, Brown, et al., 2016; Panadero & Romero, 2014; Sadler & Good, 2006; Wiliam & Leahy, 2015). The effect of using rubrics in self-assessment activity was investigated by Panadero and Romero (2014). This quasi-experimental study involved 218 third year pre-service teachers in Spain. The study revealed that groups employing rubrics for assessment criteria have higher learning strategies, performance and accuracy. Other studies document several factors that need to be taken into account when applying self-assessment, such as learners’ current and previous knowledge; learning context; the requirements of the subject matter; practicality; and readiness (Boud, 1995b; Bullock, 2011; M. Harris, 1997; Heritage, 2010; P. M. Sadler & Good, 2006).

Collectively, these studies indicate that self-assessment has the capacity to improve learning because of its potential benefits; thus, it suits the notion of formative assessment provided that the results of self-assessment are used to adjust learning. Nevertheless, there remain several aspects of the application of self-assessment about which relatively little is known. For example, how do teachers apply self-assessment in their own context without researcher intervention; to date, most studies in self-assessment have been experimental.
This section has reviewed five key strategies used to implement formative assessment, namely, clarifying, sharing and understanding learning intentions and successful criteria; engineering effective classroom discussion, questions, and learning tasks; providing feedback that moves learners forward; activating students as owners of their own learning; and activating students as instructional resources for one another. Each of these strategies was explained in terms of its principle, benefits, prevailing criticism, and implementation in classroom settings.

Nevertheless, this section has shown that there are some gaps in the literature that need to be addressed. To date, much of the research in the field of formative assessment can be divided into two categories: first, studies which are descriptive in nature (for example, Wiliam & Leahy, 2015; Wylie, 2008). Second, studies which involve researcher intervention in the process of implementing formative assessment. Although these studies have shed light on how formative assessment can be carried out, they would have been more useful for teachers if they had focused on how teachers, in an authentic classroom setting, might apply formative assessment strategies and use the result to improve learning. Providing such examples is crucial because translating these strategies can be challenging for teachers, who require practical assessment instruments which are readily implemented.

In the following section, I present a review of literature concerning the challenges encountered by teachers implementing formative assessment and discuss the ways that teachers cope with the challenges.

### 2.4 Identifying formative assessment challenges and their potential solutions

The previous section has shown that there are formative assessment strategies which can be used by teachers to implement the practice; nevertheless, translating these strategies can be
challenging for teachers unless clear and practical examples are provided for them. This section discusses the challenges encountered by teachers in implementing formative assessment, such as large class sizes, high-stakes tests, and level of assessment literacy. How teachers overcome these challenges is also presented in this section.

2.4.1 Uncovering formative assessment challenges

The literature on formative assessment has highlighted several factors that can present barriers to teachers when implementing formative assessment. The first factor is large classes. There are multiple definitions of what constitutes a large class but they are dependent on the learning context (Hayes, 1997; Nunan, 1991; Richards, 2017; Z. Xu, 2001). Despite these varying definitions, researchers are generally in agreement that large classes can be challenging (Ballantyne, Hughes, & Mylonas, 2002; Buchanan & Rogers, 1990; Carbone & Greenberg, 1998; Hancock, 2010; Hayes, 1997; Locastro, 1989; Schanzenbach, 2014). Large classes can be a barrier for formative assessment because teachers need to attend to students’ different learning needs, especially when giving feedback in order to move learning forward and to organize students as they are doing peer learning (Brady, 2012). As reported by OECD (2005), teachers in secondary schools, in eight different countries, claimed that applying formative assessment on a daily basis is not feasible due to large classes and the need to cover the materials outlined in the curriculum.

The second factor that can hamper the implementation of formative assessment are high-stakes tests which are defined by Wright (2008, p. 4) as ‘educational measures that have significant negative consequences for failure’. These tests have been considered to provide a benchmark in measuring quality in education. Often, in some countries formative assessment is enacted through their assessment policy, but at the same time these countries also administer high-stakes testing; as shown in Gu’s (2014) study. This study focused on how
teachers in China practise new curriculum and assessment policies which emphasise formative assessment. This single case study used interview and video recording as the source of data. Their findings showed that it is high-stakes tests that determine the materials and methods used in the classroom rather than the curriculum standards. This situation has often placed teachers in a dilemma between implementing formative assessment and prioritizing their teaching to enable students to pass high-stakes testing (Black, 2015).

High-stakes testing may deplete teachers’ innovation and creativity: two factors crucial for performing formative assessment. Instead, teachers are under pressure from stakeholders, such as principals, parents, and local authorities to ensure that students pass the exam. High-stakes tests have created a testing culture which permeates teaching and learning (see Section 2.1 for further details of testing culture) as revealed in a study conducted by Box, Skood and Dabbs (2015). This multiple case study, which involved three secondary high school science teachers in a Western Texan suburban community, focused on factors that may constrain the implementation of formative assessment. Drawing on the data from classroom observation, teacher interviews, informal communication, and physical artefacts, their findings revealed that high-stakes examination is a significant barrier to teachers implementing formative assessment. This is because they are under pressure to cover the materials that may be included in the high-stakes tests. This study also showed that high-stakes tests do not only affect teachers’ practice, but they also influence students to pay more attention to high-stakes test-related materials. A study conducted in a higher education setting revealed that the testing culture may also shape students’ learning; and as a consequence, students preferred summative assessment to formative assessment (Umer, Javid, & Farooq, 2013). The researchers found that 283 English major students, in the Department of Foreign Languages in Saudi Arabia who were accustomed to high-stakes tests during their secondary years, wanted their teacher to use exam-focused learning materials so that it would increase their
scores. They were inclined to memorize instead of gaining in-depth knowledge from the learning materials. This study implies that high-stakes tests create long-term impacts.

The third factor that can hamper the implementation of formative assessment is a low-level of assessment literacy. A recent meta-analysis study conducted by Quyen and Khairani (2016) showed that large classes and high-stakes testing can indeed hamper formative assessment. However, these researchers found that in the Asian context, formative assessment barriers were directly linked to teachers’ perception of formative assessment as time consuming and as a burden, and teachers’ lack of understanding of what formative assessment is and how to implement it in their daily practice. In other words, the substantial challenge may actually relate more to teachers’ understanding of formative assessment. A study in a higher education setting in Saudi Arabia, conducted by Al-Alwassia, Hamed, Al-Wassia, Alafari, and Jamjoom (2015), which investigated challenges encountered by students and faculty members in applying formative assessment found that teachers have inadequate knowledge of assessment for learning. In addition, they found that students place higher value on grades than teachers’ feedback. These studies have shown that teachers’ assessment literacy level varies significantly.

Assessment literacy is the teacher’s ability to conduct quality assessment (Fulcher, 2012; Popham, 2011; Stiggins, 1999). Gu (2014) points out that assessment literacy consists of the why, what, and how of assessment; the feedback and learning opportunities and the awareness and practice of ethical assessment. Stiggins (1999, p. 535) characterizes teachers who are assessment literate as those who ‘… have basic understanding of the meaning of high and low quality of assessment and are able to apply that knowledge to various measures of students’ achievement’.
In a similar way, Webb (2002) defined assessment literacy as the knowledge about how to assess what students know and can do, interpret the results of those assessment, and apply the results to improve student learning and program effectiveness. Boyles (2006) argued that the concept of assessment literacy goes beyond choosing, constructing, and implementing assessment instruments to include the expertise required to analyse empirical data and enhance learning. These definitions have extended the notion of assessment literacy and have postulated a convergence between the notion of formative assessment and assessment literacy: teachers’ formative assessment practice may relate to their level of assessment literacy.

In addition, it has been widely accepted that teachers hold the central role in improving students’ achievement, and learning in general (Darling-Hammond, 2013; Marzano, 2001; Moir, Barlin, & Miles, 2009). In fact, Wiliam and Leahy (2006) argued that a teacher’s attitude and commitment has a more significant impact than factors such as large classes or socioeconomic factors. This view is supported by Davison (2004) who argued that improving learning can be formulated into curriculum or other educational policies by government, but the fundamental role of teachers is to translate the visions included in the government policies into actual teaching and learning activities.

Assessment literacy is pivotal to formative assessment and teachers need to incorporate the notion into their own classroom practice (Black & Wiliam, 1998b, 2009); in other words, formative assessment demands teachers be creative and innovative. This can be daunting for teachers because the benefits are not immediate; as mentioned in a number of studies, formative assessment is not a ‘quick fix’ (Black & Wiliam, 1998b; Colby-Kelly & Turner, 2007; Higgins & Grant, 2010; Wiliam et al., 2004; Wylie & Lyon, 2015).
According to Bullock (2011, p.114) ‘innovation places new demands on teachers, and may end in failure’; hence, Young and Jackman (2014) argued that if teachers view formative assessment as demanding in terms of time and resources, they very likely will avoid incorporating it into their instruction. Similarly, Volante and Beckett (2011) maintain that teachers may develop a negative attitude as soon as they discover difficulties in implementing the suggested techniques of formative assessment. According to Fulcher (2012) teachers want clear, practical activities that are readily implemented, rather than general guidelines that require them to spend more time customizing the concept to their teaching. Thus, formative assessment may become challenging for teachers unless it is supported with examples of how it can be carried which then will be customized in teachers’ own context; as Pham and Renshaw (2015) emphasised that any assessment reform has the potential to be adopted, adapted, or rejected by both teachers and students although it depends on whether the reform requires them to adjust their previous beliefs and knowledge.

2.4.2 Coping with challenges

A considerable amount of literature has highlighted the potential of integrating educational technology to overcome challenges in the implementation of formative assessment (Beatty & Gerace, 2009; Hargreaves, 2003; Looney, 2010; Trucano, 2012). Specifically, educational technology can overcome problems in dealing with large classes and lack of time: the two common challenges to the implementation of formative assessment (Russel 2010). Educational technology is defined as ‘a combination of the processes and tools involved in addressing educational needs and problems, with an emphasis on applying the most current tools: computers and other electronic technologies’ (Roblyers & Doering, 2010, p. 8).

Integrating technology for formative assessment purposes may include various devices or software, one of the most prominent devices which is used for this purpose is CRS
(Classroom Response System). The device may have different names such as: clickers, student response systems, audience response systems, electronic response systems, personal response systems, and zappers (P. White, Syncox, & Alters, 2011). Basically, the device enables teachers to receive immediate feedback from multiple choice questions posed by teachers. CRS benefits teachers who have to deal with big classes. In addition, CRS helps teachers to make teaching adjustments. A number of studies have explored the effectiveness of implementing CRS for formative assessment purposes. For example, a study conducted by Feldman and Capobianco (2008) in US focused on the integration of technology-enhanced formative assessment using CRS. Participants in the research were physics teachers who were interested in improving their use of CRS, and the data were collected through interviews, direct observation, and collaborative discussions. Based on the data, researchers were able to construct a model for using CRS effectively. The findings of this research revealed that educational technology is commonly implemented to overcome teachers' workload while conducting formative assessment, especially in writing. The effectiveness of clickers was also investigated by Yourstone, Kraye, and Albaum (2008). This experimental study into an undergraduate operational manager course revealed that clickers have positive impacts for students’ learning.

The use of technology is ‘undeniable and widespread’ (Trucano, 2012, p.101); the potential benefits of educational technology has urged schools to integrate technology into classroom instruction. Nevertheless, a number of studies document that in some cases teachers have not implemented technology effectively. As Falchikov and Thompson (2008, p. 51) argue, some teachers use technology only for ‘fulfilling the formal requirements’; in other words, because the school has equipped them with the technology devices, they may use those devices without knowing exactly how or why (Wang & Reeves, 2008, p. 65). As Cuban (2003) points out simply providing technology is insufficient to ensure teachers use it effectively.
According to Feldman and Capobianco (2008), there are two possibilities behind the ineffective use of technology: lack of knowledge of how to use it for instructional purposes, and lack of time for developing that knowledge. Further, Feldman and Capobianco (2008) suggest that teachers should understand the nature of formative assessment and should understand how to integrate educational technology into it.

Professional development is also an important means to overcoming the difficulties in implementing formative assessment. As Trumbull and Lash (2013, p. 7) said ‘substantial professional development is needed in order for teachers to embrace formative assessment resulting in a change in practice’. A large and growing body of literature has investigated how teachers’ professional development programs can be a means to enhancing teachers’ formative assessment literacy. One of the most prominent ones is KMOFAP (see Section 2.2 for details about the project); several reports of the project document that their program had positive impacts on teachers’ formative assessment practice (Black et al., 2003; Harrison, 2005; C. Lee & Wiliam, 2005).

Another teachers’ professional development program which has the same purpose is the Keeping Learning on Track (KLT) program. It is a two-year program which consists of a two day workshop, and two day TLC (Teaching Learning Community) leader workshop. KLT consists of three main components: content component, the knowledge of the feature of assessment for learning or formative assessment on a daily basis; process component, an ongoing program of school-based collaborative learning; and empirical/theoretical component, including reasons why it is important for teachers to apply formative assessment in their assessment practice. This program has been adopted in several studies such as the Northwest Evaluation Association (NWEA) program (Thum et al., 2015). The researchers
found that when KLT was implemented in 20 schools, it contributed positively, over time, to teachers’ practices and students’ engagement.

Besides KMOFAP and KLT, another professional development program that focused on improving teachers’ formative assessment practice was Classroom Assessment for Student Learning (CASL) which was developed based on the work of Sadler (1989) and Atkin, Black, and Coffeey (2001). CASL emphasises matching learning targets with assessment methods, providing descriptive feedback, and activating student involvement in learning. CASL has been replicated in several teachers’ professional development programs. For example, Randel, Apthorp, Beesley, Clark, and Wang (2016) applied the program in public schools in Colorado. Findings from the study showed that the treatment group, which was selected randomly out of 67 schools, significantly improved their assessment knowledge and the students’ involvement.

Other professional development programs also showed to be successful in enhancing teachers’ formative assessment practice. For example, McGatha, Bush, and Rakes (2009) reported that their program in Louiseville, entitled Formative Assessment in Middle Mathematics (FAMM), had an impact on the cognitive level of teachers’ questioning, use of peer assessment, and types of questioning strategies. In another program, Dibiase (2014) focused on the effectiveness of professional development regarding Linking Learning and Assessment policy in Rhode Island. The study was designed using mixed-methods, and data were gathered through questionnaires (n=85) and focus group interviews (n=5). The study revealed that teachers’ knowledge of formative assessment improved gradually; those who were in the early stages began to realize the power of formative assessment, and those who were at a later stage could see meaningful changes in their practice. The research also revealed that participants considered collaboration, active learning opportunities, and
coherence to be meaningful for teachers. Furthermore, participants commented that professional development should be differentiated based on teachers' individual needs.

Nevertheless, some programs end up with disappointing results. For example, Frohbieter, Greenwald, Steecher, and Schwartz (2011) reported for the National Centre for Research on Evaluation, Standards, and Student Testing (CRESST) project 802 which focused on the analysis of three middle school mathematics formative assessment programs in US. Findings of the project revealed that as teachers become more familiar with the notion of formative assessment, the frequency with which they integrate the practice increases. Further, the pre-existing assessment practice helped teachers to incorporate formative assessment concepts into their practices. However, the researchers mentioned that that they were ‘somewhat disappointed that we did not find more instances of highly nuanced insights from assessment, highly responsive uses of information, and patterns of practice that we would characterize as “true” formative assessment’ (Frohbieter et al., 2011, p. 41).

In summary, I have reviewed the challenges encountered by teachers when implementing formative assessment, specifically large class sizes, high stakes testing and the teachers’ level of assessment literacy. Studies indicate that challenges in implementing formative assessment may relate to the level of teachers’ assessment literacy rather than external factors such as large classes and high-stakes testing. I also reviewed two potential solutions for overcoming the barriers to teachers practicing formative assessment, that is, integrating educational technology and teachers’ professional development. The following section explains formative assessment in an ELT setting.

2.5 Formative assessment in ELT

To date, there is a relatively small body of literature that concerns formative assessment in ELT. Studies in formative assessment have been mostly conducted in the areas of science
and mathematics (for example, Buck & Trauth-Nare, 2009; L. R. Harris et al., 2014; Hill, 2011; Hondrich et al., 2015; Panadero & Brown, 2017; Yin & Buck, 2015; Zhang & Misiak, 2015). It appears that the language learning field has not yet embraced the notion of formative assessment and hence, it is comparatively unexplored.

Thus far, only Jones and Wiliam’s (2008) study has discussed formative assessment strategies in language learning. Surprisingly, Jones and Wiliam (2008) fail to take account of two crucial factors in implementing formative assessment: clarifying learning intentions and addressing the prominent role of teachers in designing formative assessment instruments. This is despite the fact that these strategies were developed by reflecting on the strategies in the KMOFAP program (see Section 2.2 for details about the project). In addition, their study aims for language learning; thus, Jones and Wiliam (Jones & Wiliam, 2008) overlook the fact that there are several different contexts in ELT such as English as a foreign language (EFL), English as a second language (ESL), and English as an additional language (EAL). Other studies of formative assessment in ELT mostly discuss the effectiveness of formative assessment and specific assessment instruments under the notion of formative assessment. In fact, some of these studies only refer to formative assessment construct implicitly.

This section presents an insight into the practice of formative assessment in English Language Teaching. It discusses the changing context of English language learning and teaching, and then looks at the factors that have contributed to language assessment reform. Last, this section talks about formative assessment in English language learning which can be divided into two categories: studies which relate implicitly to formative assessment and formative assessment as a specific construct.
2.5.1 Overview of learning and assessment in ELT

As in other learning fields, the current English pedagogy is also influenced by the changing paradigm of learning theory; the dominant role of behaviourism was replaced by cognitivism, and later by constructivism (see Section 2.1 for details). Although, Woodward (1996) argues that new paradigms cannot be considered as an antidote for the previous learning theory, how learning is perceived and translated into teaching practice by the behaviourism and constructivism learning theories can be easily distinguished as they are relatively opposite to each other. For example, teacher-centred, discrete point of tests, language-focused, rote learning and individual learning were some of the characteristics of language learning during the behaviourism era which is now commonly referred to as the ‘traditional paradigm’ (H. D. Brown & Abeywickrama, 2010; Renandya, Lee, Kai Wah, & Jacobs, 1999). In comparison student-centred, communication-focused, collaborative learning and an emphasis on meaning construction and higher-order thinking are some of the characteristics of learning under constructivist theory (Aljohani, 2017; Birenbaum, 2016; Liu & Zhang, 2014; Wilson, 2009).

Besides the paradigm changes in learning theories, the current ELT practice was instigated by the changing needs of ELT learners - from knowing about the language to using the language for communication. As the Second World War ended, the use of English expanded significantly. According to Richards (2017) a number of factors have contributed to the increased use of English, for example immigrants, refugees, foreign students who studied in English speaking countries; the use of air travel as a means of transportation; and international trade, and commerce which use English for communication. Further, Richards (2017) adds that the expanded use of English has created various learning needs. Previously ELT focused on grammatical accuracy at the sentence level and vocabulary in terms of definition and translation, but this does not accommodate the differing needs of English language learners. Learners’ language needs have shifted from knowing about the language
to using the language for real communication action in the relevant contexts. As the use of English has become globalised, the need for oral and written communication has increased, and the contexts in which it is used have changed.

In addition, the widespread use of English has produced numerous different contexts in which the language is learned, and several different labels are used to make reference to fields in which the language is used. For example, the term EFL refers to the context in which language use is limited to the context of the classroom. In comparison, ESL refers to learning contexts in which English is used in everyday life (H. D. Brown 2007; Meskill, 2010). EAL is used to refer to a context in which the person’s first language is not English, but they do learn it in schools where it is the language of instruction. Thus, these learners need to develop their English proficiency in order to be able to communicate in English (Rea-Dickins, 2001).

To accommodate the need to communicate, students are required to have communicative competence which includes grammatical competence, discourse competence, sociolinguistic competence, and strategic competence (Canale & Swain, 1980). A large and growing literature of ELT has investigated using CLT as a potential means for achieving communicative competence (for example, H. D. Brown, 2007; Larsen-Freeman, 2000; Nunan, 1991; Richards & Rodgers, 2001). As Richard and Rogers (2001, p. 165) emphasise, CLT ‘enables learners to attain the communicative objectives from the curriculum, engage learners in communication, and requires the use of such communicative processes as information sharing, negotiation of meaning, and interaction’; further, they maintain that CLT is not a teaching method, but it is an approach that encompasses any teaching methods that improve learning. Specifically, Nunan (1991) explains that CLT has the following features:

1. An emphasis on learning to communicate;
2. The introduction of authentic texts into the learning situation;
3. The provision of opportunities for learners to focus, not only on language, but also on the learning process itself;
4. An enhancement of the learner’s own personal experience as important contributing to classroom learning;
5. An attempt to link classroom language learning with language activation outside of the classroom (Nunan, 1991, p. 279).

Since assessment is an integral part of learning, the shifting sands of the learning paradigms and learning needs entail changing the assessment processes. Previously, as in other learning fields, assessment in ELT was dominated by standardised testing until social constructivism made its mark and led to a major reform of assessment in ELT. The learning theory which emphasises collaboration, communication, and meaning construction is considered to be incompatible with the hegemony of standardised testing and does not assess students’ communicative competence (Brindley, 2008; Stoynoff, 2012).

Specifically, Spolsky (1978), using the term ‘language testing’ instead of ‘language assessment’, elaborates three eras of language assessment. In the pre-scientific era, language testing was dominated by the grammar translation method, and the concept of reliability and validity were not the main concern. In the second phase, known as psychometric-structuralist, the testing approach was dominated by discrete point items, and focused on reliability and validity. According to H. D. Brown and Abbeywickrama (2010), psychometric-structuralist language assessment was strongly influenced by behaviourism. The learning itself emphasised sentence-level grammatical paradigms, definitions of vocabulary items, and translation from the first to the target language. H. D. Brown and Abbeywickrama (2010) noted that it was in this era that standardised-tests became very popular. The most recent phase was the psycholinguistic-sociolinguistic era which criticised testing and focused on the integration of language skills and tests that can measure students’ communicative ability.
One of the significant impacts from the reform is the emergence of various assessment types which contrasted to standardised testing (Birenbaum, 1996, 2016; J. D. Brown & Bailey, 2008; Davison & Cummins, 2007). These assessment types were variously referred to as alternative assessment, performance-based assessment, authentic assessment, classroom-based assessment, teacher-based assessment, and dynamic assessment. There were several differing opinions on the generic names for these varieties of assessment types. Meskill (2010) proposed referring to them as performance assessment, while J. D. Brown and Hudson (1998, p. 657) suggested labelling them as alternatives in assessment, rather than alternative assessment, because ‘… these new procedures are just new developments in that long tradition’.

Nevertheless, despite the flourishing of assessment types, the change from testing culture to assessment culture, or from standardised testing to teacher-generated assessment is still evolving. In other words, the powerful role of testing remains, and this can be seen from two perspectives. For instance, textbooks on language assessment make an explicit distinction between tests and assessment, and most of their content is dedicated to designing a test (Bachman & Palmer, 2010; H. D. Brown & Abeywickrama, 2010; Cohen, 1994; Hughes, 2013). However, it is important to note that their emphasis is on how to design communicative language testing. According to a study conducted by J. D. Brown and Bailey (2008) on the features of a language testing course, teacher training programs are still focused on testing rather than assessment, and student-teachers are not sufficiently equipped for language assessment.

In summary, in this section I have shown that the current practice of ELT is influenced by two factors: the changing of learning theories, and students’ needs to achieve a competent level of communication. The factors have also made a considerable impact on assessment
practice in ELT. Assessment which previously focused on standardised testing received adverse criticism because it was incompatible with socio-constructivist learning theory and unable to fulfil students’ learning needs. This resulted in a number of assessment types as opposed to the standardised; nevertheless, these assessment types have not yet ended the domination of testing in ELT.

In the section that follows, I present the implementation of formative assessment in an ELT setting.

2.5.2 Overview of formative assessment studies in ELT

As mentioned in the previous section, there have been multiple labels for the notion of assessment types which criticised the domination of testing culture such as alternative assessment, performance-based assessment, or classroom-based assessment. Nevertheless, these assessment types are in line with the notion of formative assessment and they share common features. First, they emerged as an alternative to standardised tests which rely on a single instrument to measure students’ overall ability. They encourage multiple assessment instruments to measure students’ abilities such as portfolios, learning logs, dialogue journals, observations, peer assessment, and self-assessment. Second, assessment is seen as a process which can be integrated during learning with teachers holding a prominent role in the assessment process. Lastly, assessment materials, as well as learning materials, are required to be authentic and contextual because they relate more to the use of language as a communication tool (Davison & Leung, 2009; Meskill, 2010).

Among the multitude of assessment types, it is dynamic assessment that is most likely to be compatible with the notion of formative assessment. As Lantolf and Poehner (2005, p. 1) stated, ‘FA (Formative Assessment) might be reconceptualized according to DA principles’. Both formative assessment and dynamic assessment reside in socio-constructivist learning
theory. The learning theory that embraces the concepts of peer learning, autonomy, the move from lock-step learning, and high challenge-high support can be found both in formative assessment and dynamic assessment. For example, Lantolf and Poehner (2004, p. 50), defined dynamic assessment as that which ‘…integrates assessment and instruction into a seamless, unified activity aimed at promoting learner development through appropriate forms of mediation that are sensitive to an individual’s (or in some cases a group’s) current abilities.’ This notion resembles formative assessment in which teaching and assessment are seen as having an interpenetrated connection (see Section 2.1). Further, the essence of formative assessment is in line with the twin constructs of teacher-mediated learning through intervention and interaction and the ZPD (Lantolf & Thorne, 2006; Poehner, 2016; Poehner & Lantolf, 2005; Shay, 2008). That is, making relevant teaching adjustments as a follow-up from the evidence gathered from assessment activities (Black & Wiliam, 1998b, 2009; Popham, 2006; Stiggins et al., 2006).

Another important concept that relates to formative assessment is scaffolding, a term which is often equated with teaching and support. However, Wilson and Devereux (2014) have argued that those terms are not interchangeable with scaffolding. Scaffolding is not just any teaching or any support because support may not help students to improve unless it encourages students to be autonomous learners. Drawing on Mariani’s (1997) work, Wilson and Devereux (2014) conclude that ‘scaffolding involves challenging students to make leaps forward into their ZPD. Intellectual challenge must be high, while explanations must be explicit in terms of what is expected, how to achieve it and why it is important.’ Further, Hammond and Gibbons (2005) recognized two types of scaffolding: designed-in and contingent scaffolding. Simply put, the former refers to a planned scaffolding that teachers design to accomplish the learning objective, while the latter refers to unplanned scaffolding that occurs throughout learning. Scaffolding is closely linked to formative assessment
because it shares the tenet of helping students to improve during learning. Contingent scaffolding (Hammond & Gibbons, 2005) is in line with Wiliam and Thompson’s (2008) argument that formative assessment may occur in several cycles: long cycles (four weeks to one year); medium cycles (one to four weeks); and short cycles (day by day, minute by minute).

Other types of assessment may not be completely in line with the notion of formative assessment. For example, classroom-based assessment which focuses on teacher-generated assessment may not always be formative. As Rea-Dickins (2007, p. 514) points out ‘an activity or elicitation procedure in itself is neutral. It is only in its implementation and the use to which the data that emerges from a given activity that there develops its formative or summative potential’. Hence, the interrelationship between formative assessment and types of assessment can be illustrated by the figure below:
A large and growing body of literature has investigated many assessment instruments which are commonly used in the implementation of formative assessment, as explained in Section 2.2 (for example, Azarnoosh, 2013; Baleghizadeh & Masoun, 2013; Basturkmen, 2001; Vacilotto & Cummings, 2007). However, these studies use the construct of the instrument per se, or the construct of one of the types of assessment, rather than the construct of formative assessment, and studies of formative assessment in ELT remain scant and unexplored.

In the section that follows, two categories of formative assessment studies in ELT will be presented. The first category comprises those that do not directly espouse language assessment as their goal. In this category, formative assessment is mentioned only as a possible benefit. Usually, formative assessment is mentioned as being of secondary potential
to be used as alternative assessment, classroom-based assessment, or teacher-based assessment. The studies do not commonly use formative assessment strategies, although several underpinning theories of formative assessment are applied. In other words, such types of assessment may be applied either with a formative or summative purpose. Studies in this category are referred to here as ‘studies which relate implicitly to formative assessment’.

The second category comprises studies that focus specifically on formative assessment as a construct. These studies focus on formative assessment per se, and not potential of various assessment types. They are referred to as ‘studies which relate explicitly to formative assessment’.

### 2.5.2.1 Studies which relate implicitly to formative assessment

In several studies, formative assessment is mentioned only as one component of the assessment types being discussed. For example, formative assessment is mentioned in relation to alternative assessment, performance-based assessment, or classroom-based assessment (Rea-Dickins, 2001, 2006; Rea-Dickins & Gardner, 2000). Although these studies do not directly apply the construct of formative assessment, they open the door towards the understanding of formative assessment.

In fact, these studies have shown that they share similar principles to formative assessment in other learning fields and have established firm ground for formative assessment in ELT. For example, Rea-Dickins and Gardner (2000) studied how teachers carry out classroom assessment in nine inner-city primary schools in EAL program in England. They revealed that classroom assessment provides an opportunity for both formative and summative purposes, and it is complex to differentiate between them. Another study in a similar context showed that it is the implementation and the use of assessment data that determine formative assessment (Rea-Dickins, 2001). Findings of these studies are in line with much of the
literature in formative assessment which emphasises that the essence of formative assessment depends on how the evidence of learning is being used (see Section 2.2 for details).

Furthermore, as in other learning fields there has not been any agreement on how formative assessment in ELT is defined and what constitutes its practices (Davison & Leung, 2009; Llosa, 2011; Stoynoff, 2012). Rea-Dickins (2001, p. 432) proposed the following definition, ‘action taken as a consequence of an assessment activity’. This definition is in line with much of the academic literature in formative assessment which emphasises teaching adjustment as a result of formative assessment. In addition, formative assessment in ELT also shares similar features with formative assessment in other learning fields. Gattulo (2000) characterizes formative assessment in ELT as follows: it is a continuous multistage process conducted on a daily basis; it gives immediate feedback; its purpose is to modify learning and improve teaching.

Another issue which resembles formative assessment in other fields are self-contradictory policies: the implementation of formative assessment is supported by government policy, yet at the same time standardised testing is also administered. Consequently, applying formative assessment can be challenging. Specifically Leung and Rea-Dickins (2007) list several barriers to using formative assessment in ELT:

1. The pressure to meet the requirements of national tests and to move learning forward;
2. The national tests limit learners from extending their learning;
3. Disagreement over what formative assessment is, consequently, teachers are unprepared to engage in the activity.
2.5.2.2 Studies of formative assessment in ELT

In this second category, the studies focus specifically on formative assessment. Meskill (2010, p. 204) characterized formative assessment in English language learning as ongoing, individual, additive, proving information to guide subsequent instruction, and as a primary source of evaluative input. Although studies of formative assessment in English language learning are relatively scarce, this relatively small body of literature has provided an insight into the issue. To date, the majority of studies in the area of ELT formative assessment concern a number of aspects such as the effectiveness of formative assessment for teaching and learning; the measurement of teachers’ formative assessment practices; the practice of formative assessment in writing skills; and the role of technology in ELT formative assessment.

One of the studies that focused on the effectiveness of formative assessment was conducted by Colby-Kelly and Turner (2007). This mixed-method study investigated whether teacher-student interactions, when used as formative feedback, have positive impacts on learning. The researchers categorised formative feedback into two components: the motivational component and the learning component. Based on data from curriculum document analysis, questionnaires, interviews, and classroom observation, the researchers concluded that those students who considered teachers’ formative feedback seriously could improve their speaking performance.

Another study which concerns the effectiveness of formative assessment was conducted by Wei (2015). Specifically, she focused her study on the contribution of formative assessment to enhancing teachers’ instruction. The study was carried out in higher education in Vietnam and the data collected included students’ course experience and classroom observations. The study showed that using classroom observations as an instrument for formative assessment
can enhance teachers’ performance provided that teachers understand what constitutes good teaching and are able to make meaningful feedback.

Measuring teachers’ understanding of formative assessment is also one of the topics explored in this learning area. For example, Dargusch (2010) focused her study on factors that shape teachers’ assessment practice. This case study, which involved two English teachers in a year 12 ELL setting in Queensland, revealed five factors that shape teachers’ formative assessment practice: policy in both local and school contexts; teachers’ experience and beliefs; assessment documentation; and formative assessment habits. In the same learning context, a study conducted by Jian (2014) focused on teachers’ understanding and use of formative assessment. This qualitative study involved 98 English teachers in secondary schools in Beijing, China. After analysing the data from questionnaires developed by Taras (2008), the researcher concluded that participants have a misconception of formative and summative assessment.

In a different setting, Gattulo’s (2000) study, using four teachers and 86 students, examined how teachers, in an EFL context in a primary school in Italy, carried out formative assessment and used the incidental formative assessment processes identified by Torrance & Pryor (1998) as the framework for the research. Their findings revealed that several formative actions such as questioning, correcting, and judging are more common than others.

Another study with a slightly different focus was conducted by Naghdipour (2016). Specifically, this study investigated the effectiveness of formative assessment at improving writing skills. The study involved 34 undergraduate students in Iran. Participants were taught using the process genre approach, and a combination of the process writing approach and genre approach within the framework of formative assessment strategies from Black and Wiliam (2009). Based on the data from pre-and post-writing tasks, questionnaires, and semi-
structured interviews, the researcher concluded that formative assessment is a potential means for improving several aspects of writing. Its success is determined by factors such as contextual affordances, the student cohort, the English proficiency of the students, and their motivational profiles.

One of the interesting features of ELT formative assessment studies is that, of the four language skills, most studies concern writing skills. For example, Lee (2011) focused on the impacts of teachers’ formative assessment practice. The study, involving 138 students and four English teachers in a secondary school in Hong Kong, used interviews, questionnaires, and classroom observations. The researcher concluded that formative assessment can significantly change teachers’ instruction and assessment practices. In addition, Lee (2011a) maintains that formative assessment improves students’ motivation in writing. In a further study on formative assessment in writing skills, Lee (2011b) revealed that there are some prerequisites that teachers need to fulfil so that formative assessment can promote teaching and learning: teachers need to believe in the need for change; teachers must be determined to innovate and be proactive in seeking ways to improve their practice; teachers are supported within the school system; and collaborative work among teachers is pivotal so that formative assessment can be an essential part of their teaching. Bruner (2016) also conducted a study in a similar learning area. His mixed-method study focused on teachers’ and students’ perceptions of using formative assessment in a writing course. The study involved 100 students and four English teachers in secondary schools in Norway. Drawing on data from interviews and questionnaires, the researcher concluded that teachers and students have different perceptions over how formative assessment implementation feedback, grades, text revision, self-assessment, and students’ involvement. The researcher suggests that there should be a common understanding between teachers and students in order to make the practice of formative assessment more meaningful.
The tendency to integrate educational technology to overcome difficulties in implementing formative assessment can also be found in EFL settings. The study conducted by Karagianni (2012) focused on computer assistance to facilitate the process of formative assessment and enhance students’ responsibility to monitor their learning progress. Drawing on the work of McKenna and Bull (2000) and McKenna (2001), Karagianni (2012, p. 255) defined Computer Assisted Assessment (CAA) as ‘the use of computers to deliver, mark and analyse student assignments or examinations and can cater for various types of assessment such as diagnostic, formative or summative assessment’. The study took place in Greece and involved 25 secondary school students. A range of research instruments were applied such as questionnaires on motivation and learning styles, software containing quizzes, and questionnaires of self-evaluation and perception towards the application of the software. The study showed that computers support formative assessment by providing continuous measurement of students’ learning progress. Further, CAA can benefit both teachers and students; it can encourage students to become active participants in learning, and it can provide feedback for students.

Nordrum, Evans, and Gustafson (2013) conducted a study on feedback in an EFL higher education setting in Sweden. Specifically, the study focused on comparing two types of feedback in writing: in-text commentary on writing techniques and rubric-articulated feedback which provided a brief paragraph of general comments. This study was designed using action research, and the data collected comprised reflective journals, questionnaires, and interviews. Findings revealed that each type of feedback has its own specific use; in-text commentary focused on lower-order issues of language proficiency, while rubric-articulated feedback dealt with higher-order concerns related to an overview of writing achievement.
To sum up, changes in the paradigms of learning theory and students’ needs to acquire communicative competence have led to reform in ELT assessment. Nevertheless, the reform has not yet brought formative assessment to the forefront in language assessment, despite its potential to improve learning; studies of formative assessment in ELT are relatively scarce and underexplored. In fact, some of the studies mentioned formative assessment only implicitly. Further, studies that use a formative assessment construct in ELT are limited to several topics such as its effectiveness, its practice for writing skills, and the role of educational technology in its practice.

Notably, formative assessment in ELT needs further research; in particular, how formative assessment strategies should be implemented in the field of ELT. Studies of formative assessment investigating how teachers use the interplay in daily classroom talk for formative purposes, or how teachers make use of the assessment which may be ‘fast and rarely tidy and neat’ (Meskill, 2010, p. 207) would improve teachers’ understanding of how formative assessment can be carried out in their own contexts. In the following section, I will present formative assessment in ELT, in Indonesia.

2.6 Formative assessment in Indonesia

The notion of formative assessment has been adopted in Indonesia for decades. However, its implementation has not been widely accepted by teachers in Indonesia (see Chapter One for details). Presumably, this is because the notion of formative assessment is contradictory to the teaching and learning culture in Indonesia which is still teacher-centred and dominated by rote learning (OECD, 2015; Zuhdi, 2015; Zulfikar, 2009). In addition, the hegemony of high-stakes standardised tests has shaped the testing culture which is entrenched in Indonesian educational systems. Nevertheless, despite the prevailing conditions, a small body of literature has shed light on formative assessment in Indonesia.
This section discusses the practice of formative assessment in the Indonesian context. The first part will describe the assessment system in Indonesia, followed by a description of how formative assessment is practiced in Indonesia.

2.6.1 Educational systems in Indonesia

Before proceeding to explain the assessment system in Indonesia, it is necessary to give an overview of several curriculums that have been applied in this country because essentially, the assessment system is determined by the applied curriculum. There have been eight curriculums applied in Indonesia since it gained independence in 1945. These are illustrated in the table below:

![Curriculum in Indonesia](image)

Figure 2.9 Curriculum in Indonesia
Adapted from Yulia (2014, p. 16)
The most notable change in the curriculum in Indonesia was brought about by KBK in 2004. As implied by its name, the curriculum is competency based and places learners in an active role, with the teachers acting as learning facilitators. This is in contrast to the previous curriculum, Kurikulum 1994, in which students were positioned as the recipients of knowledge, and teachers were the sole knowledge provider. The differences between Kurikulum 1994 and KBK are illustrated in the table below:

Table 2.3 Differences between Kurikulum 1994 and KBK

<table>
<thead>
<tr>
<th>Kurikulum 1994</th>
<th>KBK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centralist</td>
<td>Decentralist</td>
</tr>
<tr>
<td>Contains no standardized competence</td>
<td>Contains standardized competence</td>
</tr>
<tr>
<td>No activities to familiarize students with content and concepts</td>
<td>Integrated and programmed activities to make students familiar with content and concepts</td>
</tr>
<tr>
<td>No ICT (Information Communication and Technology)</td>
<td>Introduction of ICT (Information Communication and Technology)</td>
</tr>
<tr>
<td>Multiple-choice assessment</td>
<td>Classroom-based assessment</td>
</tr>
<tr>
<td>Thematic approach for grade 1 &amp; 2 primary school students (recommended only)</td>
<td>Thematic approach for grade 1 &amp; 2 students of primary school (compulsory)</td>
</tr>
<tr>
<td>No continuity of competence</td>
<td>Continuity of competence stratification from grades 1 to 12 (over school levels)</td>
</tr>
<tr>
<td>No curriculum diversification</td>
<td>Curriculum diversification: special and international curricula</td>
</tr>
<tr>
<td>Syllabus developed by the local education authority or school depending on needs</td>
<td>Gives opportunities to teachers, schools, and local authority for program elaboration and adaptation or analysis of materials</td>
</tr>
</tbody>
</table>

Source MoEC (2013, p.87)
Another significant change occurred when the government issued KTSP, a curriculum developed from KBK. Decentralisation was applied and schools were allowed to prepare their own education plan such as school curriculum, syllabus, and teaching materials (OECD, 2015); curriculum decentralisation was considered to be unsuccessful. There were a number of problems encountered in its implementation: many teachers still relied on textbooks prepared by publishing companies; most teaching was still teacher-centred; and teaching to the test was a common practice (Nababan, 2014; Nur & Madkur, 2014; OECD, 2015). In response, the government decided to issue another curriculum, entitled Kurikulum 2013, which is currently being implemented.

There are several characteristics that differentiate Kurikulum 2013 from KTSP. First of all, Kurikulum 2013 is a centralized curriculum designed by MoEC; and it is an integrated curriculum (Zuhdi, 2015). Other elements of the curriculum such as competency-based, student-centred, interactive teaching, and active learning are still implemented and have been strengthened (OECD, 2015; Zuhdi, 2015). With regard to assessment, KTSP and Kurikulum 2013 are similar in their purpose; principles, and administrator. One significant difference is Kurikulum 2013 introduced several concepts such as authentic assessment, behavioural assessment, new scoring systems, and so on. Further, it is important to note that students’ active involvement in assessment, such as peer assessment and self-assessment, was stated explicitly in Kurikulum 2013, whereas in the previous curriculum it was only implied.

Since its implementation, Kurikulum 2013 has received criticism from teachers, in particular. One such criticism was that teachers considered the assessment systems required by Kurikulum 2013 too complicated, burdensome, and time consuming (Kurniasih & Sani, 2016). In response to the criticism, the government decided to postpone its full
implementation and asked several schools to use KTSP while other schools continued using Kurikulum 2013.

In 2016, Kurikulum 2013 was re-issued with some revisions, and this curriculum is referred to as Kurikulum 2013 revised-version. The policy regarding assessment in Kurikulum 2013 revised-version, MoEC Decree Number 23 Year 2016, stipulates that there are three parties responsible for students’ assessment: teacher, school, and government; each of which has the following purpose:

1. assessment by teachers aims to monitor and evaluate process, progress, and improvement of students’ learning simultaneously;
2. assessment by the school aims to assess Standard Kompetensi Lulusan (Graduate Standard Competence, hereafter SKL) for each subject taught in the curriculum;
3. assessment by the government aims to assess SKL based on the national standard for some subjects in the curriculum.

According to MoEC Decree number 23 Year 2016, the assessment system in Indonesia is characterized as follows. In terms of principles, assessment should be: valid, objective, fair, comprehensive, open, holistic and simultaneous, systematic, criterion-based, and accountable (Peraturan Menteri Pendidikan nomor 23, 2016). Further, the decree stated that teachers are advised to use the following assessment techniques: tests, observations, assignments, and/or other techniques of assessment as appropriate. In other words, teachers in Indonesia are allowed to use any assessment technique that enables them to assess students’ ability accurately. In addition, the decree also mentioned that the assessment results should be used for learning improvement and teaching adjustment. The features of the assessment system stated in MoEC Decree Number 23 Year 2016 and MoEC Decree Number 20 Year 2007 are in line with the characteristics of formative assessment as mentioned by several studies described in the academic literature (for example, Black & Wiliam, 1998a, 1998b; Chappuis
& Chappuis, 2008; D. R. Sadler, 1989; Shepard, 2005); thus, it can be concluded that since 2006, Indonesia has adopted formative assessment.

2.6.2 Assessment systems in Indonesia

Assessment practice in Indonesia is characterized by the domination of testing culture (see Section 2.1 for further details of testing culture), specifically high-stakes testing which has long been entrenched in the assessment system in Indonesia. Historically, the original test was entitled Ujian Negara (State Examination). This national standardised testing was administered in the period 1965-1971. The test was replaced by Ujian Sekolah (School-based Examination) applied during the period 1972-1979. As implied by its name, this examination was school-based, but the government set forth the general guidelines for the test. During the period 1980-2002, Ebtanas (Evaluasi Tahap Belajar Nasional – the National Evaluation of Level of Learning) was administered as a substitute for Ujian Sekolah. Ebtanas was managed by the government, and it was replaced by Ujian Akhir Nasional (The National Final Examination). It was in this period that the government began to set the threshold score which had to be achieved by students as a requirement for graduation; however, this test had the shortest period of administration, that is from 2003 to 2004, before being replaced by Ujian Nasional (National Examination, hereafter UN). Previously, these national standardised tests have mostly employed a multiple-choice format; today, however, UN employs not only multiple-choice format but also essay questions.

Since the first year of its implementation, UN has received criticism, and a considerable amount of literature has reported negative effects. For example, Afrianto (2011) and Murtiana (2011) report that teaching to the test, anxiety, cheating and narrowed-down curriculum are some of the negative backwash effects from UN. Another negative impact from UN revealed from a study conducted by Sukyadi and Mardiani (2011). In a multiple
case study involving three secondary schools, categorised as low to high achiever schools based on the UN result in the last four years. Data from observations, interviews, questionnaires, and documents led the researchers to conclude that UN had a strong influence on overall classroom interaction; and had more negative effects on low achieving schools. This was confirmed in a more recent study conducted by Furaidah, Saukah, and Widiati (2015). The 11 teachers who participated in the study were divided into two cohorts: high achieving and low-achieving schools. The data collected from interviews and observations showed that low-achieving schools have more intense negative backwash effects than their higher-achieving counterparts.

In response to the criticism, a significant change was made by the government concerning the role of UN. The policy on the education standard was enacted through Law Number 13 Year 2015 which in detail stipulates that students’ graduation is now determined by USBN (Ujian Sekolah Berstandar Nasional -school examination based on national standard). However, the shifting role of UN has not yet provided positive impacts. For example, a study conducted by Saukah and Cahyono (2015) concluded that UN may remain high-stakes because it is taken into consideration when students are being selected for entry into the next level of education. As stated on the policy of the current assessment system, enacted through MoEC Decree Number 23 Year 2016, article 8 (Peraturan Menteri Pendidikan nomor 23, 2016), the role of UN is:

1. to map the quality of the program/unit education;
2. in the selection of students into the next education level; and
3. to provide guidance on the formulation and the provision of assistance to a unit of education in order to improve the quality of education.
In addition, a study conducted by Ginting and Saukah (2016) revealed that teachers were still unprepared to design their own writing test as part of the USBN, and it was recommended that teachers undergo training on how to design a quality test for school examination.

2.6.3 Overview of formative assessment studies in Indonesia

A large body of literature has highlighted the fact that high-stakes tests can become a barrier for the implementation of formative assessment (for example, Green, 2013; Popham, 2001; Rosemartin, 2013). Further, high-stakes testing may also decrease teachers’ ability to perform formative assessment, as revealed by Sulistyo’s (2009) study. Drawing on data from questionnaires of 24 Mts (Madrasah Tsanawiyah/Islamic junior high school) teachers, he found that all participants agreed that classroom-based formative assessment should have a greater role in the assessment system compared to the UN; however, participants prevaricated when they were challenged to take more responsibility for determining students’ competencies. In other words, despite teachers’ objection to the dominant role of UN, they still lack confidence to take on the role themselves. Sulistyo (2009, p. 36) suggested that teachers’ attitudes may be due to their ‘insufficient knowledge and competencies in assessing students’.

Another reason that makes implementing formative assessment challenging for teachers, in particular, is related to their beliefs in teaching and assessment. In Indonesia, a teacher may work until they reach 65 years old; this is especially so for teachers who are employed as civil servants. Considering that the educational system in Indonesia has favoured student-centred, active learning, and formative assessment for less than two decades, the currently available teachers are those who are deeply accustomed to using a teacher-centred approach, with an emphasis on rote learning. Consequently, formative assessment may be in conflict with their beliefs in teaching and assessment. Teachers’ beliefs are defined by Borg (2001, p.
186) as ‘a proposition which may be consciously held, is evaluative in that it is accepted as true by the individual, and is therefore imbued with emotive commitment; further, it serves as a guide to thought and behaviour’. Teachers’ beliefs may be formed by their knowledge and experience, and shape the way they teach (Thomas, 2013). In addition, as Johnson (1994) claimed, teachers’ beliefs determine how they process new knowledge about teaching and translate it into classroom practice. Resistance by teachers to novel concepts such as formative assessment may occur because they often have more profound effects on teachers’ pedagogical knowledge (Remesal, 2011).

According to Zuhdi (2015, p.149), bringing about change in Indonesia is complicated because ‘The conventional way of teaching has become culture within schools, with teachers continuing to practice his style from generation to generation’. Zuhdi (2015) also added that it is not only teachers who are not accustomed to a ‘new’ way of teaching, but also students; as a result, an ‘old’ way of teaching practice is preferable since changes are difficult both for teachers and students. Therefore, it is not surprising that teachers in Indonesia may still believe in a teacher-centred approach and emphasise summative assessment despite the government’s efforts to impose teacher-centred and formative assessment concepts. In a mixed-method study (Azis, 2014) which focused on teachers’ conceptions and practices at junior high school level, questionnaires (based on the teacher conception of assessment) were distributed to 107 teachers and semi-structured interviews conducted with 12 teachers. It was concluded that teachers were interested in formative assessment; however, in-depth analysis showed that their beliefs were closer to summative assessment. The researcher argued that this was because the applied assessment system gave little attention to classroom-based assessment.
In addition to UN and the prevailing culture that shapes teachers’ beliefs, implementing formative assessment in Indonesia can be challenging due to the lack of support from the environment, especially from the government (OECD, 2015; Zuhdi, 2015). For example, policies are sometimes contradictory; teachers are required to apply formative assessment yet, they are also required to prepare students to pass the UN. It is very unlikely that teachers will prioritise formative assessment when the consequence of failure in UN is high. Further, in the case of Kurikulum 2013, which introduced the notion of formative assessment, it was considered that the dissemination program was rolled out in a rush that left teachers with insufficient time to learn and understand the curriculum. As a result, a number of studies mentioned that teachers found implementing assessment, as required by Kurikulum 2013, to be challenging. For example, Retnawati, Hadi, and Nugraha (2016) reported that the 22 teachers involved in their study encountered difficulties in almost every stage of the assessment activities, such as formulating indicators and rubrics, and gathering evidence from multiple assessment instruments.

It is essential for teachers to understand the curriculum because they are the front liners who decide whether they want to change their practice or not. As stated in the OECD report on education in Indonesia (2015, p. 270): ‘Teachers need to feel a sense of ownership of the curriculum reform if it is to achieve its aims. It needs to be ensured that teachers know what the changes are and why they are important’. Further, the report also mentioned that in order to implement curriculum successfully in Indonesia, several requirements need to be fulfilled:

1. strengthen networks of practice (both face to face and virtual);
2. build professional and leadership capability;
3. engage in ongoing monitoring and evaluation;
4. align resources, and provide equitable student access to textbooks;
5. use a range of methods to focus teachers’ attention on things that make the biggest difference to student achievement (OECD, 2015, p. 271).
Nevertheless, despite all the challenges, formative assessment is applied by some teachers in Indonesia. Several studies document the benefits of formative assessment, for example, Nahadi, Firman, and Farina (2015) investigated the impacts of feedback in formative assessment on students’ learning activities in a chemistry course. This quasi-experimental study involved 32 students, and gathered data from tests, interviews, questionnaires and observation sheets. The findings revealed that formative feedback provides positive impacts in terms of enthusiasm, motivation, and participation. Another study on the benefits of formative assessment was conducted by Edyanto (2014) in physics. Specifically, the study focused on the effectiveness of web-based design for the purpose of formative assessment. The researcher argued that it is essential to have a web-based program so that formative assessment, which is often considered to be time consuming, can be conducted outside school hours. Having gone through a number of stages in developing the site, the researcher concluded that the web-based program could improve students’ conceptual understandings of the materials being taught. Another study on the benefit of formative assessment was conducted by Purnomo (2013). This experimental study involved 81 students in a mathematics II course. Drawing on data from documents, questionnaires and tests, the researchers concluded that formative assessment is more effective than traditional assessment in terms of students’ motivation and achievement.

Besides focusing on the advantages of formative assessment, one of the studies focused on instruments that can be applied to measure pre-service teachers’ assessment orientation (Listiani, 2014). The instrument, POSTT (Pedagogy of Science Teaching Test), was translated and adapted into Bahasa Indonesia. Having conducting a pilot study on 55 Indonesian biology teachers, the researcher concluded that POSTT is a valid and reliable instrument for examining pre-service science teachers’ teaching orientation. Although the
study was conducted in the higher education context, the researcher stated that it can be applied in other learning areas.

In the ELT field, one study of formative assessment was conducted by Widiastuti and Saukah (2017). The study focuses on the barriers and opportunities of implementing formative assessment in EFL in the junior high school setting. This qualitative, descriptive study involved only three teachers and three students, and used interviews as the source of data. The researchers found that teachers’ lack of formative assessment knowledge hampers their application of follow-up teaching activities. Although this study has given a valuable insight into formative assessment in EFL in Indonesia, it does not give a comprehensive description of teachers’ difficulties in implementing formative assessment due to small sample size and the reliance on interviews as the sole source of data.

Another study in the ELT field in Indonesia was conducted by Arrafii and Suhaili (2015). It was conducted in two secondary schools in Central Lombok, Indonesia, and four English teachers participated in the study. Drawing on data from interviews and classroom observations, the researchers concluded that teachers’ understanding of formative assessment needs to be improved; further, they suggested that teachers’ understanding did not correlate with their practice as they found one of the participants who practised ‘sophisticated’ formative assessment, but had a poor understanding of the assessment concept. Similar findings were revealed in a more recent study with a larger number of participants (Arrafii & Sumarni, 2018). The survey design explored teachers’ understanding (243 English teachers in Central Lombok) using an instrument entitled TFALTQ (Teacher Formative Assessment Literacy Questionnaire). Research revealed that teachers do not yet have comprehensive knowledge of formative assessment, and their assessment practice has the following features: ‘… serve the accountability purpose, examination/test driven learning, the procedural
approach to learning and assessment, and the receptive role of students’ (Arrafi & Sumarni, 2018, p.45). These two studies have provided important information about teachers’ competencies at performing formative assessment; however, they make no attempt to give concrete solutions on how to improve teachers’ formative assessment practice.

As can be seen from the paragraphs above, most formative assessment studies were conducted in the context of science subjects; to date there has been only one study on formative assessment in the social sciences, in EFL in Indonesia. In addition, most studies have focused on the benefits of formative assessment, as opposed to how formative assessment can be carried out, or how it would work in a secondary school setting.

To sum up, despite the vast area, education in Indonesia is now available to most students in Indonesia. The main concern now is to provide quality education. Efforts for improvement have been made in various ways, one of which is to revise the curriculum in line with modern approaches to learning. The current curriculum, Kurikulum 2013, should improve learning because it upholds the notion of active learning and strongly recommends formative assessment. Unfortunately, it also stipulates the high-stakes tests that hamper the aim of the curriculum itself. Despite various studies that document the negative effects of summative assessment, the government insists that the test be administered to students in grades 9 and 12. Since the test has a significant effect on students’ future lives, and on the reputation of schools, teachers focus more on preparing students to pass the test instead of focusing on learning improvement. In addition, the entrenched teaching-learning culture favours teacher-centred pedagogy. As a result, active learning and formative assessment are often neglected by teachers; both the prevailing culture and UN are barriers to formative assessment. Fortunately, some teachers are still applying formative assessment and studies have documented its benefits; however, these studies are mostly in the science area. To date, only
three studies in a secondary school setting in EFL have focused on teachers’ difficulties and teachers’ understanding in performing formative assessment. As a consequence, how, and to what extent, formative assessment is carried out in EFL study remains largely unknown.

2.7 Summary

To sum up, in this chapter I reviewed the existing literature on formative assessment. I have presented a discussion on how formative assessment is defined in this thesis and a thorough elaboration of formative assessment strategies. Challenges to implementing formative assessment, followed by possible solutions, recommended by several empirical studies, were also discussed. Furthermore, formative assessment in the ELT context, including a general description of language assessment, is also explained. This chapter also presents ELT formative assessment in the Indonesian context, that is the Indonesian education and assessment systems.

It is apparent from the literature that there is no agreed definition of what constitutes formative assessment, despite a number of definitions having been proposed. Nevertheless, there are several common elements present in the various definitions of formative assessment:

- **process**
  Formative assessment is not a one-off activity, rather it is a continuing process.

- **feedback**
  Teachers’ assessment should involve formative feedback which is evidence-based so that follow-up action can be taken.

- **students’ active involvement**
  Assessment is not the teachers’ sole responsibility: peers and learners should actively participate in the process.

- **teaching adjustment**
Evidence gathered from assessment should be used by teachers to adjust teaching; in fact, this is the core element that differentiates formative assessment from summative assessment.

Overall, these studies highlight the distinct feature of formative assessment; that is, using the evidence of learning to scaffold learners. Therefore, in this study formative assessment is defined as the process of assigning value towards students’ competencies, including feedback, student involvement, and subsequent adjustment to teaching.

Most of the studies reviewed here agree that formative assessment is not confined to a set of assessment instruments; however, a set of formative assessment strategies are available, and teachers can incorporate these into their own teaching contexts. Unfortunately, these strategies remain too broad to be translated to teachers’ classroom practice. Although it is important to note that several studies attempted to provide this information, most were conducted with researcher intervention, and some were only focused on providing guidelines on how formative assessment can be implemented.

Furthermore, a review of the literature also revealed that large class sizes and high-stakes testing are still considered to be a barrier to the implementation of formative assessment. However, a deeper analysis reveals that the substantial challenges to applying formative assessment may relate more to teachers’ assessment literacy and the nature of formative assessment itself, which has multiple interpretations.

Turning now to formative assessment in ELT, there is a relatively small body of literature in this area. To date, there is only one study concerned with the overall formative assessment practice in foreign language learning. Moreover, there has not been any study in the ELT area which discusses how formative assessment strategies can be implemented. Studies in formative assessment in ELT can be categorized into two: firstly, studies that do not refer to
formative assessment as a construct per se; rather, it is seen only as part of assessment types such as alternative assessment, classroom-based assessment, or performance-based assessment. Secondly, studies that used formative assessment as a research construct have only focused on measuring the effectiveness of formative assessment practice on students’ achievement, and evaluating teachers’ understanding of formative assessment. In addition, these studies were only carried out in the context of teaching one of the language macro skills despite the fact that teaching language is not conducted in this manner.

Taken together, the literature review in this chapter identifies research gaps in two areas: focus and context. Specifically, previous studies have not provided teacher-generated examples of how formative assessment can be carried out and have not explored the practice of formative assessment in an ELT context. Unfortunately, the same research gaps also prevail in the Indonesian setting. There are only limited studies of formative assessment in Indonesia, and those studies are mostly conducted outside the EFL context. To the best of my knowledge, only three studies of formative assessment have been conducted in the EFL field in Indonesia; however, the first two studies focused on evaluating teachers’ formative assessment practice (Arrafii & Suhandi, 2015; Arrafii & Sumarni, 2018), and the second, focused on the barriers to and opportunities for formative assessment practice in lower secondary schools. Hence, how and whether teachers practice formative assessment on a daily basis in EFL, and what can be learned from teachers’ practice in secondary high schools in Indonesia remains unknown.

Thus far, the aim of this thesis is to address the aforementioned research gap. Specifically, this current study seeks to answer the following research questions:

1. How do ELT senior high school teachers in Indonesia perceive formative assessment?
   a. What do they understand by the term formative assessment?
b. What do they believe are its advantages?

c. Ideally, how do they think formative assessment should be carried out?

2. How are teachers’ perceptions of formative assessment manifested in their practice of it? What instruments and strategies do teachers use to implement formative assessment?

3. What challenges are encountered by teachers when implementing formative assessment? How do they cope with these challenges?

The procedures and methods used in this investigation will be discussed in the following chapter.
Chapter 3
Research methodology

The previous chapter has identified the paucity of studies in formative assessment in ELT, especially in the Indonesian contexts. This study is pivotal considering the potential benefits shared by formative assessment practice in improving learning. This study aims to address the gap; specifically, it aims to provide examples how formative assessment in an ELT context can be carried out. This chapter presents the research methodology applied to seek answers to the proposed research questions which are:

1. How do ELT senior high school teachers in Indonesia perceive formative assessment?
   d. What do they understand by the term formative assessment?
   e. What do they believe are its advantages?
   f. Ideally, how do they think formative assessment should be carried out?
2. How are teachers’ perceptions of formative assessment manifested in their practice of it? What instruments and strategies do teachers use to implement formative assessment?
3. What challenges are encountered by teachers when implementing formative assessment? How do they cope with these challenges?

Specifically, this chapter would firstly discuss the rationale behind selecting multiple case study design for this study. Other topics presented in this chapter are how participants were recruited; reasons for employing semi-structured interviews, classroom observations, and formative assessment documents as data sources; ethical considerations and confidentiality; and how the data were analyzed. The last section discusses the summary of this chapter.
3.1 Research design

Relevant to the nature of the research questions and theory underpinning the study, case study is selected as the research design. A case is defined as a ‘bounded system’ (B. Johnson & Christensen, 2012). It may include an entity with a clear identity, such as a person, group, or organization; or it can also include an event, an activity, or a process (Creswell, 2005; Merriam, 1998; R. K. Yin, 2003). In an educational setting what constitutes a case may be related to a person such as a student, a teacher, or a principal, or a setting such as a classroom, school or workplace. An event, such as a program, an activity that is related to educational practice, can also be categorised as a case (Merriam, 1998; Scott & Morrison, 2007).

Case study was selected for the current research because it allows the researcher to obtain detailed information (Willis, 2007). It is in line with the aim of the research which is not only to identify teachers’ practices in conducting formative assessment but also the beliefs which underlie their practices. Further, case study is selected due to its strength in dealing with a variety of research instruments such as documents, interviews, observations and artefacts (R. K. Yin, 2003). This research employed the multiple (or collective) case study model, which is one of the three types of case studies mentioned by Stake (2006). In other words, this study involved more than one case and each case is considered to be part of the overall study.

Multiple case study design was chosen over a single case study design because teachers in Indonesia are teaching in diverse contexts. Examining how teachers conduct formative assessment in their various contexts would lead to better understand decisions other teachers might make in other contexts in terms of formative assessment.

Further, multiple case study offers deeper research results, as Johnson and Christensen (2012) elaborated, as studying more than one case provides the following benefits: cases can be
compared based on similarities and differences; it is effective to test a theory; and it generates richer and more reliable results than one case. Similarly, Freebody (2003) claimed that multiple cases can enhance the development of a solid theory. Stake (2005, p.446) emphasised that ‘it is believed that understanding them (the cases) will lead to a better understanding, and perhaps better theorizing, about a still larger collection of cases’.

To sum up, multiple case study design is selected to address the gap in the literature of formative assessment practice in ELT which is still underexplored. The design is considered to suit the aim of the study because it can allow useful, practical, and deeper insight into understanding and practice. Further, it also has the potential to give a comprehensive insight on how teachers face challenges in conducting formative assessment including the solutions that they implement.

3.2 Case selection

As suggested in the previous section, this study employed purposive sampling (Creswell, 2005). Participants were selected based on their relevant profiles to the purpose of the study as well as the theoretical framework underpinning the study (Flick, 2014; Nunan, 1992). Further, Stake (2006) expounded three criteria in selecting cases, they are: the cases should be relevant to the study; they should provide diversity throughout various contexts; and they should afford opportunities to learn about complexity and context.

For this reason, three criteria were applied in selecting the cases for the current study. First is that a diversity of teaching contexts was one of the selection criteria for this research; as Stake (2005) has said using heterogonous backgrounds in purposive sampling can produce lucid and insightful data. Further, information from participants would be able to be corroborated by having participants from diverse contexts (Bryman, 2012). To ensure that the
study represents various contexts, participants from MoEC or MoRA schools were also selected based on their different teaching context.

In addition, to gain information from settings in which formative assessment was most likely to occur, participants were selected based on their competence (Merriam, 1998; Silverman, 2010; Stake, 2006). Hence, the second criterion for selecting participants is teachers who were considered to be competent to provide a model for other teachers in the context of theory and practice. As mentioned in the Chapter 1, the notion of formative assessment was strengthened in the assessment systems in Indonesia through Kurikulum 2013. However, due to adverse criticism and disproportionate curriculum dissemination, the government decided to postpone its implementation for some schools; it is important to note that the curriculum was re-issued in 2016. In other words, during the period of 2014-2015 a limited number of schools implemented Kurikulum 2013 while some schools applied KTSP (see Figure 2.6 Curriculum in Indonesia for details).

The participants in this study were teachers whose schools were implementing Kurikulum 2013 during that particular period. Teachers whose schools are implementing Kurikulum 2013 were considered to have had the experience in using formative assessment because they had had training on how to implement formative assessment. This implies that teachers who implemented Kurikulum 2013 had at least a theoretical understanding of formative assessment, as well as experience in implementing it.

The last criterion in selecting participants was peer recommendation. In Indonesia every discipline is allocated to an academic community entitled Musyawarah Guru Mata Pelajaran (panel of secondary school subject teachers, hereafter MGMP) or Kelompok Kerja Madrasah (panel of secondary school subject for madrasah teachers, Hereafter KKM). It is a regional disciplined-based teacher association. In this community teachers consult each other about
teaching and learning difficulties, syllabus, disseminating government educational programs, writing tests for regional use, and so on. Through this interaction, teachers are able to identify those peers who are considered to be able to give good examples of how formative assessment is carried out.

Based on the criteria mentioned above, eight teachers were invited to be the participants of this study. Stake (2005) and Merriam (1998) pointed out that there is not any precise guideline on how many cases should be employed in multiple case study design; however, Eisenhardt and Graebner (2007) argued that researchers should not employ less than four cases in using multiple case study because it is not insightful and could not reveal the complexity of cases. In addition, Miles and Huberman (1994) claimed that more than 15 cases may become ‘unwieldy’ (p. 30). This research involved eight cases because this number would be sufficient to ‘expand and generalize theories’ in a process of literal and theoretical replication (Yin, 2003, p. 10).

The cases selected were divided into two cohorts: four MoEC teachers and four MoRA teachers. I applied two types of procedures in gaining access to these teachers. In some cases, I contacted the school principals; and I explained my request to conduct research in their schools. A brief description of the research was also elaborated. To support my explanation, I provided the copy of the Ethics Approval and support letter from my supervisory panel. The principal then appointed one of the teachers to be a participant in my research. This teacher agreed willingly to participate. In other cases, I made direct contact with the potential participant. I described the research that was going to be carried out and what would be required from the teacher as participant. Once teachers agreed to participate, permission from the principals was obtained.
3.2.1 Participants’ profiles

Based on the selection criteria elaborated in the previous section, eight teachers were selected to participate in the study. The summary of their profiles is presented in the following table, followed by their detailed profiles. Each participant is given a pseudonym to protect their privacy.

**Table 3.1 Participants’ profiles**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Educational background</th>
<th>Teaching experience</th>
<th>Participation in TPD program</th>
<th>School context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nelly</td>
<td>Doctoral’s degree in Education</td>
<td>36 years</td>
<td>TEFLIN Board, Chair of MGMP, has participated in several teachers’ professional development programs such as in ICT, curriculum, and teaching methods</td>
<td>It is a public school under MoEC management. It is a public school which each class consists of more or less 40 students</td>
</tr>
<tr>
<td>Adam</td>
<td>Bachelor’s degree in TESOL</td>
<td>23 years</td>
<td>Has participated in various teachers’ professional development programs such as in curriculum and teaching methods</td>
<td>The school is a public school under MoRA management. Each class consists of approximately 40 students</td>
</tr>
<tr>
<td>Wendy</td>
<td>Master’s degree in TESOL</td>
<td>28 years</td>
<td>Participated in various workshops and seminars. Presented several topics related to ICT and curriculum</td>
<td>The school is a public school under MoEC management. Classes are divided into accelerated class which consists of less than 15 students and regular class which consists of approximately 40 students</td>
</tr>
<tr>
<td>Alfred</td>
<td>Bachelor’s degree in TESOL</td>
<td>8 years</td>
<td>Has participated in workshops and seminars mostly in curriculum and teaching methods</td>
<td>It is a private school under MoRA management. Each class consists less than 35 students</td>
</tr>
<tr>
<td>Anne</td>
<td>Doctoral’s degree in Language Study</td>
<td>17 years</td>
<td>Participated in various workshops and seminars. Presented in various workshops and seminars, especially on topics related to classroom action research and literacy</td>
<td>It is a private school under MoEC management. Each class consists of less than 35 students</td>
</tr>
<tr>
<td>Name</td>
<td>Degree</td>
<td>Years</td>
<td>Experience</td>
<td>School Information</td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
<td>-------</td>
<td>------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Brenda</td>
<td>Master’s degree in TESOL</td>
<td>20</td>
<td>Participated in various workshops and seminars. Presented in various workshops and seminars especially on topics related to curriculum as she was one of the National instructors who was assigned to disseminate Kurikulum 2013.</td>
<td>It is a public school under the management of MoEC. Each class consists of approximately 40 students.</td>
</tr>
<tr>
<td>Evelyn</td>
<td>Master’s degree in TESOL</td>
<td>18</td>
<td>Participated in various workshops and seminars. Presented in various workshops and seminars especially for secondary schools managed by the Ministry of Education and Culture.</td>
<td>It is a public school under MoRA management. The maximum number of students in each class is 20 students. The school adopts moving class concept.</td>
</tr>
<tr>
<td>Daisy</td>
<td>Master’s degree in TESOL (on progress)</td>
<td>13</td>
<td>Participated in various workshops and seminars.</td>
<td>It is a public school under the MoRA management. Each class consists of approximately Teaching 40 students.</td>
</tr>
</tbody>
</table>

**Nelly**

Nelly is a high school teacher who has been teaching for 36 years. She started her teaching career in junior high school for two years, then she continued teaching in high school. She has a bachelor’s degree in English education, master’s degree in English Education, and completed her doctoral’s degree in Education program. She has participated in various Teacher Professional Development programs with various topics, and the last one is in Kurikulum 2013. She was once MGMP chair for English subject for two periods. Once, she was also one of the members of TEFLIN (Teaching English as Foreign Language in Indonesia) boards.

Nelly is teaching in a public school in the city which is managed by MoEC. The government appointed the school to be one of the pilot schools for the 2013 curriculum. As a consequence, trainings and workshops for other schools in the same region were often held by the school.
Nelly was a teacher of 12th grade students. There were 40 students in the class. These students were going to have their National Examination by the end of the school year. Nelly’s class reminded me of my school years with the teacher’s desk in front of the class and students were sitting in four rows. Each desk was occupied by two students. The class was decorated with some pictures. Most are posters related to class organization and school subjects.

Adam

Adam has taught since 1992. He holds a bachelor’s degree from one of the Islamic-based universities in Indonesia. His teaching career began in junior high school. During his early career, he taught for two different junior high schools, but, currently he focuses on teaching in one high school. He has taught in the school for 20 years. The school is managed by the Ministry of Religious affairs. At an early stage, the school was one of the few schools which implemented Curriculum 2013.

Adam has participated in several workshops and training courses concerning curriculum, such as CBSA, genre-based curriculum and Kurikulum 2013. Since it was an Islamic-based school boys and girls had to sit separately. They could not sit on the same desk nor the same row, and they didn’t usually work in the same group. There were 38 students in his class. The class itself was almost empty; there were no posters, LCD Projector nor air conditioning. Perhaps, it was because the building was relatively new. The school had just finished renovating the class, and the parking area was still under renovation.

Wendy

Wendy has taught English at the same school for 29 years. She holds a master’s degree in English Education. Wendy is particularly interested in Information and Communication
Technology (hereafter, ICT) in education. She was excited when she showed me how she integrated technology, and enabled her students to interact with other students from different parts of the world and made use of their English. She tried to use videoconference, writing blogs, using screencast-O-matic, and so on.

Her interest in ICT in education has led her to participate in various ICT related workshops and trainings. She was often asked to be the keynote speaker for implementing ICT in education for her colleagues in the same region.

Wendy has participated in various workshops and training courses on curriculum and other related topics. Wendy taught an accelerated class which accelerates students over the curriculum in two years rather than three. They are students selected by the schools according to various criteria and requirements. These students were outstanding in terms of achievement. In every school year, there is one accelerated class. Their academic excellence varies, they might be outstanding in language, science or social subjects. There were 13 students in the class. Different from other classes, these students were sitting individually. The class was filled with students’ projects in various subjects and equipped with LCD projector, sound systems, air conditioner, and television. Although there were only 13 students, they occupied a similar space to non-accelerated classes.

The classroom atmosphere itself was very different. The pace of learning was fast. The teacher would move from one activity to another. The teacher did not need to give detailed information of the task that she assigned. What is extraordinary from this class is their English competence; they were all almost fluent in using English. They did some codeswitching to Bahasa Indonesia, but spoke mostly in English.

Alfred
Alfred graduated from the English Education program in one of the Islamic-based universities in Indonesia five years ago. Before he began teaching in the high school, he had the experience of teaching junior high school students for extra-curricular courses. Alfred was teaching in a private Islamic-based school. The class that he taught consisted of 32 students. His class was the first lesson of the day; however, students often seemed to be less enthusiastic. The class was a bit dim; however, it was fully equipped by LCD projector, air-conditioning, and sound system.

Alfred had participated in various teacher’s professional development programs during his career. Some were organized by the school and some were self-initiated. A Kurikulum 2013 workshop was held by the school. The school invited a speaker who was well-recognized in the area.

**Anne**

Anne taught in a private school under the MoEC. Anne graduated two years ago with a doctoral degree from one of the universities in Indonesia specializing in language learning. Her career began in 1997. She has experience of teaching at various levels of education in Indonesia such as in kindergarten, primary school, junior high school, and university level. Her passion in improving her competence has led her to attend various seminars, conferences, and workshops in various parts of the world such as in America, Cambodia, and Australia. She often shares her experience in conducting action research with other teachers.

The class that Anne taught consisted of 35 students. Those students were mostly coming from advantaged families; they had access to learning resources and were able to fulfil any academic necessities such as accessing reliable internet services to support their studies, buying software to improve their writing skills, or joining student exchange programs to
English speaking countries. Some students had participated in homestay programs in English speaking countries held by the school during school holiday.

**Brenda**

Brenda taught at public school in a sub-urban area in Indonesia managed by MoEC. She holds a master’s degree in English Education. Brenda obliged herself to present in conferences and training courses at least three times a year. She was one of the national instructors for the dissemination program of Kurikulum 2013 of whom there were only 60 instructors throughout Indonesia. Recently, she was nominated as one of the outstanding teachers in Indonesia. In conducting her role as national instructor she was often assigned to disseminate the curriculum in various places in Indonesia.

The school where Brenda is teaching shared common features of many schools in Indonesia: a large number of students in the class and minimum facilities. The class was spacious; however, it was only equipped with an LCD projector. Thus she had to carry a sound system and laptop to class.

There were 45 students in one class. Every desk was occupied by two students. Students were sitting in four rows. There were some posters hanging on the wall. The posters were mostly figures of Indonesian national heroes, the president and vice president of Indonesia, and class organization.

**Evelyn**

Evelyn taught in a public school managed under the MoRA. At the beginning, it was a private school and was the first Islamic boarding school in Indonesia. Parents from around Indonesia send their children to this school due to its reputation in Islamic and science education. The
school changed its status from private to public school a few years ago, but it holds similar characteristics.

Teachers in this setting were also coming from different parts of Indonesia. They were tightly selected and the process of selection itself took almost a year. Evelyn had taught in this school for 19 years. Prior to teaching in the school she taught in a high school in another city. She has master’s degree in English Education. The school often hosted workshops or training courses for Islamic-based schools in other regions. Evelyn sometimes is a keynote speaker for such events.

Every class in this school consisted no more than 20 students. The school adopted a peripatetic system in which students move classes based on the subject that they learn. The class was equipped with stationery for every students, LCD projector, and dictionaries. Some posters of English tenses and students’ projects can be found on the wall. One of the most striking characteristics of this class was that every student attempted to speak in English, even when the teacher was not in the class, and also during classroom discussion with their peers.

Daisy

Daisy was an English teacher at one of the public Islamic-based schools in Indonesia. The school was located in the outskirts of the city. Daisy was in the process of achieving her master’s degree. She has taught for 7 years in the particular school. She likes to participate in various teachers’ professional development programs, but she had never presented in any conference or seminar.
She actively participates in the MGMP discussion which was commonly held in this school. Daisy taught in 10th grade and the class consisted of 40 students. Daisy stated that most of these students come from the disadvantaged families and lacked the support of their parents.

3.3 Data collection

Three sources of data were used to collect the data: interviews, observations, and documents. As Yin (2009) pointed out, in multiple case study the data attained should be confirmed by other sources of data. Freebody (2003) argued that multiple sources of data in a case study will enable researchers to: compare and contrast; increase project’s relevance; and discover findings which support or are against any hypotheses. In addition, Bryman (2012) points at the importance of data triangulation which is defined as ‘the use of two or more methods of data collection in the study of some aspect of human behaviour’ (Burns, 1997, p.324). Each source of data is elaborated below.

3.3.1 Interview

Interviews are defined by Berg and Lune (2012) as ‘(a) conversation with a purpose to gather information (p.105)’. Further, Gray (2014, p.378) characterizes interview as ‘a verbal exchange in which one person, the interviewer, attempts to acquire information from and gain an understanding of another person, the interviewee’. Participants’ beliefs, knowledge, past experience, feelings, and so on, can be identified through an in-depth interview. Consequently, using interviews as one of the sources of data for this research is essential because in depth information about the participants’ perception can best be acquired through interviews. Further, participants’ practices which may not occur during observation can also be discovered.

Among the three types of interviews mentioned by Creswell (2005), -structured, semi-structured and unstructured interviews, I applied semi-structured interviews or semi-
standardised interviews as Berg and Lune (2012) refer to them. As elaborated by Nunan (1992) semi-structured interviews allowed me to: have a certain degree of power and control during the interview; have flexibility in asking the sequence of questions in the interview process; get privileged access to other people’s practices and beliefs in order to gain insights.

An interview protocol was established prior to conducting interviews. The formulated research questions and theoretical clarification of formative assessment became the basis for developing questions for interviews (Brinkmann & Kvale, 2015). Eleven questions were designed for the interview. The first three questions focused on participants’ perceptions of formative assessment, including the definition, advantages/disadvantages, and the features of formative assessment. The next six questions concerned participants’ practices in conducting formative assessment in terms of techniques and feedback. Finally, challenges that participants might encounter was the last question on the interview protocol (See Appendix 6 for further details).

The interview questions were then piloted. Piloting the interview is essential because it gives the researcher an insight on how to ask, probe or clarify some questions on the interview protocol (Merriam, 1998). Further, it gave me experience on how to conduct interviews so that by the time I conduct the real interview I felt at ease. Pilot interviews were conducted with two participants who were having similar background to the participants of the research. Each participant was interviewed twice. The first interview focused on participants’ perception of formative assessment, practices they commonly used in formative assessment, and challenges that they encountered in implementing formative assessment. The second interview functioned as a follow-up of the first interview. Questions mainly focused on issues which I still considered to be unclear in the first interview. I also asked for some clarification on activities that happened during observations.
Since the interview was the first source of data that I collected, a consent letter was given prior to the interview. General information of the research was also provided. A brief explanation of formative assessment was conveyed as an aim to have common understanding about the topic. To ensure that the participants felt at ease during interviews, I asked participants to choose the place for the interview. All participants preferred to have the interview in a place where they can have privacy and remain undisturbed. Before the interview began, I asked the teacher for permission to record the interview. Each interview lasted for approximately 30 minutes. The interviews were conducted in Bahasa Indonesia; the aim was to build good rapport with participants and to enable the participants to put forward their opinion fluently.

3.3.2 Classroom observation

Besides interviews, this research also used observations as one of the sources of data. The aim was to gain an insight of how participants put their knowledge of formative assessment into practice. As Flick (2014) has said, it is essential to use observation because there might be a gap between attitude and behaviour. An observation protocol was established prior to conducting the observation. I adapted the observation framework from Oswalt (2013) in identifying formative assessment during learning (see Appendix 7 for further details). The framework consists of five points as follows:

a. Learning targets: clarifying intentions and sharing criteria for success
b. Monitoring: engineering effective classroom discussion, questions, and learning tasks that elicit evidence of learning
c. Feedback: providing feedback that moves learning forward
d. Self-assessment: activating students as the owners of their own learning
e. Peer assessment: activating students as instructional resources for one another
This framework aims to measure how much formative assessment occurs in the class. I adapted the framework and to identify how participants conducted formative assessment. Besides the five points mentioned above, other important points which were intriguing or seemingly related to formative assessment were also included. For example, I added pre-observation notes and post-observation notes. I also added my reflection and any questions that emerged during observation.

During observations I took the role as participant-observer (Flick, 2014; Merriam, 1998); thus, I was more as an observer than a participant. The observations were carried out in a naturalistic environment (B. Johnson & Christensen, 2012). This implies that I did not arrange the class environment; it occurred in the real world. I applied the stages in carrying out observations as mentioned by Merriam (1998), they are; entry, data collection, and exit. On the first observation, participants introduced me to their students. I always sat at the back; remained friendly but unobtrusive throughout observations. All observations were audio recorded. The number of observations for each participant varied, but they all fell within a ‘medium-cycle of formative assessment’ (Thompson & Wiliam, 2008), as they covered a period of 1 to 4 weeks within and between instructional units. As Johnson and Christensen (2012, p. 186) have said, observation is “the watching of behavioural patterns of people in certain situations to obtain information about the phenomena of interest”. Therefore, I stopped observing as patterns of how teacher conducted formative assessment were identified.

### 3.3.3 Documents

Documents were also one of the sources of data for this study. As Creswell (2005) said, documents are considered to be good a source of data to enhance researchers’ comprehension about the emerging phenomenon for qualitative study because they are ready to use. I collected documents during the observation processes. Some documents were requested from
the participants, and some were voluntarily offered. The collected documents include: teachers’ assessment instruments, lesson plans, samples of students’ work, and sample of teachers’ feedback. Those documents were copied and scanned. The original documents were returned to participants.

3.4 Ethical considerations and Confidentiality

Prior to collecting the data in the classroom, I met the principals, and explained the research project in person. As mentioned in the previous section, there were two procedures in gaining access to the classroom. First, I contacted potential schools, and I met the principals and explained the research project. Second, I contacted the teacher who met the criteria that I had made before. I gave general information about the project. As soon as the teacher agreed to participate, s/he took me to the school’s principal; afterwards, I explained the research project. Besides explaining the general information about the research project, the principals were also informed that they could withdraw at any stage of the research. Further, I would conceal their schools’ identity in writing up the thesis. If any teaching materials uncovered the school’s identity, it would be hidden. In the process of collecting the data I tried not to disrupt the school’s agenda.

Likewise, the teachers who participated in the research project were given comprehensive information about their involvement in the data collection. First, I gave them the general information about the nature and the purpose of the research in written form (see Appendix 3). Afterwards, I explained and answered some questions as these teachers tried to clarify what was written in the general information. It was made clear that their involvement in the research project was voluntary, and they could withdraw from the project at any time. A consent letter was given for teachers to sign after I ascertained that they had understood their involvement in the project. In the process of collecting the data, I tried to make the teacher feel comfortable during interviews by letting them choose the place to conduct interview.
Similarly, during the observation process I ensured that I did not disturb the teaching and learning activities.

Before I began the observation, the teacher introduced me to the students. Similarly, I explained my presence and the general information of the research project in the language style which is comprehensible based on the students’ age. I distributed the consent letters to students as they agreed to participate in the research.

Further, I also informed that their anonymity and confidentiality were guaranteed through the use of pseudonyms for participants and schools. Any identifying information from the documents such as test, teachers’ lesson plan, teachers’ record books, etc. was removed. Soft copies of files of observation, interview, and documents are stored in password protected computer. Only members of my supervisory panel and I can access the data. Participants had the access to read the interview transcript, and field notes if they wished. I also informed participants that they would be free to veto any information that they do not wish to be made public. Further, participants were offered the opportunity to view the final thesis.

3.5 Data analysis

Simply put, data analysis is the process of making meaning from collected data. Gay, Mills, and Airasian (2014, p. 223) defined qualitative data analysis as ‘a process of breaking down data into a smaller units, determining their import, and putting the pertinent units together in a more general, interpreted form’. The data analysis process in this research was adapted from Creswell (2005)
Figure 3.1 The process of data analysis
Adapted from Creswell (2005, p. 230)

The data analysis began as the data was being collected. Afterwards, I transcribed the data from the recording of interviews, observations and field notes (Johnson & Christensen, 2012). Once the transcripts were ready, the following stage was to read through all of the data again to get a general sense of the information and to engage with the data.

The initial reading, would led on the process of category construction (Merriam, 1998; Travers, 2001). According to Travers (2001), coding is the most important stage in organizing data into a set of themes. A code is ‘a word, phrase, abbreviations or symbol which is applied to a segment of text’ (2009, p. 176). The coding phase involved several
stages, namely open coding, axial coding, and selective coding. In the open coding stage, data was read and notes were taken at the end of the first reading of important, interesting, or significant data. The second stage is axial coding. Similar words, phrases, events were allocated into the same category. A categorisation is ‘classification of ideas or concepts: when concepts in the data are examined and compared to one another and connections are made’ (Gay et al., 2014). This stage also involved making connections between the data and the theoretical frameworks of the study. Categories were re-examined and ensured that they were linked. The last stage of coding was selective coding. At this point, categories were refined and allocated into themes. Since interviews were carried out in Bahasa Indonesia, Data from the interviews was analysed in Bahasa Indonesia, and only the displayed data which was translated to English. Data from observations and documents are (mostly) in English; thus, it is analysed in English.

Throughout the coding process I did not use the available software for managing qualitative data. I used an Excel spreadsheet in analysing the data since it provides flexibility in managing data, in comparison with using manual ‘cut and paste’ method or using software such as NVivo. In addition, I believe that it is the researcher’s capabilities that hold the prominent role in analysing data. As Yin (2009) made clear that general analytic competence cannot be substituted by any computer-assisted tool.

The last stage in data analysis is interpreting the data; Gay and Airasian (2003, p. 245) defined interpreting as ‘the reflective, integrative, and explanatory aspect of dealing with a study’s data’. This process was basically making connections and identifying patterns, common aspects and linkages among categories that emerged from data. The process of transferring the collected data into display data is iterative and simultaneous (Creswell, 2005). It was iterative because I needed to refer back and forth from the data collection and
Having being immersed in the data, eventually five different categories emerged; namely participants’ background, learning, perception, practice and challenges.

3.6 Summary

This chapter provides detailed information of how the current research was carried out and the rationale for this process. In line with the aims of the study and the proposed research questions, the study is designed using multiple case studies. Eight participant teachers were recruited based on the following criteria: had the knowledge and experience in implementing formative assessment and were recommended by peers. To gain a comprehensive vignette of how these teachers conduct formative assessment, semi-structured interviews, classroom observation, and classroom assessment artefacts were collected and used as the source of data for the study. Permission was secured from the University of Canberra Ethics Committee prior to collecting the data and all data has been handled confidentially and anonymously. The collected data was analysed using Creswell’s (2005) framework for data analysis (see figure 3.1); the results of the analysis data are presented in the following chapter.
Chapter 4
Findings

The previous chapter presented the research methodology used in this study to answer the following research questions:

1. How do ELT senior high school teachers in Indonesia perceive formative assessment?
   a. What do they understand by the term formative assessment?
   b. What do they believe are its advantages?
   c. Ideally, how do they think formative assessment should be carried out?

2. How are teachers’ perceptions of formative assessment manifested in their practice of it? What instruments and strategies do teachers use to implement formative assessment?

3. What challenges are encountered by teachers when implementing formative assessment? How do they cope with these challenges?

This research employed a multiple case study design. Four teachers whose schools are under MoRA’s management, and four teachers whose schools under MoEC’ management, participated in the current study. Data were collected from three instruments, namely semi-structured interviews, classroom observations and artefacts.

A thorough analysis of the collected data showed that the participants have diffuse understandings in defining formative assessment; nevertheless, the participants had a common understanding of the key elements of formative assessment and its potential benefits. In fact, these participants emphasised that the key element of formative assessment is using the assessment results to adjust teaching and eventually enhance the overall learning. Furthermore, findings from the study show that these participants made adjustments to their
teaching based on the results of assessment and applied several recommended strategies when implementing formative assessment. This resulted in a number of excellent examples of how formative assessment can be carried out, including how to overcome the emerging challenges in formative assessment such as large class sizes and limited time.

This chapter presents a detailed representation of the research findings. It is divided into three sections based on the proposed research questions. The first section presents an analysis of teachers’ understandings of formative assessment, based on the data from the semi-structured interviews. The second section elaborates a detailed analysis of how formative assessment strategies were applied by the participants. This is followed by a discussion of the challenges encountered by the participants during the implementation of formative assessment, and also how participants overcame these challenges. Data from semi-structured interviews, classroom observations, and artefacts formed the basis for the second and the third sections. I conclude the chapter with a summary.

4.1. Understanding formative assessment

The first set of questions aimed to find out how the participants perceive formative assessment; this constitutes the way formative assessment is defined and/or its key elements, as well as what the participants believe to be its potential benefits. Data from this section is mainly derived from semi-structured interviews which were conducted twice for each participant: prior to and after conducting classroom observations. Data from formative assessment artefacts were also used to support the discussion in this section. The findings of the study are summarized in the following table, and then further discussed.
Table 4.1 Participants’ understanding of formative assessment

<table>
<thead>
<tr>
<th>Participants</th>
<th>Definition</th>
<th>Elements of formative assessment</th>
<th>Benefit(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nelly</td>
<td>Formative assessment is a quiz at the end of chapter</td>
<td>Integrative, continuous, feedback, teaching adjustment</td>
<td>• Identify students’ needs</td>
</tr>
<tr>
<td>Adam</td>
<td>Daily score/assessment</td>
<td>Integrative, continuous, feedback, teaching adjustment, simple</td>
<td>• Identify students’ needs • Monitor students’ progress • Provide feedback for teachers</td>
</tr>
<tr>
<td>Wendy</td>
<td>Formative assessment can be conducted during and/or at the end of learning</td>
<td>Integrative, continuous, feedback, teaching adjustment</td>
<td>• Identify students’ needs • Monitor students’ progress • Feedback for teachers</td>
</tr>
<tr>
<td>Alfred</td>
<td>Formative assessment is assessment which is conducted during learning and is repetitive</td>
<td>Integrative, continuous, feedback, teaching adjustment</td>
<td>• Enhance students’ competence • Monitor students • Identify students’ ability • Identify students’ learning needs</td>
</tr>
<tr>
<td>Anne</td>
<td>Formative assessment is a type of assessment which is conducted during learning that records students’ learning progress and behaviour</td>
<td>Integrative, continuous, feedback, teaching adjustment</td>
<td>• Identify learning points that need to be improved or revised • Provide feedback for students, teacher, and schools</td>
</tr>
<tr>
<td>Brenda</td>
<td>Formative assessment is formal and informal assessment conducted during learning</td>
<td>Integrative, continuous, feedback, teaching adjustment, students’ involvement</td>
<td>• Encourage self-assessment • Provide tools to describe students’ learning progress • Enhance the institution support for students • Identify teaching areas that need to be improved</td>
</tr>
</tbody>
</table>
| Evelyn  | Formative assessment takes various forms such as quizzes, assignments or projects. Formative assessment can be conducted during and/or at the end of learning. Formative assessment can be conducted by teachers and peers. Formative assessment (includes) instruments to identify students’ comprehension. | Integrative, continuous, feedback, teaching adjustment, students’ involvement. | - Inform teachers about instruction and assessment instruments.  
- Enable students to identify their learning mastery. |
| -------- | ----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------- |-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Daisy   | Formative assessment is a test at the end of each chapter.                                                                                                                                                                                                         | Integrative, continuous, feedback, teaching adjustment, simple.                                                                                                                                                                                                  | - Identify students’ own comprehension and difficulties in learning.  
- Is feedback for teaching.  
- Enable students to measure their own competence. |
It can be seen from the data presented in Table 4.1 that participants had different understandings of how formative assessment is defined. Although all participants agreed that formative assessment is conducted during learning, they had a diverse range of perceptions of what formative assessment is, especially in terms of timing. For example, three out of eight participants considered formative assessment to be a quiz which is administered at the end of the chapter or unit of a lesson. As Nelly said in the first interview:

So in theory, after every few classes, right, or after a topic, there needs to be formative assessment. As for that… Oh, formative assessment is usually for every chapter. But after a topic there’s an exam or a test. (Nelly/Int1/[06:06])

In comparison, other participants did not limit formative assessment to tests at the end of unit of a lesson or chapter; they had broader perceptions of formative assessment. For example, Brenda referred to it as an assessment which includes ‘any formal and/or informal assessment instruments’. Brenda clarified what she meant by ‘formal’ and ‘informal’. She referred to tests which are administered at the end of a chapter or unit, the middle or end of a semester, as formal; and she referred to any assessment instrument other than tests such as teachers’ observations, portfolio, or journal, as informal. Similarly, Evelyn described formative assessment as any form of assessment instrument applied by teachers to measure students’ competence; in addition, she added student involvement as one of the features of formative assessment.

Interestingly, although Alfred agreed that formative assessment is an activity measuring students’ learning competence which is conducted during learning and involves various assessment instruments, he excluded the mid-term test and final test from his definition of formative assessment. Further, he emphasized the idea that formative assessment relates to a repetitive process of assessment, as he stated in the interview:

Formative assessment is a form of assessment to measure students’ ability, but not, what’s known as UTS UAS (mid or end of semester exam) isn’t formative
assessment. Formative assessment is assessment that is done during the students’ learning process. (Alfred/Int1/[06:34])

In addition to defining what formative assessment is, participants were also asked about its key elements. Surprisingly, although the participants defined formative assessment differently, they seemed to agree on its key elements. There were some suggestions of what elements should be present during the process of formative assessment. First, most participants agreed that formative assessment should involve process; it is not a one-shot assessment which only concerns the end product of students’ learning. In the interview, Wendy argued that assessment aims to classify students based on their performance, and formative assessment should assist teachers to gain reliable data. In doing so, on-going assessment is essential or else teachers may come up with inaccurate assessment results of their students’ achievement. Similarly, Brenda and Anne considered the process as a crucial element of formative assessment; further, they added that, if teachers were only focused on students’ end product, they might derive false conclusions about students’ competence. In the interview, Alfred described what he meant by the formative process in his class:

    I observe the students’ processes while working. For example, structure. They analyse and draw conclusions. Within that, I see they study the most important processes, understand them, and they repeat the processes. (Alfred/Int2/[18:51])

Regarding process as a crucial element of formative assessment, one of the implications is that formative assessment should involve various types of assessment instruments. Wendy commented that applying various assessment instruments would enable teachers to identify students’ progress because students’ learning progress changes from time to time; while Daisy argued that applying various instruments would ensure the representativeness of her assessment.

Besides process, the majority of participants mentioned that formative assessment should be integrative in terms of learning domain and language skills. Two out of eight respondents
suggested that teachers should focus not only on the cognitive domain, but also on the affective and psychomotor domains. What they meant by affective domain included students’ motivation and attitude during learning; however, participants did not give further explanation on points that they assessed in terms of psychomotor domain.

With regard to assessing the extent of students’ language skills, all participants were of similar opinion that it was very difficult to conduct discrete point language assessment as formative assessment. They argued that teachers need tangible evidence of learning, and some skills can only be assessed when they are supported by other skills. In the interview, Brenda explained how she needed to integrate the language skills:

> We can’t suddenly assess speaking. The process emerges from reading or writing first. So, writing and speaking are final products, and they would definitely have come from reading or listening first. It’s indirect, I have done it like that; it wasn’t successful. (Brenda/Int1/[42:50])

Likewise, Alfred also agreed that formative assessment involved integrative skills; further, he provided examples of how he assessed several language skills in one activity. He entitled the activity, ‘I Can Speak’ (ICS), and it is described below:

> Individually scheduled, every class must nominate the best speaker, and they will convey a message or give their opinion in front of everyone. The other students will listen, taking notes. Their writing shows that they were listening. So the notes are collected and we assess them. (Alfred/Int1/[16:43])

He further explained that from the activity, there were three language skills that could be assessed: speaking skills, that is, students’ ability to talk in public; listening skills, demonstrated by the students’ ability to understand the message from their peers who spoke; and writing skills, taken from their note taking.
One of the interesting findings in this study was that two out of eight participants mentioned simplicity as a crucial element of formative assessment. They further explained what they meant by simple; that is, it should be practical and meaningful instead of complicated in terms of administering and scoring. Their opinion of this element may relate to their understanding that formative assessment is confined to a set of assessment instruments.

Table 4.1 also shows teachers’ perceptions of the main benefits of formative assessment. All participants agreed that assessment was about measuring students’ ability and that formative assessment is central to improving students’ learning. Prior to elaborating the benefits of formative assessment, Brenda explained that formative assessment functioned as a lens to identify students’ learning progress. As she explained in the interview:

So it’s like this, teachers have to monitor and follow students’ learning progress. They’re teachers, right, their job is to provide learning experiences, and see how far those learning experiences have been internalized by students. So, it’s formative assessment that helps teachers get the full picture. Because we can’t get the picture without a strategy. (Brenda/Int1/[12:43])

Overall, all participants considered formative assessment as being beneficial, not only for students, but also for teachers and schools. As can be seen from Table 4.1, all participants agreed that formative assessment can improve learning because it informs learning progress for both teachers and students. As Evelyn said in the interview:

Yes, of course, for both us and students. It’s two ways. If for example, maybe from our observations, if we see that perhaps children aren’t motivated, we’d of course ask, is the presented material uninteresting, monotonous or what? That is also a type of assessment for us. Or in written assessment, if the children’s marks are bad, it should be several points above KKM, meaning it’s our problem. Maybe it was difficult to read, ambiguous or so on. On the other hand, if everyone does really well, it means the questions were too easy, unable to differentiate the good from the bad. As for the children themselves, it’s like that too, right, they can assess how much they’ve absorbed. It’s important, it can’t be left out of the KBM process. (Evelyn/Int1/[26:53]).

Note: KBM (Kegiatan Belajar Mengajar, teaching and learning activities), KKM (Kriteria Ketuntasan Mimimuns, minimum completed criteria)
Likewise, Alfred shared a similar opinion that formative assessment benefits both teachers and students. Further, he added that in order to get the benefit from formative assessment, the activity should be conducted repetitively. By doing so, students’ comprehension towards learning materials would be enhanced. As he said in the interview:

Very much so, first we can see how much the children have understood the materials given to them, and also what their difficulties are, how to solve the problem, and it’s also a form of reflection for us teachers. For students, it’s also important to be able to measure themselves. Because of that I already remind them that the aim of language isn’t purely theoretical, but to communicate. So, it’s not purely chasing grades, but learning in a broader sense so you can independently understand. (Alfred/I1/[07.11])

Nevertheless, although all participants agreed that formative assessment will benefit students because it informs students about their learning progress, none of the participants explained how students can make use of the information to improve their learning. This finding strongly implies that these participants believed that students would automatically feel motivated to close their learning gaps when they received the information about their learning progress, or the information itself would provide learning direction for the students.

In comparison, the participants delineated their arguments on how formative assessment would benefit teaching and overall learning. In the interviews, the participants claimed that formative assessment informs teachers on the learning area in which they need to improve. In other words, formative assessment provides information so that teachers are able to make suitable adjustments to their teaching. Daisy considered formative assessment to be a reflection of her instruction; it gives a description of how well her students have understood the learning materials. Specifically, Wendy explained that formative assessment is beneficial for identifying whether or not students are making progress and in which areas they need further assistance.

The issue of making teaching adjustments was also raised by Brenda, in whose opinion, there should be some follow up action after a teacher has conducted assessment. If the results of
formative assessment are not satisfying, the teacher should adjust their teaching methods, instructional language, or teaching realia. Brenda also believes that formative assessment benefits the school and that it provides an opportunity for schools to improve their services.

As she said in the interview:

Actually, really, it has more than two functions. For students, formative assessment lets them know how well they’ve mastered a particular competence. For teachers, when they see that the results aren’t actually very good, adjustments will be made. Adjustments to improve assistance – for this, perhaps making changes to methods, language, or resources. While for the school, if teachers undertake formative assessment, then it can be reported to the school…resulting in institutions taking action to help students. (Brenda/I1/[13:19])

To sum up, the teachers’ definitions of formative assessment varied, ranging from a test at the end of chapter or unit of lesson, to any formal and informal assessment. Although all participants agreed that formative assessment is conducted during learning and excluded several compulsory tests, they had a diverse range of perceptions of what formative assessment is, especially in terms of timing. Further, the participants shared similar perceptions of the key elements and the benefits of formative assessment, namely the need for feedback, the continuous nature of formative assessment, the need for teaching adjustment to be made, and the integration of language-skills. In addition, all participants agreed that formative assessment has the potential to improve learning because it provides information about learning progress so that teaching adjustment can be made accordingly.

4.2 Implementing formative assessment strategies

Essentially, the existence of formative assessment can be seen from two perspectives: most importantly whether or not the evidence from assessment is used improve learning and the application of formative strategies. This study has shown that formative assessment opportunities occurred throughout learning (Rea-Dickins, 2006); and this confirms the interpenetrated connection between formative assessment and teaching (see Chapter 2 for details). Furthermore, this study showed that the participants applied the recommended
formative assessment strategies (Thompson & Wiliam, 2008; Wiliam & Leahy, 2015), and used them to improve learning.

One of the most crucial strategies of formative assessment is to provide feedback to the students, and it should be given as an ‘information communicated to the learner that is intended to modify his or her thinking or behaviour for the purpose of improving’ (Shute, 2008, p. 154) (see Section 2.2.3.2 for details). Findings from the study showed that the participants used the evidence of learning to adjust their teaching in order to improve the students’ learning. As mentioned in the previous section, the participants considered teaching formative assessment as an essential because it provides information about the areas of learning that need to be improved (see Section 4.1 for details). This finding is confirmed by the data from classroom observations. For example, in one of the public speaking activities in Alfred’s class, Alfred explained how to ask the presenter a question. This arose because according to Alfred’s assessment during the activity, none of the students had managed to correctly ask the presenter a question. He did not explain the same materials in the following session as he considered that students had made improvements in asking questions, during the public speaking activity.

This section discusses the strategies used by participants when conducting formative assessment. The results are generated mainly from interviews and observations. Data from documents is also used as one of the data sources in this section. The framework for this section is divided based on the underpinning theory of implementing formative assessment strategies, namely: clarifying, sharing, and understanding learning intentions and success criteria; engineering effective discussion, tasks, and activities that elicit evidence of learning; providing feedback that moves learning forward; activating students as learning resources for
one another and as owners of their own learning (Leahy, Lyon, Thompson, & Wiliam, 2005; Thompson & Wiliam, 2008; Wiliam & Leahy, 2015). The diagram below serves as a framework for formative assessment strategies discussed in this section.

**Figure 4.1 Formative assessment strategies**

### 4.2.1 Clarifying, sharing, and understanding learning intentions and success criteria

Learning intentions are defined by Wiliam and Leahy (2015, p. 31) as ‘… the things that we want our student to learn’ and they describe success criteria as ‘… the criteria that we use to judge whether the learning activities in which we engaged our students were successful’ (Wiliam & Leahy, 2015, p. 31). Findings from the study showed that all participants applied this strategy in their formative assessment practice. This strategy was applied using several techniques, namely by stating learning agenda, by establishing assessment criteria and/rubrics, and by giving examples. This section presents a detailed description of how the
participants applied the strategy, and it is divided into two sections: clarifying, sharing, and understanding learning intentions; and clarifying, sharing, and understanding success criteria.

4.2.1.1 Clarifying, sharing, and understanding learning intentions

Findings from the study showed that most of the participants informed students about their learning intentions for the class. Findings from the classroom observation showed that all participants announced learning intentions when they began the lesson; some even restated them at the end of the lesson as a way of summarizing the lesson and checking students’ understanding of the materials from the particular session. As can be seen from the observation data in Brenda’s class:

Brenda: Before we end this session, what have you learned today?
Student: news item
Brenda: What is the function of news item?
Student: Eh …
Brenda: What is the function of news items?
Student: to
Brenda: to what? Inform …
Student: To inform (unclear voice)
Brenda: What are the characteristics of spoken news?
Student: There is hello
Brenda: There is greetings. How about written news?
Student: No greetings
(Brenda/ O1/P6)

Interestingly, some of the participants informed students about learning intentions at the beginning of the school year. These participants informed the students of several assignments that they would have to complete during a semester or school year. Evelyn is one of the participants who applied this strategy. She said in the interview:

So, since the beginning (of a school year or semester) we (Evelyn and the students) have already had an agreement. For example, on the recount topic there are several assignments that they need to submit … they (the students) know what their assignments are, the topics (of the lesson), and the due date for the assignment.
(Evelyn/Int 2/[07:37])

Other participants only informed the students what they were going to study, as can be seen in Adam’s class:
Adam: Today we are going to learn passive voice. What is passive voice? (Adam/O4/P1)

4.2.1.2 Clarifying, sharing, and understanding success criteria

The following criteria of success were used by the participants: establishing success criteria/rubrics and giving examples. Alfred was one of the participants who established assessment criteria for a public speaking activity in his class. In this activity, individual students were asked to present a topic to the class. After the presentation, the audience (the student’s peers) was given an opportunity to ask questions. In the interview, he said that students were assessed based on the following criteria for public speaking:

I use assessment criteria, I assess three things: the first is appearance. For appearance, I look at how they are dressed, they need to be well-dressed at the front of the class, neat. Then their expression needs to match what they are saying. Then their gestures also play a role. So that’s appearance. The second component is verbal, which relates to the choice of words they use to convey the message, together with sentence structure, whether they’re creative in their use of sentences. The third component is their vocals. They have to be able to, you know, play around with their delivery, not be monotonous in their explanations, that would make us sleepy, so they need to play around with the vocals. But there’s a final one which has to be done. They need to break the ice. (Alfred/Int1/[11:58])

Data from classroom observation confirmed that Alfred shared and clarified not only the learning intention, but also the assessment criteria, as captured in the observation

Alfred: Here I will tell you some points that you need to avoid when you are in front. I will conduct an evaluation, I don’t mean to criticize, but the aim is for everyone’s improvement right? The first, I scored from this session is visual points… (Alfred/O1)

Although the majority of the participants established assessment criteria, deeper analysis of the data revealed that some of the participants kept the assessment criteria broad. For example, data from the documents (see Appendix 8) showed that Evelyn assessed students’ essays based on the following criteria: grammar, content, organization, and punctuation.

Data from the observation in Evelyn’s class confirmed that Evelyn used and explained the criteria to assess students’ essays; however, she did not clarify how each element would be
measured. Thus, it might not give further information on the quality of work or learning objectives that students should achieve.

Surprisingly, participants who did not clarify the success criteria reasoned that they purposefully made their assessment criteria broad. They said this enabled them to assess their students’ overall performances using their teaching experience and without detailed assessment criteria. Further, Daisy added that the most important thing in learning language is nurturing students’ confidence; and she thought that establishing detailed criteria might disrupt the process of encouraging students to use the language. Another surprising finding from the study showed that some of the participants included elements not related to language learning as part of their assessment criteria. For example, some participants included effort, appearance, handwriting, and students’ engagement as criteria in assessing students’ performance.

An example of the interesting techniques used by the participants for sharing and clarifying learning intentions was used by Brenda. She presented the learning objectives taken from her lesson plan (see Appendix 9). On the first observation, Brenda began a new topic, entitled News Items. She explained the learning objectives documented in her lesson plan and stated that by the end of the lesson the students should be able to achieve the first three learning objectives. In the following session, she re-presented the same slides of learning objectives. She informed the students which learning objectives they should have achieved in the last session, and which learning objectives they should achieve in the present session.

In addition to informing students about learning objectives and establishing assessment criteria and/or rubrics, the research data also revealed that the participants gave examples with the aim of clarifying their learning intention. Data from the observations also revealed that in general participants provided examples prior to giving assignments to their students.
For example, Adam used to give answers to the first two practice questions that he assigned to students. Other participants provided examples which were taken from the previous school year. For example, on the first observation, Wendy assigned students to make a poster of an event as a project. Before she asked students to do the project, she firstly gave an example of a poster of an event which she took from the previous school year. In addition, Wendy explained that she had a collection of students’ works in a blog. The blog could be accessed by other students from any school year, as she said in the interview:

So, each year I create a blog containing all of the students’ assignments. Then we can enter the assignments in here (showing the blog). These are the children’s assignments so others can access them in the future. (Wendy/Int1/[26:21])

Similarly, Brenda provided examples of previous students’ portfolios to the students. She divided the portfolio into three varieties; one of which was to give examples for other students which she referred to as an open access portfolio. She elaborated the portfolio as follows:

…for open access, for example if there were a unit on news items, they would write an article… The article would be kept in the library. The purpose is, if, for example I need an example. For example, last year there was something… and now there’s new material, in order to make pamphlets, and cards, for friends of Save the Earth. Explanatory material, and how floods happen. Then they’ll make a card and that will also be made open access. Then, if students want, they can take them. (Brenda/Int1/[35:25])

Examples were not only taken from the previous school year, but participants also made use of one of the students’ work as an example. For instance, Alfred showed the result of one of his students’ work in free writing (see Appendix 11 for the activity details). Free writing was a classroom activity implemented by Alfred. In the activity, Alfred gave two keywords, gave their definitions and examples of how each word is used in a sentence. On the third observation, Alfred gave an example of what constituted quality free writing, as described in the observation below:
Alfred: Before we start free writing, I will show you a good example of free writing. Helen showed that she understands what the words mean, the example is written in red. It shows that Helen understands (how the words are used in the sentence) (Alfred/O3/P5)

In addition, besides presenting successful learning outcomes, Alfred also showed examples that needed revision. Together with the students, he corrected the examples, as captured by the observation data:

Alfred: Ok, everybody please pay attention. So, I will give you correction if you make mistake, (I) underlined. There will be new words or sentence that you need to add. This is the example, and I put it in bold. Let’s see the example ‘contohnya’ my phone is gone, I turned my house upside down but I cannot find it, looking for it otherwise’ Here Andrew wrote can chatting. Because there is an auxiliary ‘can’ then it should be infinitive, so I gave this sign (Showing the sign)’ I can chat …finally I find it on my desk’ this is an example. Another example is from Norton, he is very good. Let us have a look at your friend’s essay. I show you this because it is a good example.(Alfred/O2/P5)

The findings from the study show that all participants stated the learning intentions to students in various ways, mostly they were stated at the beginning of the lesson, but some restated them at the end of the lesson. In addition, there were some participants who stated the learning intentions at the beginning of a school year or a semester. However, further analysis showed that these learning intentions were not clarified by most of the participants; they only stated what they were going to teach in that particular session. Fortunately, these participants clarified their intentions for the students’ learning with the use of success criteria such as assessment criteria or rubrics, and concrete examples. However, it is important to note that in some cases, rubrics and assessment criteria may not have helped students to understand learning objectives. They can be too broad or not related to the notion of language learning as has been discussed in the previous section.
4.2.2. Engineering effective discussions, tasks and activities that elicit evidence of learning

In order for formative assessment to improve student’s learning, it must intervene while learning is still in progress. The interventions conducted by teachers on the basis of formative assessment in the form of feedback or an adjusted approach will align students’ needs and teaching goals. Thus, improving learning entails a ‘past-to-present perspective of learning’ (Heritage, 2012, p. 180) which is captured and recorded as evidence of learning. This data is important because it is based on the evidence that teachers are able to make inferences about students’ learning and adapt their teaching when necessary; two important formative mechanisms of formative assessment (Bennett, 2011). Therefore, eliciting evidence of learning is central to formative assessment because it serves as the basis for learning improvement; in other words, it is important for teachers to select relevant assessment instruments in order to gain valid data for enhancement of their students’ learning.

Since formative assessment is not related to any particular instrument, assessment instruments used by teachers when performing formative assessment can be varied. This section discusses the instruments implemented by participants when seeking evidence of students’ learning. Firstly, I give an overview of participants’ main concerns when designing formative assessment. The section is then divided into two sub-sections, common and distinct instruments applied by all participants, and their individual practices.

4.2.2.1 Designing effective instruments for formative assessment

Comments from individuals concerning what formative assessment instruments should look like were varied. Two out of eight participants who claimed that formative assessment should be simple put priority on practicality in designing assessment instruments for formative assessment; the instruments needed to be easy to construct, administer, score and interpret. Other participants pointed out that teachers should put priority on validity rather
than practicality. For example, Anne emphasized that teachers should have a clear idea about what they intend to measure including the desired outcome. She said in the interview:

So firstly, it has to fit in with the syllabus and framework. Then it also needs to be clear. Then its validity, whether it is valid or not, reliable or not… The students should also know the purpose of the assessment, it must be well-written, the purpose must be clear, our output needs to be stated as clearly as possible.

(Anne/Int1/[09.57])

Similarly, Brenda pointed out that she could not administer formative assessment instruments merely based on practicality. According to her, it was unfair if teachers were only concerned with practicality when they designed assessment instruments. They should be designed based on the learning indicators that had been established previously. In the interview she explained how she designed her assessment instrument:

It depends right? It depends on what skills are being assessed. If the skill is speaking, then the focus will be on how I can gauge that. If it’s through a rubric, it will be based on the indicators for speaking skills. It’s different for writing. So the priority is to appraise what I’m measuring; for speaking the measurement is speaking. So, I can’t just choose assessment because it’s practical, because it isn’t fair for students.

(Brenda/Int2/[26:12])

Further, Brenda also mentioned that her assessment instruments must enable both her and the students to learn: students learn the materials and she learns whether the materials taught to students are comprehensible or not (Brenda/Int2/[27:22]).

Another factor mentioned by some participants was innovation. Data from the interview revealed that some participants considered creativity to be an essential element when designing assessment instruments. According to Wendy, teachers should design assessment instruments which are interesting to students, or else students would get bored and will not give their best performance (Wendy/Int2/[12:00]). In a similar way, Alfred claimed that he needs to keep making innovations to the assessment instruments. As he said in the interview:
Each year has a new format. Free writing is a new scheme. And short stories are also a new scheme. Because I feel it’s a shame if good students just sit, listen, and write and that’s it. I need to assess them and encourage their abilities.

(Alfred/Int2/[05:16])

4.2.2.2 Implementing formative assessment

Data from the study revealed that the instruments used by the participants to implement formative assessment can be divided into three categories: those applied by all participants, those applied by some participants, and distinct instruments applied by only one of the participants. This finding is interesting because to a certain extent it confirms theories that describe formative assessment as being contextually bounded; that is, teachers need to incorporate the practice in line with their setting. Yet, there are some instruments applied by all participants. Further, the data showed that in several cases, although some participants applied the same instrument, they applied it differently. The instruments applied by the participants are summarized in the table below:

<table>
<thead>
<tr>
<th>Instruments applied by all participants</th>
<th>Instruments applied by some participants</th>
<th>Distinct instruments applied by only one of the participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Tests</td>
<td>- Portfolio</td>
<td>- Three-line writing</td>
</tr>
<tr>
<td>- Observations</td>
<td>- Reading journal</td>
<td>- Short story</td>
</tr>
<tr>
<td>- Questioning</td>
<td>- Early diagnostic,</td>
<td>- Sampling</td>
</tr>
<tr>
<td></td>
<td>- Reading journal</td>
<td>- Video conference</td>
</tr>
<tr>
<td></td>
<td>- Peer assessment</td>
<td>- Team-blogging</td>
</tr>
<tr>
<td></td>
<td>- self-Assessment</td>
<td>- Speaking assessment by checking attendance list</td>
</tr>
</tbody>
</table>

4.2.2.2.1 Instruments applied by all participants

As shown in Table 4.2, there are three formative assessment instruments applied by all participants, namely tests, questioning, and observations. Data from the interviews strongly indicates that the participants considered tests as one the most crucial instruments for
assessing students. This may be the result of the myriad tests which are administered to students throughout a learning period.

Further, the UN has also promoted the popularity of tests (see Table 2.4 for details). Since English is one of the subjects included in the UN, the tests affect the school’s overall program and teachers’ instruction in general. As a result, it influences the formative assessment practice of all EFL teachers in Indonesia, especially those who taught the 12th grade students.

Besides test preparation for the UN, data from the observations also revealed that UN has promoted tests, or as Popham (2001) called it: ‘teaching to the tests’. For example, in every session, Nelly gave UN test samples taken from the previous year, or she specifically gave a set of tests which she thought were going to be on the UN. In a similar way, some participants often mentioned UN in class, for instance Brenda and Anne both mentioned UN as they began their lesson. Their aim may be to motivate students and familiarize them with the tests, as shown from the observation data:

Brenda: Ok, good. The question in UN would like this, for example, what is the communicative purpose of news item? Or what is the social function, It may sound like that. There are two possibilities social function atau (or) communicative purpose. In the syllabus it is stated social function .... (Brenda/O1/P5)

Similarly, the data from the classroom observation shows that Anne discussed Standard Kompetensi Lulusan (Graduate standard competencies, hereafter SKL) before her lesson. SKL are the competencies which are supposed to be measured by UN. Data from the observation implied that students were all familiar with SKL:

Anne: I will read and explain about SKL, do you know SKL?
Students: Yeesss....
Anne: What is it?
Students: Standar Kelulusan... (graduate standard)
Anne: Ya standar kelulusan ...(graduate standard)
Students: Standar Kompetensi Lulusan ...(Graduate Competency Standard)
Anne: Ya... ya... ya...Ok, this a from goverment that you
should have this qualification, why I asked you here because you have to prepare for the examination and one of the topics is about news, it means that you can make use of the facility to prepare yourself for examination

(Anne/O3/P1)

One of the striking findings from the data is how positively the participants regarded the UN and tests in general. For example, the participants considered the national examination central to learning because it helps to determine students’ competencies. As Brenda said in the interview:

Because we are constrained by SKL (the graduate competency standard), our tool to measure success is their ability to manage the graduate competencies. For example, the students need to understand several types of texts, their functions, and their purposes. It’s explicitly stated. Because of the orientation on graduate competencies, we automatically steer toward it. It doesn’t mean teachers aren’t creative, but adjust to the requirements of the competency. UN (National Examinations) are the government’s way of measuring whether someone is competent or not.

(Brenda/I2/(04:55))

In addition to the UN, the other compulsory tests that impact the practice of formative assessment are UAS (Ujian Akhir Semester - end of semester examination) and UTS (Ujian Tengah Semester - mid-semester examination). UAS is an examination at the school level; that is students, in the same grade in the same school, will have the same test. Commonly, teachers in the same subject in one school assign which teacher writes the final test for which grade. These teachers have the freedom to write their final test including its form and content. However, findings from the study showed that schools under MoRA’s management did not write their own UAS test; rather, the UAS tests were written by a group of teachers in Kelompok Kerja Madrasah (Madrasah (teachers) working group). This means several schools in the same region have the same UAS test. In the interview, Daisy explained the process of writing the UAS test as follows:

It could be that I make it, because we have a Madrasa working group, and this school is its chair. So, this KKM supervises hundreds of private high schools too, which are then divided. So, the teachers below MGMP share the load. So, the job of writing the test is shared. Some teachers will design the framework, and others will write the test. Then after the test is ready it will be returned to those that created the
framework. Its suitability is assessed, then revised, and when it’s finalised it will be sent off to the printers. So, the tests for everyone in a KKM are the same, UTS (mid semester exams) come down to individual teachers. (Daisy/Int1/[09:20]).

The participants also put strong emphasis on UAS. For example, Brenda considered UAS to be the end point of learning that measures whether or not students have achieved the learning indicators stated in the curriculum. As Brenda said in the interview:

As for the result of the end of semester exam, whatever the relationship with formative assessment, because of the system of mastery learning, the continuity of materials is gauged by the end of semester exam. The end of semester exam plays an important role to assess how these long processes end up… If those students that have carefully followed the process while studying, and they have learned, at least in the end of semester exam they will be able to reconstruct their learning, for example in a news item, the introduction will be there, the ending… at least they’ll have the structure. (Brenda/I2/[30:32])

Adam’s concern with UAS is reflected in his assessment practice. Adam whose school is under MoRA’s management, thinks that his students may not do well on the UAS test because the test is written by other teachers. Consequently, he administered a test prior to the final test. The purpose of administering a pre-final exam test was to mitigate students’ poor performance in the final examination. As he said in the interview:

Usually when I grade, my grade, the grade usually comes from what I’ve taught the children. Usually a week beforehand I’ll revise what I’ve taught. That’s what I usually do, but of course not all teachers are like that. So, for me the results of the UAS (end of semester exam) are not the benchmark. (Adam/Int2/[10:15])

In comparison to the UAS, UTS is written by the class teacher. Interestingly, although UTS is a compulsory test, Brenda considered the test should not be administered because she already had more reliable data. In the interview she argued:

Actually, the mid-semester exam could be removed, because we already have extensive data comparative to the mid-semester exam. In the mid-semester exam three basic competencies are formally assessed, but it can be a burden on students, they can get nervous, right, because they know it’s a test. But it’s more real in everyday activities. So I think the mid-semester exam should be removed, because the extensive everyday assessment is already a substitute. I believe assessment shouldn’t be too rigid, the assessment process is very important. So formative assessment should actually help me learn whether the students are succeeding or not. (Brenda/I2/[26:37])
Besides UN, UAS, and UTS which are compulsory, the participants also administered smaller tests or quizzes to students. Commonly, the test was administered at the end of a unit or lesson. Interestingly, although these tests are optional, all participants seemed to consider them compulsory. For example, Evelyn uses a range of assessment instruments but she always includes tests as part of her assessment practice. Another example is Alfred, who administers two types of tests to the students: Ulangan Harian (daily quizzes) and Quiz. He explained the differences between the two tests as follows:

A quiz is for structural materials, while UH (daily quizzes) assess whether students can do things yet, whether they understand things yet. A quiz is specifically for grammar, UH (daily quizzes) are for comprehension in its entirety. (Alfred/Int2/[07:48])

Furthermore, data from the study revealed that in some cases the participants have made use of these compulsory and optional tests to improve learning. For example, Brenda and Nelly usually administer tests at the beginning of a school year or semester. They use the results of the test to categorize students so that it will be easier for them to provide instruction in line with the students’ capacities. In addition, data from the classroom observation showed that summative tests could also be used to encourage peer learning. This was evident in Nelly’s class where she re-administered the summative tests from UTS. Although Nelly did not assign students to work in groups or pairs, the students collaborated to find the best answers for the questions on the test; in other words, tests stimulated peer learning (See Section 4.2.2.2.2 for further details).

Interestingly, despite the myriad tests that participants administered for students, some participants claimed that they dislike tests. Anne, for instance, preferred to assess students based on observation or on the products of learning, instead of tests (Anne/Int1/[25.11]). Likewise, Wendy also claimed that she disliked administering tests because writing tests is as time consuming as marking them. In addition, data from the observations revealed that tests could have negative backwash effects on students, such as causing anxiety. In the
observations; anxiety occurred throughout the test administration. For example, in Nelly’s class students pretended that the teacher had not informed them about the test, as shown from the observation:

Nelly: Students, today we are going to have test…
Student 1: Really?
Student 2: You didn’t warn us
(Nelly/O1/P1)

Other students looked busy trying to remember information that they thought was going to be on the test. Data from the observations showed that a similar situation also occurred in the other participants’ classes. Wendy, who taught in an accelerated class, gave a test after a sequence of activities which focused on reading. The students had to present the text to their peers; decide on a title for the text, and make a conclusion. Surprisingly, students looked anxious as Wendy administered the test, despite the fact that the previous activities had been more challenging.

The data suggest that both the participants and their students disliked tests. The exception was Brenda. In her opinion, tests could be a potential means for learning, providing that the teachers have the capability to construct quality tests. She also mentioned that tests can measure higher level of thinking and students’ ability to think critically. However, teachers need to remember to test only what they have taught. As she said in the interview

If teachers can develop quality tests, then the test will be a good tool. The problem is, can teachers really make tests? Teachers could make a great test, for example a test that assesses critical thinking, but if they didn’t teach in the same way then it wouldn’t work. Teachers have to guide students step by step and the test and instruction should be aligned. (P6/Int1/[25.30])

Another formative assessment instrument which was applied by all participants is questioning. A number of studies have indicated that questioning has the potential to improve learning because it can provide immediate feedback (Glasgow & Hicks, 2009; Shen & Yodkhumlue, 2012; Wong, 2010). Further, questioning has the flexibility to measure both
lower and higher order thinking; thus, it is not surprising that it is one of the most common assessment instruments applied by teachers.

Data from the interviews revealed that none of the participants mentioned questioning as one of their instruments in conducting formative assessment. This is a surprising result considering that data from the observations showed all participants applied questioning in their instruction; moreover, questioning was applied for multiple purposes in learning. For example, the participants used echoic questions (Long & Sato, 1983) to check students’ comprehension. Several expressions such as: ‘Do you understand?’ or ‘Is it clear?’ were very common in all of the participants’ classrooms.

Further, the participants also used questioning to assess students’ prior knowledge before they began new materials. For example, before Brenda started to explain news items, she used questioning to elicit students’ previous knowledge:

Brenda: My question is, what is news?
Student: News, things, like stuff.
Brenda: News, item?
Student: stuff, like things or articles
Brenda: So, new things (laughing)
Student: No, article
(Brenda/O1/P3)

Questioning was also used to guide the process of students’ thinking. This technique was commonly used by most participants, especially when they were teaching grammar points. For example, when Alfred was explaining the notion of a correct sentence, he used questioning as follows:

Alfred: Do you think it’s a good sentence? So, what is the subject here?
Student: The sky.
Alfred: The sky. What about verb?
Student: Was
Alfred: Can you find more?
Student: Was
Alfred: What else?
Student: Was
Alfred: Okay, good
Participants also used questioning to move learning forward. The participants sometimes
used questioning to extend students’ knowledge about the language simply for and to
improve their students’ vocabulary. For example, when Anne was teaching reading skills,
one of the students asked her about a particular word in the text, and she tried to enhance
learning in the following way:

Student: Ms, the word ‘potato’ is used with ‘e’ instead of ‘o’,
what is the difference? If the Irish… eh Ireland
Anne: Come on, Irish or Ireland? Which one?
Student: Ireland is the country right…? (continued discussing the
difference between Irish and Ireland)

(Anne/O1)

In addition, data from the study showed that the participants used questioning to elicit
information regarding students’ knowledge, and the same question was used for several
students. These participants sometimes sought alternative answers, as can be seen from the
observation data below:

Wendy: Now I show you another one, what is it?
Student: It is one of the competition to design …
Wendy: How about you Beth? Is it the same thing? Or are they different?
What are the similarities and differences?
Student: The same thing is about competition
Wendy: What do you think Alma, which one is better? Why? (Wendy/O1)

Observation is another instrument applied by all participants to elicit evidence of learning.

Data from the interviews revealed that all participants monitored their students throughout the
lesson, especially when they were doing language practice. Previously, observation was one
of the compulsory instruments implemented in the assessment process, as explicitly stated in
the current curriculum. However, the curriculum does not specifically mention what factors
teachers should emphasise nor does it explain just how the teachers should perform the
observations. As a result, the way participants conduct observations varies. Data from the
interviews revealed that the majority of the participants focused their attention on low and high achieving students. As Evelyn said in the interview:

I usually focus on those that are smart or those that are struggling. For example, is A passive or active today. I can’t watch all the students at the same time.
(Evelyn/I2/[05:06])

Further, data from the interviews showed that there were differing opinions on the necessity of using instruments such as checklists of when teachers observed students. Some participants also claimed that observing or monitoring had been encapsulated in their teaching practice; thus, they did not need to use specific instruments to observe the students. In contrast, other participants thought that it was important to use instruments to help them focus on particular students. Interestingly, regardless of whether or not these participants used observation sheets, they tended to focus their attention on the affective domain during the observation process. Students’ motivation, attitude and/or behaviour are several aspects that participants’ paid attention to during observation. As Anne stated in the interview:

So, every meeting I watch their disposition, their motivation, their attention, their consistency during reading. I focus more on students’ behaviour. (Anne/I1/[20:50])

However, there seemed to be a contradictory finding from the observation data; it showed that the participants were actually focused on students’ comprehension of the materials that they taught rather than on the affective domain. For example, Adam re-taught the materials when he saw that students seemed to have difficulty in doing the task that he had assigned. These findings imply that there is a tendency for participants to interpret students’ competence through their behaviour.

4.2.2.2.2 Formative assessment applied by some participants

As shown in Table 4.2 there are several formative assessment instruments commonly applied by the participants, namely diagnostic tests, portfolios, self-assessment, and peer assessment.

In this section, I focused on diagnostic tests and portfolios while self-assessment is discussed in Section 4.2.5 and peer assessment in Section 4.2.4
As mentioned before, the Indonesian assessment system is filled with tests, and some of the participants found ways to use these for formative purposes. Diagnostic tests were administered at the beginning of a school year or a semester with the aim of identifying the students’ capacities prior to teaching. Findings from the interviews revealed that four out of eight participants applied diagnostic tests. For example, Nelly would put marks on her attendance list as she was able to identify students’ abilities in the first few sessions (Nelly/Int1[7:31]). In comparison, Wendy and Brenda said that they commonly tested the students by asking them to write about themselves in the first session of a school year. In the interview, Wendy stated her reason as follows:

The first task is always “Write about yourself” in the first posting about “My identity”, so I say “It’s me, it’s all about me.” From that I can understand a student’s writing ability, I can also understand their character traits, their background and so on. (Wendy/Int1[64:22])

Further, Brenda pointed out that although it was actually a test to identify students’ ability, she tried to disguise it so that students did not feel anxious as they did the test. In the interview, she explained how she conducted the diagnostic tests, and how she used them formatively.

At the beginning of semester, we get information from a test that doesn’t appear to be a test, but assesses students. For example, from the question “Why don’t you like English or why do you like English?” They’ll answer both verbally and in written form. From the verbal response, we can conclude “Oh their verbal competence reflects this level”. Through a rubric, for example, if there’s hesitation it shows which group they’re in. Diagnostics at the beginning will inform future directions. Writing assessment can also be done through prompting questions that they don’t actually feel are a test, but it can become data. For example, “Tell me about your family”, and then the students write. Their competence can be seen through the flow of the writing. Then all kinds of competencies can be seen in writing, and that will inform future directions. So from the beginning I choose groups. So there are three groups. They’ve been marked. In the attendance list they are marked based on their group with a moon or a star (Brenda./Int1[15:09])

Besides diagnostic tests, another formative assessment instrument applied by some participants is portfolio. Similar to the implementation of diagnostic tests, there were some
differences in terms of media and content in each application of portfolio. For example, Anne mentioned that she preferred online portfolios in which students could include materials either they or the teacher had chosen. In contrast, Daisy preferred hard-copy portfolio which is submitted at the end of school year. Daisy further explained she assigned students learning materials that they should put in their portfolio. She explained the detail of the portfolio content as she said in the interview below:

…. For example, they make a summary adding in previous tasks. Making a summary of lessons, for example three chapters together with a reflection on what was studied, together with where the difficulties lay, then what they’ve done, what the minimal effort was. “What’s your suggestion to your teacher”. But unfortunately it’s only for several children because of constraints. (Daisy/Int1/[14:11])

In comparison, Brenda divided the portfolio that she applied in her class into three categories. First is the students’ portfolio in which they include their best work over a certain period of time. Second is open access portfolio which is a collection of the student’s works which is considered to be excellent. In line with its name, open access portfolio is put in the library so that it is accessible for every student. The last portfolio is the teacher’s own portfolio. Commonly, Brenda scans or takes a photo of students’ work which she considers to be either very poor or excellent, before she put marks on the assignment.

Another instrument which was applied by two of the eight participants is the reading journal. In Anne’s case, she assigned students to make a list of the new vocabulary that they learned, and also a summary of the reading materials. In comparison, Alfred who also assigned reading usually checked students’ reading progress during every session. He assessed students’ reading journals in the UH by asking questions which related to the materials that they had read.
4.2.2.2.3 Instruments applied by only one participant

Table 4.2 also showed that there are some formative assessment instruments which were implemented by only one participant in line with their contexts and beliefs in learning. For example, Adam had a teaching strategy that he used to elicit students’ speaking skills. Data from the interviews and from observation showed that at the beginning of each session Adam assesses students’ communication skills by engaging a number of them individually in a conversation. He conducted this as he was checking the attendance list; thus, students did not realize that they were actually being assessed. For example:

Adam: Andy, I like your cap, you look handsome  
Student 1: Thank you Sir  
Students: Weeeiiiii …  
Adam: Andrea?  
Student 2: I am here  
Students: Weeiiii …  
Adam: Andrea, are you sick today?  
Student 2: No, I’m not  
Adam: Reina … are you sick too?  
Student 3: Itu lagi itu …  
Adam: Oh lagi not good ya? Sam? Are you sleepy Sam?  
Student 4: No Sir  
Adam: Or as usual? Biasanya seperti itu (usually like that)?  
Students: Yes … yes he is usual  
(Adam/03/P2)

Based on the responses given by students, Adam would put some notes on his attendance list.

As he said in the interview:

I’ve made some notes. Maybe make a point or example that the student was quiet even after I’ve made enquiries several times, that’s how the student was.  
(Adam/Int1/[13:32])

Deeper analysis from the interviews and classroom observation data indicated that Adam put more weight on students’ speaking skills. In his opinion, speaking skills can represent a students’ overall ability of the other language skills, he considered speaking to be the most important because communication in the target language will most likely be conducted orally. Therefore, this technique gave a context for the materials taught in the particular session, or
at least it made students pay attention. Further, he considered this technique was effective because there were a large number of students in his class. Wendy also had her own strategies for eliciting evidence of students’ learning. She believed that the students in her accelerated class already had considerable prior knowledge; thus, all she needed to do was to sharpen it, and develop their learning habits (Wendy/Int2/[42:59]). In addition, she also believed that English is a potential means to learn any knowledge, and technology is a potential means to facilitate communication. This probably explains why assessment instruments implemented by Wendy commonly involve students making use of their English in a real context. For example, she implemented video conferencing and team blogging. In the video conferencing, students had the opportunity to discuss some topics such as faiths, beliefs, and attitudes with three or four schools from different parts of the world. This activity was facilitated by an online moderator who would then give feedback to the students. Data from the observations revealed that during the activity Wendy did not participate herself, but she took notes on the students’ performances which she discussed with them in the next session.

Another instrument implemented in Wendy’s class is blogging. Wendy divided this activity into individual and team blogging. This activity basically required students to write and put their essay on the blog. The topics on the blog were mostly talked about events, tourism sites, or the environment. These blogs would be read by other students from different parts of the world because Wendy had established a network with those schools. In the interview she said:

So, team blogging is like this, I’m a part of facetofaithonline.org community that I’ve talked about. On facetofaith.online.org we can book team blogging, and video conferences. I can add the children into the names of registered students on face to faith. So, I can add them, so they have to be registered in order to link up or make you posts, we have to become a member of face to faith. Yesterday we were with students from Pakistan, America and Indonesia. Every post from the two countries
will receive comments, or other friends in the face to faith community will later respond. The activity only lasts three weeks. (Wendy/I2/[00:15])

This activity contributes to learning improvement because the students receive additional audience and feedback from their peers who are required to respond to the essay on the blog. In addition, it gives flexibility for Wendy to assess and give feedback because it is online and can be accessed at any time.

In comparison, Alfred emphasized repetition in his instruction including during formative assessment. He believes that repetitive activities have a positive impact on students’ learning. As he said in the interview:

> It’s really, really important, because, honestly formative assessment must be done. If it’s not done, then we can’t control and see students’ competency. And it needs to be done continuously and repeatedly, because I myself understand formative assessment is done repeatedly so students will understand it. I must emphasise that if it’s done repeatedly they’ll understand the concepts, they’ll understand the strategies and will understand the presented material. (Alfred/I1/[07.11])

This probably explained why Alfred had some ritual activities as shown by data from the observations. There were two activities that were always present in every session: free writing and public speaking. In free writing activity, Alfred gave two keywords and their definitions and as well as examples of how each of the word can be used in a sentence.

The students were required to write an essay which contained the given keywords. The essays were assessed based on the students’ abilities in using the words in a relevant context. Alfred explained the exercise in the interview:

> I consider, in free writing, right, whether they are capable or not of expressing ideas with the key words, but freely. It’s here that my scrutiny comes in: their understanding of the key words, then besides the understanding, their self-confidence while writing. (Alfred/I2/[03:56])

Public speaking was another activity that was always present in Alfred’s class. In this activity, one of the students would present a topic that had been prepared previously. Peers were given an opportunity to ask questions at the end of the presentation, and Alfred gave
immediate feedback based on the notes he took during the presentation. Alfred’s feedback was addressed not only to the presenter but also for the audience who asked questions after the presentation. Further, Alfred also gave some suggestions to the next presenter.

In Brenda’s case, evidence of learning was documented in the form of scores, and each score represented each KD (Kompetensi Dasar, Basic Competence). Basic competence is the minimum competency that students must achieve, and in every session there are several Basic Competences that students need to achieve; thus, Brenda commonly recorded students’ scores several times per session. In the interviews, she explained her reasons for gathering students’ scores as learning evidence

   In every meeting, because there are IPK [Competency Achievement Indicators]; for example, in this meeting there are three indicators, meaning there needs to be at least three instruments to measure the achievement indicators, whether the competency has been achieved or not. Because in our teaching practice, being successful or not is viewed as being based on the achievement of those competencies. So, the competency achievement indicators need to be written into lesson plans. That’s how things need to be these days. If it says in the lesson plan that students need to be able to imitate, is there evidence or not to assess that imitation? For example, the imitation needs to be like this, and it will be assessed like that. Then analysis. Can the students analyse things? Which means analysis needs to be assessed. As a measuring tool for the achievement of competencies. (Brenda/I2/[08:51])

Data from the interviews and classroom observation strongly indicate that Brenda put priority on learning mastery in her instruction, and this was also reflected in how she assessed her students. Data from the interview below support the indication:

   So for speaking like this, I tried the narrative Jack and the Bean Stalk. So at the beginning we listened to the person (from the recording) telling the story. “Oh so it’s like this right” in the tone of the person on “The goose and the golden eggs”, but it was very short. So someone talked and we listened. Then I asked some questions, and they answered. So indirectly I’m saying that, if someone is telling a story, they really have to understand its contents. At that point people started to unpack that when you tell a story there should be this, this and this feature. Then to make the story interesting, there should be props, to make it interesting. Only then should they develop the narrative. That part is unconstrained, it could be fiction, legends, myths. Then they write, and that will be assessed, I’ll assess it, then I’ll return it, and there will be this and that feedback, but I won’t ever change the plot. (Brenda/I1/[42:50])

Another example:
As an example of a text discussion, at the beginning (students) will read texts in the form of discussions, then they have to make some points, the pros and cons. Then they try to re-write it. This is the original text, right. For example, dental braces. Many of them will have an opinion, this one will say this, that one will say that and they take stock and then make a new text. So writing cannot suddenly appear by itself, it’s a process. (Brenda/I1/[47:00])

Data from the Observations confirmed the finding that Brenda carefully guided students before they actually arrived at the desired learning product. For example, when students were learning about news items, the terminal objective was that students should be able to make a news item in the form of television news, using screencast-O-matic software. Prior to achieving this objective, the students had to complete a set of learning activities which led the students to arrive at the terminal objective. Even in the process of writing itself, Brenda assigned students to use different coloured paper, so they were able to see the difference between the draft and the final product. In the interview, Brenda described another technique of guided writing that she applied:

In writing it is very clear, because there are my scribbles. My notes. For lower classes there are three lines. The first line is their writing, the second line is for my writing once it’s been collected. And the third line is for writing following revision. And then the fourth is for the final version. And the feedback results in improved writing. (Brenda/I1/[56.40])

To sum up, the findings of this study document some common instruments that are used by all participants, namely questioning, tests, and observations. In addition, teachers found their own individual ways to incorporate formative assessment into their classroom contexts. Even when participants used the same instrument, they had different procedures for implementing them.

4.2.3 Providing feedback that moves learning forward

Feedback is defined by Hattie and Timperley (2007, p. 81) as ‘information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one’s performance or understanding’. The potential means by which feedback can improve
learning has been well-documented (Kumar & Stracke, 2017; D. R. Sadler, 1989; Shute, 2008). Nevertheless, no researchers have developed an empirical formula on how feedback should be given in order to enhance learning. It is important to note, however, that a number of studies discuss several techniques of giving effective feedback are suitable in certain contexts. This implies that every teacher needs to find their own formula for effective feedback, and in the next section, I discuss how the participants provided feedback to their students, in order to move learning forward. It is divided into two sub-sections: teachers’ beliefs about feedback and the types of feedback provided by the participants.

4.2.3.1 Beliefs about feedback

Data from the interviews revealed that all participants agreed that feedback is essential for learning improvement. As Evelyn said in the interview:

I believe they need to know what their weaknesses are. I don’t say it’s obligatory, right, but it’s best if we tell them, otherwise they sometimes won’t know where their weaknesses lie. But by being told [they’ll go] “oh right”. Sometimes we can’t assess ourselves, we need someone else to let us know. (Evelyn/Int1/[34:23])

Despite its crucial role, Brenda considered giving feedback to be burdensome for teachers because it is time consuming and it doesn’t always provide significant proof of improvement. However, she argued that improvement can only be made if teachers provide feedback. Thus, the focus was on how to give constructive feedback instead of on whether or not feedback should be given. In the interviews, she described how her feedback improves students’ competencies:

When writing doesn’t receive feedback, then they will have errors in conveying ideas. But through feedback; although feedback is time consuming and not fun for some people. Some feedback isn’t useful because the teacher hasn’t made it useful. As for me, I’ll say “I’ll read this text, but who will read this text”. Later I’ll tell your friend to read it. So then I’ll give feedback to you. Then you write something new. It must be new. Then they’ll be aware “Oh it turns out I wrote this incorrectly”. So it means they must be forced to learn indirectly from what I’ve indicated. (Brenda/Int1/[55:18])
Data from the interviews revealed that participants were predominantly in agreement that feedback can motivate students and improve overall learning. As Alfred said in the interview:

Very much so…and they feel they are valued “Oh my work has been checked, oh yeah, this, this and this”. In written form, it is really good. But it’s very necessary and I’ll give it. (Alfred/Int1/[31:24])

Likewise, Daisy agreed that feedback will motivate students and enable them to be more confident in using the target language. In addition, Daisy believed that feedback also functioned as a reminder to students not to make similar mistakes in the future. As Daisy explained in the interview when she was asked whether her feedback benefited students’ learning:

Yes of course, from the test later, in the end of semester exam they’ll get the picture, because the problems will more or less be the same. Then if they’re familiar with the daily revision, then they’ll be ready. If they don’t ever do it they’ll get a shock. (Daisy/Int2/[08:32])

Nevertheless, in all cases, the participants agreed that feedback will benefit learning only if teachers can provide effective feedback. Further, the majority of the participants agreed that the efficacy of feedback depends upon whether or not the teachers can meet students’ learning needs. Brenda and Evelyn stated the importance of giving relevant feedback as follows,

Students need different feedback. It helps. They learn from their errors and the feedback we provide for the future: “Oh the student doesn’t make the same mistakes, this means they’ve understood the feedback that we gave”. So when a student doesn’t make the same mistake we can be sure the feedback helped, it was received. Sometimes we notice “Oh they’re still making the same mistake”, meaning our feedback hasn’t been understood. (Brenda/Int2/[20:15])

It just depends on whether we convey it personally or not. Amongst everyone, yeah, it’s done without mentioning names, like this is the mistake, because it’s a common error. Usually, for example for speaking, if a certain word is mispronounced, there’s more to correction actually. (Evelyn/Int1/[34:23])

All participants agreed that feedback is considered to be effective when changes are made by the students after they have received it. For example, Anne considered her feedback to be
effective when students were able to produce a learning product that incorporated her feedback. The efficacy increases when the students also enjoy the activity, as she said in the interview:

So we know our feedback is effective if students can produce good work. Especially if they enjoy the process. (Anne/Int2/[21:50])

Adam was one of the participants who had a slightly different opinion of what constitutes successful feedback. He believed that affective feedback leads to improved cognition and also improved overall behaviour (Adam/I2/[20:15]).

Further, some of the participants pointed out that the behaviour changes might take time to form. For example, Alfred claimed that it took a year before he could see the changes arising from the feedback that he had given. He gave an example, in the interview, of how he taught students to ask and answer questions during presentation:

I think yes (changes of students’ behaviour is the result of effective feedback). One of the ways actually. For example, in Year 11 I always remind them to be appreciative, in the first meeting they haven’t yet. The nature of it is indeed repetitive, asking and answering like that, directly. Finally, in Year 12 it can be seen. So, in the end the students know what is being assessed and required after I keep repeating the point. (Alfred/Int2/[14:02])

Further, effective feedback that results in behaviour changes should also be supported by teaching materials that encourage students to enhance their capabilities. As Wendy stated in the interview:

As seen from daring to perform, they express their opinions, even if they’re regarded as strange by their friends. In the beginning Amanda was shy, but now she’s confident to raise her hand, in the beginning she didn’t want to. Because they know to prepare to ask questions. When they are prepared to ask questions they listen carefully. Amanda is definitely smart, but she lacked confidence to ask questions. This means their confidence has improved, at least they have the confidence to ask questions. Other children were already great. The important thing is the material, like teaching them to expand the exploration of their abilities. (Wendy/Int2/[48:53])
4.2.3.2 Giving feedback

As mentioned in the previous section, effective feedback is not a fixed formula that can be implemented in all settings; teachers need to find their own strategies for giving effective feedback. This section talks about types of feedback given by the participants. It is divided into four sections, as theorized by Brookhart (2008): time, amount, mode, and audience.

Timing of feedback

In terms of timing, feedback can be either immediate or delayed. For example, for a performance-based instrument such as role play, grammar practice, questioning, and observation, immediate feedback is given. Whereas for other assessment instruments such as essays, portfolios, and journals, participants commonly applied delayed feedback.

Data from interviews revealed that all participants preferred immediate feedback. For example, when role play is the assessment instrument, Nelly preferred to give immediate feedback, as she said in the interview:

Yes to giving feedback. It’s best to describe things at the time. I note things down at the time, when they’re in the middle of things. But, in giving feedback, for example “Okay students, just now this group is really good especially this”, like that. (Nelly/Int1/[10:05])

According to Brenda, she preferred to give immediate feedback so that both she and the students were aware of the learning progress by the time the session ended. As she said in the interview:

I’m someone that needs to finish everything off in class so I don’t… so students know that marking hasn’t been taken home, you know they’d then have to wait. The students need to know what they’re studying including how they’re doing, except for the final writing. Students sometimes assess themselves for writing: Is there a title? They check. Is there a thesis? They check. After students assess themselves, I’ll assess them. (Brenda/Int2/[37:05])
In addition to informing students about their progress, two of the eight participants said that they needed to give immediate feedback and record it in their record books, to prevent forgetting about it as they have a large class. As Evelyn said in the interview:

I usually remember, if this particular student…but I also keep notes in a journal, today I’ll note down about a student and their weaknesses. But not always, sometimes I’ll be unsure of names. Especially average children, then I’ll need to note it down. But if I know them well I’ll remember, only if I need to I’ll put it in my journal. It’s required in Kurikulum 2013, because teacher observations is a part of assessment within the curriculum. (Evelyn/Int2/[05:06])

Data from the observations revealed similar findings, for example, when Daisy was teaching listening and checking students’ comprehension, she gave feedback as part of the IRE question pattern (Loef Frank, Kazemi, & Battery, 2007; Mcaninch, 2015):

Daisy: What is the conversation about?
Student1: Shopping
Daisy: Who went shopping?
Student2: the mother of Mica…
Daisy: yes, Mica’s mother and then … Where was the mother now?
Students: In the….
Daisy: Where was the mother now…? In the kitchen.
(Daisy/O2/P5)

Nevertheless, data from the interviews showed that immediate feedback could not be applied when the focus was on writing. The main factor that makes this difficult is large classes. The exception was for Alfred. In the interview, he claimed that he can give immediate feedback for free writing, one of the activities he applied in his class:

I use this (an iPad) so that it is easier to check (students’ work) and students can quickly see their results. (Alfred/I1/[23:40])

Data from the observation confirmed this finding. Alfred was able to give immediate feedback for almost any of the activities in his class. For example, in another activity, while a student performed a public speaking activity, Alfred sat at the back of the class and assessed the student’s performance including their response to questions from peers. At the end of the activity, Alfred gave immediate feedback.
Audience for feedback

In terms of audience, feedback can be divided into two categories: individual and group feedback. Data from interviews revealed that most participants used the two types interchangeably; however, most participants claimed that they preferred to give individual feedback. There were some reasons behind their preferences. Firstly, students are individuals with their own strengths and weaknesses, and feedback given by the teacher may not suit all students. Data from the observations revealed that most participants varied their feedback based on the individual student. For example, when Alfred was checking the students’ reading journal, he sometimes complimented those who had lower scores more than those who had higher scores in their reading journal. As demonstrated by the observation:

<table>
<thead>
<tr>
<th>Alfred</th>
<th>Andre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 1:</td>
<td>40</td>
</tr>
<tr>
<td>Alfred</td>
<td>John</td>
</tr>
<tr>
<td>Student 2:</td>
<td>17</td>
</tr>
<tr>
<td>Alfred</td>
<td>Good progress</td>
</tr>
<tr>
<td></td>
<td>(Alfred/O3/P3)</td>
</tr>
</tbody>
</table>

In the interview Alfred gave his reason for varying his feedback depending on the audience:

One, because seeing a child’s nature, it needs to be pointed that they need to work hard, while another child may already be doing well, so it’s different. So a child that isn’t doing so well, but wants to read, that needs to be appreciated.
(Alfred/I2/[19:42])

In a similar way, Evelyn also agreed that she gave individual feedback because students were different in terms of competencies. In addition, individual feedback would accommodate those who felt reluctant to ask questions in front of their friends. She said in the interview:

For me, it is to ensure whether the children have really understood what we’ve been passing on. Because, you know, students are different from each other, right? Perhaps it’s the expectation, whether we’ve given appropriate information, whether our instructions have been understood, so that’s first. Secondly, related to that, we can see directly, so individually we can do collective corrections which hit the mark. Maybe it’s a small error, but when we look it’s from the structure or choice of phrasing. Being more personal the children are usually more positive about us, and are quick to understand, as well as being closer to them. If we just stand out the front of the class there’ll be a distance between us and the students. If we approach them, they’ll feel closer to us, and not be shy to ask questions. (Evelyn/I2/[01:03])

163
Further, data from the interviews revealed that some participants had their own distinct way of providing feedback to their students. For example, in Brenda’s case, she would firstly identify difficulties encountered by students: whether she was addressing an individual or group. For individual feedback, she would assign the particular students some assignment that would scaffold their learning, or ask one of the student’s peers to be a mentor. On the other hand, if most of her students were encountering difficulties, she would reflect and find a solution for how to teach the materials. As she said in the interview:

Now for example if they’re stuck. Getting stuck doesn’t just happen to the low achievers, it can happen to smart students too. They can get stuck, for example while we’re teaching them… Cognitive capacity happens in stages. Perhaps they could all pass from the point of view of comprehension. (Brenda/I2/[15:09])

In comparison, Adam had his own strategies for giving feedback to students. He actually preferred to give individual feedback; however, the number of his students made it difficult to do this. He then implemented a technique which he referred to as ‘sampling and spread’. When he assigned an assignment, he asked students to come forward to show their work so that he could give feedback. Afterwards, the students were asked to spread their individual feedback among their peers. If, for example, ten students came forward, and none of them did the assignment correctly, he would ask students to stop working and re-explain the particular materials he had taught. In the interview, he explained the technique:

There’s willingness. That’s what I usually do. If it’s discussed one by one usually, if it’s too… If work has been collected from 10 children, and no one is correct… usually it’s “Okay sit down, and we’ll try it together”. Yeah, sampling those 10 people mean that the answers will spread to others, which I think is perhaps better. They’re more active too like that. (Adam/I2/[11:54])

In Anne’s case, feedback was given by involving students. Anne claimed that she gave not only individual feedback, but small group feedback (Anne/I1/[17:16]). She asked students to give feedback on their peers’ works. Later, she would also give feedback, and she described the process as follows:
Yeah… They also listen to my explanation and they also sometimes have ideas “Miss, how about we do it like this for example”. So it is not just feedback from the teacher, but also from their friends. So we give feedback together. (Anne/I1/[17:30])

**Amount of feedback**

Amount has something to do with how much feedback that participants gave to improve students learning. Most participants focused their feedback on correcting students’ sentences in terms of grammar and vocabulary regardless of the topic being discussed. For example, the criteria that Nelly used to assess two macro skills, speaking and writing, did not relate specifically to the lesson that she taught. She mentioned criteria such as pronunciation, gestures, and expression, for assessing speaking; whereas spelling and content were used to assess writing. As she said in the interview:

> use it [criteria]. For example, for speaking it’s the pronunciation. Their performance, expression, how their body language is, gestures. And of course the content…Is it correct or not to say it like that. For writing, you know it’s that, the spelling is checked. And the content too. The connections between sentences, and words. (Nelly/Int1/[11:12])

These criteria are considered to be very broad, and are not specifically related to the learning materials being assessed; as a result, the feedback given to the students was also broad, and focused more on the accuracy in general. However, other participants did give feedback based on the assessment criteria given previously. In the interview, Brenda explained how she assessed students’ speaking skills based on the criteria that she had established previously. In the example below, she shows how she assessed students’ speaking skills when they were doing role play of a narrative story:

> (some criteria for assessing speaking in narrative story) is there an opening? If there is an opening, their grade will be so much. There will be a rubric. If it’s there, in the opening you use, use…in the opening there must be orientation. Then if you insert a moral lesson, the grade will be so much. Was it just inserted anywhere, was there eye contact, was there communication with the audience? Did you use a prop? So having a rubric will help them, that I need to fulfil this, this, and this. Then there’s the closing, it’s there, it’s helped, there’s the rubric. (Brenda/I1/[42.50])

In a similar way, Alfred also gave feedback using assessment criteria that he had established previously (See p.136)
Data from the observations confirmed the finding that Alfred indeed provided feedback based on the assessment criteria that he had communicated to the students previously. For example, he told the students:

Alfred: The first I scored from this session is visual points. You know visual point? It is something that people can look out from you. After visual point is expression. Unfortunately Dika, I mean you need to try harder to do a very good expression. It’s kind plain, but it’s ok, no problem. If you have a very good expression, your presentation will be very interesting. Sayangnya, ekspresinya kurang main. [Unfortunately, the expression wasn’t animated enough.] Next time every single session you have to prepare the slide, outfit (Alfred/O1/P2)

In terms of content, most participants provided positive feedback on students’ work; then, they continued with an area that could be improved. For example, Anne gave individual feedback on her students’ writing, as follows:

Anne: So you have a good point about Africa, where do you got this kind of idea, from reading or watching movie (students nodded and said yes). So ... would you change this one, the capital letter then if it is Ok send to class email address and also the wise words. Oh there are some new words here, how to spell it? Sundra, tsetse? (Anne/O4/P5)

**Mode of feedback**

The data from the interviews and observations revealed that most participants commonly gave three forms of feedback: scores, compliments, and corrections. These forms were conveyed in oral and written media. Data from the interviews revealed that corrections in terms of reconstruction had always been the concern for teachers. For example, Alfred and Evelyn point out their opinion on the role of personal feedback. In the interview Alfred said:

I am the one who correct them (the essays) … It’s free, it’s up to them, if the grammar is incorrect or whatever it isn’t a problem. That’s not what’s been assessed. I just give the corrections for consideration, “Oh it needs to be like this, like that”. I return it promptly online so they can see it straight away. (Alfred/Int2/[05:16])

Evelyn said

I believe they need to know what their weaknesses are. I don’t say it’s obligatory, right, but it’s best if we tell them, otherwise they sometimes won’t know where their weaknesses lie. But by being told (about their mistakes, they would say) “oh right”. Sometimes we can’t assess ourselves, we need someone else to let us know. It just
depends on whether we want to convey it personally. If it’s done communally, without mentioning any names, it’s like this is the error, because it’s a common error. Usually, for example for speaking, if the pronunciation of certain words is wrong, it’s more likely to be corrected. (Evelyn/Int1/[34:23])

In addition to corrections, all participants gave feedback in the form of scores. Data from the observations revealed that scores were used at all stages of instruction. For example, when Adam assigned students to write a paragraph, he informed the class that those who submitted early would be entitled to the best score. In a similar way, Brenda also assigned scores rapidly in her class. In one session, she recorded students’ scores more than three times. According to her, she often assigned scores to students because it was easy for the students to interpret them. As she said in the interview:

But actually it’s not like that, because they know that numbers are just temporary representations. It’s expected that abilities, not numbers, will be internalized. Numbers are just the preliminary picture. So, they go “At this point I’m in this safe position”, so they take the average. So grades help them position themselves, so they I’m with students that need extra guidance, or I can move onto the following material. It’s easier for them to interpret their abilities through scores. It’s the easiest tool for students. (Brenda/Int2/[32:00])

Evelyn was also one of the participants who commonly assigned scores as feedback for her students. Compliments are another form of feedback. Evelyn thought that scores and compliments are the two most effective feedback modes for students. As she said in the interview:

Lots of children ask about grades right, but compliments make students more motivated to learn. So the two of them are both effective. Grades will describe them more. Compliments aren’t actually feedback, right? Feedback is when we let them know about weaknesses so they can be improved. (Evelyn/Int2/[17:42])

To sum up, all participants agreed that feedback helps students to improve their learning. Although a number of studies have argued that there has not been any empirical agreement on how to give effective feedback, findings of the study indicate that there is common agreement on how feedback should be given to students. For example, all participants preferred immediate feedback compared to delayed feedback, and individual feedback rather
than class feedback, and scores and compliments were practised by all participants. Further, findings of the study showed that the nature of the assignment influenced their feedback.

4.2.4 Activating students as learning resources for one another

One of the features of formative assessment that relates closely to socio-constructivist learning is students’ involvement in the process of assessment; students learn from their environment including their peers. It is important to note that the strategy is not confined to peer assessment solely; rather, it involves various instruments that engage students in the process of assessment. According to Wiliam and Leahy (2015) collaborative learning and cooperative learning are often mentioned as the focus in this strategy. Further, they stated that this strategy does not focus solely on collaborative nor cooperative learning, rather it emphasises how peers can improve each other’s performance, but not only by judging performance. This section focuses on how the participants involved students in the process of formative assessment.

Data from the interview revealed that most participants agreed that positioning students as learning resources for one another would assist them in overcoming their difficulties. They believed that peers can be valuable resources for learning because students can share ideas and learn from each other. As Daisy said in the interview:

> Children can be better helped in groups. They can share with those who didn’t have ideas, so the smarter ones can share and provide input. So I hope that the leaders can inform their friends (Daisy/Int2/[02:27]).

Likewise, Alfred also claimed that he put students into pairs when conducting one of the school projects so that they could help each other. As he said in the interview:

> But there will always be groupwork. They work in a group, there is another form of activity, they do a written project. In first class, second semester, the project is to write a short story based on their own imagination. Of course, it isn’t straight to writing, it begins slowly with freewriting. Then they produce the work. They’re split into several groups to complement each other, to discuss with each other, to help each other, and be helped, together. (Alfred/Int1/[23.40])
Based on the data from the interviews, two out of eight participants said that they implemented formal peer assessment, which includes a peer assessment sheet. Brenda and Evelyn give vivid guidance prior to the implementation. As they said in the interviews:

Yes, I told them first. Because I told them beforehand what will be assessed, automatically those that perform will know what is being assessed. (Evelyn/Int1/[16:16])

It usually works because I give them guidance first, then you have to say this to them. As if I’m testing you. So the questions can be about this, but they may be changed. Then for example you cannot mark it off the list if they didn’t speak in English. Everyone is given some kind of script. (Brenda/Int1/[21:24])

The rest of the participants also included students in the process of formative assessment, but they did it informally. In other words, participants did activate students as learning resources for one another, but they did it without a peer assessment sheet. Each participant had different ways of implementing informal peer assessment. For example, in Nelly’s case she would ask students to assess their peers’ performance as they gave their presentation. As she said in the interview:

It can also be like this. (when students were presenting, after the presentation I asked) “which one is your favourite group?” Or I informed the students that they must pay attention to their friends’ performances because they should also assess their friends (Nelly/Int1/[10:26])

Thus, I asked the students “Does anyone have objection If I give an A for this student” by doing so, students were actually conduct assessment (Wendy/II/[43:58])

In fact, data from the observations confirm that more participants used informal peer assessment than formal assessment. For example, after Wendy had used video conferencing, she played the video recording of the interaction and asked students to assess their peers’ performance. Data from the observations revealed that throughout the process of the recording, students continue conducting peer assessment. They corrected their peers’ pronunciation and also grammatical mistakes.

In Nelly’s case, students peer assessed themselves without being assigned to do so. Nelly, who gave the students the same test that they had for their mid-term examination, did not
assign students to conduct peer assessment. However, as students were doing the same test they worked together with their friends. In the process they discussed, questioned and negotiated their answers with their peers. Further, they gave different answers from what they had given on the previous test as they discussed it with their friends.

Data from the observations revealed that activating students as learning resources for one another is mostly conducted in peer correction. All participants applied this activity. The observations showed that when a score was involved students became more critical and interested as teachers explained the correct answer. As has been mentioned in the previous section, testing has a strong influence on participants’ practice of formative assessment.

Data from the observation revealed that students felt reluctant to ask questions on materials that they had not yet understood. Adam who had 38 students made use of this situation. As he said in the interview, he would ask students to ask their peers first when there were some points which they did not understand. As he said in the interview:

I tend to give it to the children: “Please talk about it with your friends. If you don’t understand, ask your friends first, discuss it with friends first.” Then I will check whether their friends’ explanation is correct or not. Doing it like this, aside from learning, they’re also socializing. (Adam/I2/13:40)]

In terms of the purpose of using students as learning resources, data from the interviews and observations revealed that during a peer assessment activity most participants commonly assigned students to judge their peer’s performance. As Evelyn said in the interview:

For writing it’s like that. For speaking I also like to give them a sheet with maybe just five points, for example, whether your friend can understand this. The points are whether the pronunciation is right or not, whether the voice is loud or not, whether they are confident enough. I usually just give five points. So it’s enough for the children to assess using A B C, then I will do it. A is X amount. so they don’t have to worry. For example, one performs, and another will assess, for example is A confident or lacking in preparation. (Evelyn/Int1/[14.50])

In a similar way, Anne also activated students as a resource for other students. This was revealed in the observation data when she did individual feedback:
Anne: Is there friends that giving you time to discuss? Anyone that helps you in the class?
Student: I did. I discussed the grammar with three of my friends.
Anneas: ok, good. You have to check, you need discussion with me or your friends, so you have to make it better. Please revise… (Anne/O3/P7)

The data from the observations revealed that female students participate more actively compared to male students, in Evelyn’s class. In the MoRA school, male and female students sat separately. When they were assigned a writing project, male students had the tendency to finish their part individually; on the other hand, female students collaborated to produce a good result. A similar condition occurred in Daisy’s class. When Daisy assigned students to discuss and present on the topic of describing people, female students gave more attention to how to make their presentation successful, while male students divided up their roles in the presentation and had no further discussion on how to make their presentation successful.

Interestingly, in some cases, students had become accustomed to assessing their peers. In Evelyn’s class, for example, when she asked her students to present their work, their peers in the audience continued assessing the presenter even though they had not been assigned to do so by the teacher. For example,

S1: Blouse /blaws/
Ss: /blauz/ (students were in choir as they taught their friend mis-pronounced) the word.
(Evelyn, O4/P10)

Further, findings from the study also showed that when students encountered difficulties, they relied more on their friends rather than their teachers. Data from the classroom shows that in most cases, students asked their friends first instead of asking the teachers, as can be seen in Wendy’s class when one of the students presented a topic, and she did not know the word “pendaftaran” in English. Instead of asking the teacher, she asked her friends.

S1: When is the ‘injury time’?
S2: What?
S1: Pendaftaran paling lambat (the due date for registration)
S2: Oh … pendaftaran apa Bahasa Inggrisnya? (What is registration in English)
SS: Registration (Wendy/O3/P4)

In summary, the data shows that involving students in the process of assessment is an integrated part of the participants’ formative assessment practice, and this strategy was not always being implemented through formal peer assessment. Some participants did use formal peer assessment, but others implemented the practice informally; that is, they did not prepare peer assessment sheets before they involved student-peers in the assessment process.

4.2.5 Activating students as the owners of their own learning

Simply put, this strategy is concerned with encouraging students to take control and be responsible for their own learning. It is commonly related to several terms such as autonomous learning, independent learning, or self-assessment. This capacity is important because students need to continue their learning out of school, and teachers will not always be available to monitor students’ learning. This section talks about how teachers nurture and develop students to become autonomous learners.

Data from the observations showed that the participants predominantly encouraged self-assessment through questioning. Participants give follow up questions, or simply give some time for students to think back over the answer they had given. For example, in Alfred’s class as he was teaching sentence connectors:

Alfred: Helen, is it a correct sentence?
Hana: Yes…
Alfred: Why is it correct?
Hana: Because there is a sentence connector ‘or’

(Alfred/O2/P1)

In a similar way, Anne also gave her students the opportunity to self-assess their work prior to her assessing it. For example, when she was holding individual consultations to give feedback on the short stories that the students had written, she did not pinpoint students’
mistakes directly; rather she gave some ‘hints’ for several points that needed to be revised.

This is described by the observation data below:

Anne: Have you checked your work?
Student: I did Ms.
Anne: Have you revised the misspelling (words),
grammar, capital words?
Student: I did Ms.
  (Anne read the short story)
Anne: I think you need to check it one more time,
especially the capital letter, and also the flow of the
story
Student: Ok, should I continue or change the plot?
Anne: It’s up to you, whether you can develop the story
or change the plot a bit

Besides questioning, the teacher also used journals and/or dialogue journals to activate
students to be the owners of their learning. Data from the interviews revealed that two of the
eight participants claimed that they used a journal and/or dialog journal during formative
assessment. Interestingly, these journals served two functions, not only for students to
identify their own progress, but also for teachers to reflect on their own teaching. As Wendy
said in the interview:

They write about the English lessons and then write a reflection… At the end of the
year, or semester, or session I’ll ask them to write a journal. Whoever wants to write,
just send it to me. So, I have feedback. These are their journals. (Wendy/I1[53:49])

In comparison, Brenda implemented dialog journals which also functioned in two ways. She
explained the procedure in the interview:

Back and forth, it goes like this, for example: “Did you learn something today?”
Later they’ll write, “I learnt that to write a text is not easy” Then once that’s done the
student has the right to ask a question, because every person has the right to ask two
questions, but not always. If they don’t have a question they don’t need to. But then,
from there for example “I didn’t understand what you taught today”. I will comment
on everyone’s, both those that understand and those that don’t. Then I’ll write, if for
example they did understand the lesson: “I’m glad to know that you understood the
lesson today. I hope it’s the same for the next lesson.” And then the next one will
return to the student, who will write “Thank you, I will try my best in the next
lesson.” Then I’ll give another prompt, for example “Can you understand the
vocabulary written in the text?” Then they’ll write again “The words I don’t
understand are this one, and this one, and this one.” Sometimes I’ll ask “What did
you learn today?” Then the student will make a summary. So indirectly it becomes a
tool, for both me and the student to know what they know. Or “What do you think
about while I’m teaching?” Then they can write what it is, it may even be
“whatever”. (Brenda/Int1/[41:06])

Interestingly, although developing students’ self-assessment ability logically relates to self-
assessment instruments, only three of the eight participants mentioned self-assessment as one
of the instruments applied during formative assessment. Brenda was one of the participants
who claimed that she applied self-assessment:

From the beginning. This is how I do it, I show them first. “This is the ideal. If you
want to make a caption it has to be like this, the rules for captions are like this. If you
want to make a stand alone picture it must be like this.” It needs to be explained first.
Then there will be the rubric. Now we’re going to present news. For example, there
might be something accompanying their work: 1. It needs to be clear who it is for; 2.
What kind of font, should they use a caption or not, then how many lines of
sentences are provided for in the caption. So there’s a guide for while they write. So,
they just need to follow the checklist: Did I write the title? If they wrote one, they
can tick it off. Did I write in simple sentences? Do I understand what I wrote? They
can tick it off. So, I use self-assessment, because self-assessment is the easiest way to
measure someone. Sometimes I’ll request self-assessment for myself, like this,
earlier we wrote: I like self-assessing myself through the newspaper, I like the
material in the newspaper, I still have three questions for this material, so I just need
to ask what the questions are. (Brenda/I1/[38.13])

An interesting finding was that two of the eight participants claimed that the ability to self-
assess related to students’ achievement. Wendy and Evelyn claimed that low achieving
students were mostly unable to assess their own performance: their self-assessment results
often do not correlate with the results of the participants. As Evelyn said in the interview:

Sometimes students are honest in assessing themselves, sometimes they’re not. For
example, they give themselves a C where they should get a B. Maybe they don’t
want to appear superior. (Evelyn/I1/[18:23])

On the contrary, Wendy, who was teaching an accelerated class, claimed that the result of
students’ self-assessment is more or less similar to her own assessment. As she said in the
interview:

It’s relevant. The students and I are usually on the same track. But, in classes with
low-achievers sometimes there’s a difference. (Wendy/I1/[57:54])

According to Brenda, one of the ways to overcome the inaccurate judgment of students’ own
assessment is for the teacher to inform students that their results will be checked. Further,
she mentioned that over time students would develop the ability to self-assess themselves.

As she said in the interview:

> Actually, my aim is this, when it comes to measurement it should be connected to the standard. They can measure how far they themselves are from the standard, they can assess themselves. And it’s fairer, because it’s judgemental, and it’s more subjective. But actually their score isn’t final. I’ll take the paper to see whether they’ve scored themselves too high, because it refers to the standard, have they understood it or not. But because they’re used to it. This has been going for seven months, right, and it was taught. They’re given an opportunity, and I just trust that they won’t lie, because I’ll double check it. And there’s an MOU that I won’t tolerate cheating. That’s there. Maybe that’s one of the reasons for teachers checking, otherwise we wouldn’t know. (Brenda/I2/[11:57])

Data from the interview revealed that participants did not only apply instruments that encouraged students to develop self-assessment skills, although some participants nurtured students’ self-assessment skills by giving them the opportunity to plan for their own learning. Data from the interviews revealed that Brenda is one of the participants that provided this opportunity:

> Students are usually critical when I tell them about the syllabus. There’s an offer ‘this is your turn’, negotiate with them about what material I’ll include. (Brenda/I1/[15.09])

This section presents how teachers implemented one of the strategies in formative assessment; that is, how they involved students as assessors of their own learning. Findings from the study showed that all of the participants applied the strategy; the majority using questioning, while others used self-assessment and journals.

**4.3 Challenges and solutions**

The previous section discussed how formative assessment strategies were implemented by the participants. Findings from the current study show that formative assessment is context-embedded; with the participants generally incorporating strategies relevant to their setting. Each context has its own formative assessment culture depending on the school, the students or their level of ability. Although these participants showed that they had ‘figured out’ ways to implement the strategies, the majority of them still encountered several challenges when
applying formative assessment. This section focuses on the challenges encountered by the participants during the process of formative assessment. Data for this section was derived from interviews and observations. Further, this section also talks about participants’ efforts to overcome formative assessment challenges.

4.3.1 Identifying formative assessment challenges

Findings from the current study revealed that most of the participants considered time to be one of the challenges for implementing formative assessment; thus, when the government decided to reduce the time allotted to English subjects, the challenges escalated. As stated in Kurikulum 2013, the time allotted for English in high secondary was 180 minutes per week, but today it is only 90 minutes per week. Another factor that was considered to be one of the challenges to applying formative assessment is large class sizes. Based on the MoEC Decree Number 129a Year 2004 (Peraturan Menteri Pendidikan dan Kebudayaan Nomor 129a/U/2004, 2004), classes in Indonesia may consist of 30 to 40 students. Class sizes in Indonesia can be considered large, for example, Brenda and Nelly had more than 35 students in their classes.

All participants argued that that the time allotted was insufficient, and they gave various reasons for their argument. For example, Adam said that teachers have various roles, and it requires a large amount of time to perform these roles. In Nelly’s opinion, it was difficult for her to assess students’ writing and give feedback to students in such a limited time. This argument was also raised by other participants. Providing personal feedback, which the participants believed to be beneficial for students’ learning (see Section 4.2.3.1), became difficult because of the time constraints.
In the interview, the participants claimed that the time for English teaching is insufficient for them to cover the materials assigned by the curriculum and there is little time left for formative assessment. For example, Evelyn gave a description of how she had to reduce formative assessment due to time restraints:

In the past with four hours, lots of children gathered together their work, so they put it into a folder. Because it was four hours, it wasn’t too bad. In one week, they could read quite a lot. Now with two hours the children don’t have enough time. So, there isn’t a portfolio and journal. (Evelyn/I1/[16:16])

Given the combination of limited time and large classes, implementing formative assessment in the Indonesian context is undoubtedly challenging. Nevertheless, data from the observations revealed that limited time and large classes were not the major challenges for teachers attempting to implement formative assessment. For example, class size did not correlate with the time needed to give feedback. Alfred, who had 32 students, was able to hand back his students’ essays with his feedback sooner than Wendy, who had only 13 students. This was despite Wendy’s claim that she did not encounter any difficulties in performing formative assessment. As she said in the interview:

Not really, so far I’m good with it. There aren’t any problems. Yeah, for assessment. Yeah, no. I can do everything. I feel actually it’s about how we apply things, and simplify things, with the current perceptions. It’s not difficult. It’s just how it is. Not really. For me. But I don’t know about others’ opinions. “Oh it’s complicated!” like that. (Wendy/Int1/[63:33])

This implies that it is only in the perception that large classes contribute to difficulties in applying formative assessment; in fact, it is the way participants manage their workload that influences their assessment practice.

Another source of difficulty relates to the reporting systems. Three participants claimed that it is not feasible to provide a description of each student’s ability, because they have so many students. These participants claimed that assessment systems should be simple and meaningful. They also mentioned that reporting systems that require teachers to make score
conversions are complicated. In addition, teachers are required to assess the core competence of behaviour and faith. These views surfaced mainly in relation to the implementation of Kurikulum 2013 prior to the revision, although today they are no longer expected to assess students’ behaviour and faith. An interesting comment was made by Daisy. She stated that regardless of the effort that she made for the reporting systems, parents did not appreciate this because they could not understand the current assessment systems. This had the effect of discouraging her from using the reporting systems. Clearly, the participants would like to have assessment systems which are simple, but meaningful; as Adam and Daisy said ‘I don’t want something that is complicated’.

Interestingly, although data from the observations revealed that UN hampered the implementation of formative assessment, none of the participants mentioned this factor during the interviews. Teachers who were teaching 12th grade students, often projected their teaching in line with the UN and they were required to set aside a large amount of time for UN related activities. For example, Nelly and Anne had to sacrifice their teaching time for UN mock tests, and add extra hours to practice UN tests from previous years.

4.3.2 Overcoming formative assessment challenges

As mentioned in the previous section, the challenges of formative assessment can be overcome by managing the workload and the overall assessment practice. Findings from the study showed that several participants had managed to overcome some challenges when implementing formative assessment. Most of the participants preferred to integrate educational technology when dealing with barriers during formative assessment practice. For example, Alfred made use of Google Drive; a file storage software which allows users to store files in the cloud, share files, and edit documents, spreadsheets, and presentations with collaborators. With his iPad that he always carried during teaching, Alfred claimed that it enabled him to give immediate feedback on students’ essays because he could assess them
literally anywhere. It would be different if students’ essays were submitted in paper form, as he stated in the interview:

Let me tell you about it. So it’s like this, right. Giving students writing tasks is a huge amount of work when I have to give feedback, because I have four classes. To make things easy, they usually use a book for writing, sometimes I need to organize the book, sometimes it (the book) gets lost, then the students will ask about it. It gets complicated! So I use this (iPad/tablet) to make checking easier, and students can immediately see their results. (Alfred/I1/[11:58])

In the same way, Anne made use of WriteToLearn, an internet-based learning tool, to check students’ essays and store students’ learning progress. The program assessed students’ essays to the extent of vocabulary, grammar, and punctuation. The software has reduced her time to assess technical mistakes, and Anne can now focus her feedback on the content level of the essay such as organization, coherence, and ideas as revealed in the observation data. In addition, the program also stores students’ works through time; thus, it can also serve as portfolio of their work. Likewise, Wendy made use of Quipper, a web-based learning tool that provides various different practice lessons and tests. Both teachers and students have access to the lessons and tests. The web enables students to practice at any time and get feedback almost immediately. According to Wendy in the interview, this website helps her a lot because she considered writing tests to be complicated and time consuming.

Nevertheless, the study findings also showed that attempts to overcome challenges to implementing formative assessment with educational technology were not always successful. In some cases, these devices or software were not effective. For example, data from the observation showed that two of the participants who took their students to study in a language laboratory could not actually operate several of the devices; consequently they had to perform formative assessment and the overall instruction in the normal classroom. Not only did these participants not get the benefit of integrating educational technology, they did not overcome their problems with implementing formative assessment.
Besides integrating educational technology, data from the interviews showed that another solution, implemented by some participants, is to add extra time for students whom they consider to be having difficulties with learning. Four of the participants applied this technique. Evelyn who was teaching in a boarding school has the advantage of being able to conduct ‘remedial teaching’ outside class time. She usually grouped the students, and asked them to come after school hours. Interestingly, in the interview, Evelyn said that it was common for the high achieving students to come to her for extra teaching the students who actually needed her help. Brenda, who was teaching in a public school did not have the privilege of providing extra lessons outside class time; so, she gave extra assignments for students who were still having difficulties coping with the lessons. In the interview she said that the dialog journal that she used in the class helped her identify which students needed further assistance with their learning. In this way, she was able to reduce the learning gap among students.

In the interview, Brenda also said that she commonly categorized students according to their competency at the beginning of each school year. She would ask students a simple question such as ‘Do you like English? Why or why not?’ Brenda asked the students to answer in written form because time was insufficient to do this orally. Based on the students’ answers, she divided them into three categories: low, average, and high achievers. She stated this categorization was reviewed throughout the school year; however, based on her experience, not many students were categorised differently from her first diagnostic test. Categorizing students at the beginning of each school year based on the diagnostic tests helped her in several ways: first, it gave her an insight into how to give instructions from the beginning; second, it enabled her to monitor the students’ progress over time; and last, it enabled her to know which group of students needed more attention and to decide which teaching strategies were suitable for which group of students.
This section has described the challenges encountered by the participants in implementing formative assessment. Findings of the study showed that the participants’ claims concerning difficulties in performing formative assessment were different from what I observed. In the interviews, the majority of the participants claimed that large classes and limited time were barriers to applying formative assessment; however, data from classroom observation showed that class size did not influence the participants’ formative assessment practice. In this section, I also discussed how teachers overcome challenges when performing formative assessment. The study’s findings showed that teachers used integrated educational technology and classroom innovations to overcome barriers to implementing formative assessment.

4.4 Summary

To sum up, understanding and defining formative assessment is complicated. Findings from the study showed that participants had differing perceptions as to what formative assessment is; they ranged from tests at the end of a unit of work to any formal and informal assessment. It is important to note that all participants agreed that formative assessment is carried out during learning. However, despite there existing superficial differences in the way that participants defined formative assessment, a deeper analysis revealed that participants perceived the key elements and potential benefits of formative assessment reasonably similarly. Further, data from the study revealed that there is an element of formative assessment that can only be found in language study; it is the integration of language skills.

In terms of formative assessment practice, findings of the study confirm that the participants implemented formative assessment strategies in their instruction, and used them to enhance learning. Moreover, this research revealed that several teaching innovations were made by the participants to tailor formative assessment practice to their context. For example, three of the participants applied portfolio, but it was implemented differently in each setting: one of
the participants applied technology-enhanced portfolio, another participant modified the procedures, and others used standard portfolio. In addition, findings from the study also revealed that this teaching innovation helped the participants to overcome challenges to implementing formative assessment, such as large class sizes and limited time. It is important to note that there were some formative assessment practices that needed to be refined in order to ensure that the practice contributes to learning improvement.

Overall, findings from the study have shown that these participants have presented a repertoire of excellent examples of how formative assessment can be carried out in the Indonesian context. Hence, this implies that teachers in Indonesia have considerable potential to implement and gain the utmost benefit from formative assessment. These key findings will be further discussed in the next chapter.
Chapter 5
Discussion

The previous chapter presented the findings of the study in response to the research questions:

1. How do ELT senior high school teachers in Indonesia perceive formative assessment?
   a. How do they understand the notion of formative assessment?
   b. What do they believe are the advantages of formative assessment?
   c. Ideally, how do they think formative assessment should be carried out?

2. How are teachers’ perceptions manifested in their formative assessment practice?
   What strategies, instruments, and practice do teachers employ in implementing formative assessment?

3. Do teachers encounter challenges when implementing formative assessment? If so, how do they cope with the challenges?

In this chapter, I discuss the key research findings presented in Chapter 4. I interpret, compare, and evaluate them according to the theoretical framework applied in this study (see Figure 4.1). Although the teachers had different understandings of what defined formative assessment, they understood its key elements and its benefits, and this was manifested in their teaching practice. The participants applied formative assessment strategies to good effect, and some were able to overcome challenges which they encountered, during implementation, by integrating educational technology and innovating teaching methods. The findings have also provided several excellent examples of how formative assessment can be implemented in various contexts. Nevertheless, it is important to note that some practices need to be refined in order for teachers and students to get the utmost benefit from formative assessment.
In line with the aforementioned key findings, this chapter is divided into the following sections: surviving the fuzziness of formative assessment, applying formative use of tests, building a formative assessment culture, overcoming formative assessment barriers, and missing formative assessment opportunities. A summary of this chapter can be found in the last section.

5.1 Surviving the fuzziness of defining formative assessment

The first question in this study sought to determine teachers’ understandings of formative assessment in terms of its definition, potential benefits, and key elements. The current study found that all participants agreed that formative assessment is conducted during learning, and it aims to improve learning. However, when it came to defining formative assessment, the participants’ definitions ranged from ‘test at the end of unit of a lesson’ to ‘any formal and informal assessment’; moreover, one of the participants excluded mid-term and final term assessment from formative assessment. There are two possible explanations for the widely diverging definitions of formative assessment.

First, the concept of formative assessment has often been broadly defined. For example, Michael Scriven, who established the term, defined formative assessment as follows:

Summative assessment has referred to tests administered after learning is supposed to have occurred to determine whether ‘it did’. Formative assessment has been the label used for assessments conducted during learning to promote, not merely judge or grade, student success (Scriven, 1967 cited in Stiggins, 2005, p. 326).

Likewise, Black and Wiliam (1998b, p. 2), who revived the concept, also gave a broad definition of formative assessment, which they defined as:

… all those activities undertaken by teachers - and by their students in assessing themselves - that provide information to be used as feedback to modify teaching and learning activities. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching to meet student needs.
Defining formative assessment is often complicated (Harlen & James, 1997; Popham, 2006); hence, a number of studies report that this notion has often been misunderstood as ‘any assessment other than tests’ or as ‘frequent tests’ (Davison & Leung, 2009; Wiliam, 2011).

Second, this result may be explained by the fact that there has not been any agreement on the exact definition of formative assessment. As explained in Chapter 2, Section 2.2, efforts have been made to determine what formative assessment is by proposing a new definition or giving it new labels, yet none of these efforts have satisfied scholars, and as a result, formative assessment remains poorly defined. Presumably, the interpenetrated connection between formative assessment and teaching, and the nature of formative assessment, which is contextually situated, have contributed to the complexity of defining it.

Another important finding concerning teachers’ understanding of formative assessment was that despite the participants having differing definitions of formative assessment, they all agreed that formative assessment is conducted during learning and that teachers should use the results to improve learning. These findings are in line with those of previous studies that emphasise learning improvement as the essence of formative assessment (for example, Black, 2015; Black & Wiliam, 1998b; Chappuis & Chappuis, 2008; Hamm & Adams, 2009; Pinchock & Brandt, 2009; Stiggins et al., 2006; Wylie et al., 2012).

Furthermore, the current study found that all participants agreed that formative assessment has the potential to benefit learning. Specifically, formative assessment enables teachers to identify students’ learning needs, or in Wendy’s words, ‘identifying whether or not students’ have made any learning progress, and recognizing in which learning area students need further assistance’. The majority of the participants also stated that one of the benefits of formative assessment is to provide information about teaching areas that need to be improved, or in Daisy’s terms ‘formative assessment is a reflection of my teaching’. These
results are in accordance with those of previous studies which mentioned that formative assessment can provide information about learning and has the potential to improve teachers’ teaching practice (for example, Ash & Levitt, 2003; Black & Wiliam, 1998a, 1998b; Wiliam, Lee, Harrison, & Black, 2004).

Generally the participants agreed that formative assessment constitutes the following elements: process, feedback, adjustment to teaching, and students’ active involvement. The majority of the participants agreed that formative assessment is not a one-off assessment activity. They agreed that there are sequences of activities that lead the teacher to make decisions about students’ progress in learning. Participants argued that a one-off assessment may lead to inaccurate judgments about students’ competence, and as a result, teachers may not provide the learning assistance needed by students. Further, participants argued that assessment should involve a process because language learning involves four language skills, and they believed that students’ competence, in each of the language skills, may vary. This finding is in accord with previous studies that indicate process as a defining characteristic of formative assessment. There are several definitions which state that formative assessment is a ‘sequence of activities’ (Black & Wiliam, 1998a, 1998b; Cowie & Bell, 1999; OECD, 2005; Rea-Dickins, 2006; Shepard, 2005; Stiggins, 2005).

The majority of participants also mentioned feedback for students, teachers and schools as a crucial element in formative assessment. This finding is in line with those of previous studies which mentioned feedback as a central element of formative assessment (Black & Wiliam, 1998b; Irons, 2008; Moss & Brookhart, 2009; Shute, 2008). It is through feedback that teachers are able to move learning forward, although it is important to note that not all feedback benefits learning because, to a certain extent, feedback may also be detrimental to learning, if it does not provide information on how to improve learning.
Teaching adjustment was also considered by the participants to be one of the key elements of formative assessment. This finding corroborates several studies that found formative assessment would enable teachers to make relevant teaching adjustments (for example, Buck & Trauth-Nare, 2009; Cauley & McMillan, 2010; Chappuis & Chappuis, 2008; Curry, Mwavita, Holter, & Harris, 2015). In fact, teaching adjustments can be regarded as the most important element of formative assessment because this is what distinguishes formative assessment from other assessment types (Black, 2015; Black & Wiliam, 1998a, 1998b; Cauley & McMillan, 2010; Chappuis & Chappuis, 2008; Curry et al., 2015; McManus, 2008; Pinchock & Brandt, 2009; Wiliam & Leahy, 2015).

One unanticipated finding was that only two participants, Evelyn and Brenda, considered students’ active involvement to be one of the key elements of formative assessment. Evelyn said that students can also be potential assessors who can assess their peers’ work as well as their own work. In a different sense, Brenda claimed that one of the benefits of formative assessment is that it encouraged students to self-assess (see Section 4.2.5 for further details). These findings suggest that most of the participating teachers still consider assessment to be the teachers’ sole responsibility. This argument is supported by the finding that none of the participants mentioned students’ increased motivation as one of the benefits of formative assessment.

These findings suggest that it is only teachers who use the evidence of learning and not the students. In other words, the students’ role in the assessment process is still considered to be passive, while learning agency is held by teachers. These results are likely to be related to classrooms in Indonesia which are still dominated by a teacher-centred approach as documented by several studies (Ariatna, 2016; OECD, 2015; Zuhdi, 2015; Zulfikar, 2009). The teacher-centred approach puts the responsibility for learning in the hands of the teacher.
and leaves a minimal role for students; hence, assessment, and learning from assessment, is also considered to be the teacher’s responsibility.

Furthermore, the present study also contributes several new notions of formative assessment. The findings of the study reveal that the participants considered that formative assessment should integrate learning domains and language skills. According to the participants, formative assessment should not only concern the cognitive domain, but it should also include affective and psychomotor learning domains. Regarding formative assessment in the affective domain, most participants include students’ engagement as one of the factors to be assessed.

A possible explanation for this might be that the participants are required to assess students’ affective domain, as stated in the policy on the standard of educational assessment enacted through MoEC Decree Number 23 Year 2016 Article 3, which in detail stipulates that students are assessed based on: behaviour, knowledge, and skills (Peraturan Menteri Pendidikan nomor 23, 2016). The Decree specifically stated that students’ behaviour needs to be assessed in the following manner:

- Observe students during learning,
- Take notes on an observation sheet which has been preprepared,
- Follow-up action based on the results of observation, and
- Describe students’ attitudes.

As stated in the previous paragraph, the participants claimed that they assessed students’ behaviour; however, deeper analysis showed that most participants focused on students’ engagement during learning generally and only assessed students’ behaviour when identifying whether or not they have understood the learning materials. Most of the participants did not use an observation sheet to conduct the assessment (see Section 4.2.2.2
for further details). These findings raise intriguing questions regarding the teachers’ understanding of how to perform assessment for the affective domain. As several studies have mentioned, identifying students’ affective domain enhances learning (Boyd, Dooley, & Felton, 2006; Buissink-Smith, Mann, & Shephard, 2011; Hall, 2010); hence, this study suggests that an empirical study on how to assess students’ affective learning domain is required. As revealed in a study conducted by Isnawati and Saukah (2017), efforts and behaviour are part of a teachers’ assessment in Indonesia.

Besides integrating learning domains, the findings of the study also revealed that formative assessment should be integrated in terms of language skills. According to the participants, formative assessment in ELT should involve the four language skills. It could be argued that the notion of integration is specific to ELT. As mentioned in the literature, most formative assessment studies were conducted in mathematics and science. The notion that the four language skills should be integrated into the assessment procedure has occurred with the reform of language assessment, in the wake of CLT. Previously, language assessment was compartmentalized, not in terms of skills, rather in terms of discrete language elements such as vocabulary, grammar, or pronunciation. In recent years, however, the notion of CLT has brought changes in the way a second language is taught and assessed, with communicative language testing and integration of the four macro skills (H. D. Brown & Abeywickrama, 2010; Meskill, 2010; Renandya, Lee, Kai Wah, & Jacobs, 1999).

It is somewhat surprising that some participants considered formative assessment may benefit the school as a whole. They argued that if teachers were able to conduct formative assessment effectively, then they would be able to assist students with their learning. This would eventually enhance students’ achievement and overall learning. Further, if teachers are able to report reliable data on students’ achievement, schools will also be more able to
provide relevant learning assistance. Moreover, if students are able to identify their own learning as a result of formative assessment, then schools will be able to provide them with appropriate assistance. In turn, this will enhance the reputation of the school. This is an important issue for future research considering that most of the available studies of formative assessment demand support from the school in order to apply formative assessment effectively. If empirical studies show that formative assessment does benefit schools, there is a strong possibility that this will increase school support for the application of formative assessment.

To sum up, the findings of the present study indicate that participants have diffuse understandings of formative assessment; however, their understandings of its potential benefits and key elements are in agreement with the current literature. In addition, to a certain extent all participants implemented formative assessment strategies, and their practice was often richer than their definitions.

In general, it seems teachers’ understanding of the definition of formative assessment may not relate to their overall understanding of the concept and their practice. Thus, it is important to reconsider whether a narrow-focused definition of formative is still crucial. At this point, the focus of discussion on formative assessment should be shifted from theorizing about it to how can we support teachers in their implementation of it. After all, teachers do not necessarily need to be an expert in the formative assessment field, but they need to have a sound understanding of the basic concept and a broad understanding of the benefits, and how it can improve students’ learning.

5.2 Applying formative use of tests

The second question in this study concerns teachers’ formative assessment practice. Findings from the study revealed that the participants relied heavily on tests; they considered testing to
be a compulsory instrument for their formative assessment practice. For example, Evelyn stated in the interview that when she assessed students’ competence for a unit of a work she might administer various assessment instruments, and testing was always part of it. Wendy did likewise, although she was technology-savvy and integrated educational technology into her assessment practices. My observation data showed that while she administered testing as one of her assessment items, she attained abundant data from other sources.

There are several possible explanations for the teachers’ heavy reliance on testing. The national standardised high-stakes testing has been administered for decades (Section 2.6.2) and may influence teachers’ selections of assessment instruments. Although the government has made significant changes to the role of this national exam, as stipulated by the MOEC Decree Number 23 Year 2016, it is still categorised as high-stakes testing in practice (Ginting & Saukah, 2016). McNamara (2000, p. 48) defined high-stakes testing as ‘a test which provides information on the basis of which significant decisions are made about candidates’. This is indeed the case in Indonesia where high-stakes testing has significant impacts on students’ learning because it determines students’ learning for the next level. Thus, it has shaped the expectations of not only students but also parents and schools. Furthermore, to a certain extent the tests also have significance for teachers and the principal because they reflect on the school’s standing in the community.

Undeniably, testing culture has pervaded classrooms in Indonesia; it has become necessary for teachers to prepare students to pass the national high-stakes test (see Section 2.1 for further details about testing culture). A number of studies document how teachers lack confidence in conducting their own assessment practice, and keep focusing on the national tests (for example, Afrianto, 2009; Azis, 2014; Sulistyo, 2009). Moreover, testing culture has not only affected teachers’ teaching practice but also their beliefs in teaching and learning.
Teachers’ strong reliance on tests cannot be changed in a short period of time (Kennedy, Chan, Fok, & Yu, 2008).

It is also possible that the assessment policy, as stipulated by the MoEC Number 23 Year 2016 on the standard of assessment education (Peraturan Menteri Pendidikan nomor 23, 2016), may have implicitly compelled teachers to administer testing. For example, as can be seen from Table 5.1, there are various tests that students must take throughout their secondary years; these tests are administered by government, schools, and teachers, and while some tests are optional, others are compulsory.

**Table 5.1 Assessment systems in Indonesian secondary schools**

<table>
<thead>
<tr>
<th>Test</th>
<th>Test administrator</th>
<th>Requirement</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN</td>
<td>Government</td>
<td>Compulsory</td>
<td>Once in the 12th grade</td>
</tr>
<tr>
<td>Ujian Sekolah</td>
<td>School</td>
<td>Compulsory</td>
<td>Once in the 12th grade</td>
</tr>
<tr>
<td>UAS</td>
<td>School and/or teachers and/or MGMP</td>
<td>Compulsory</td>
<td>Once every semester</td>
</tr>
<tr>
<td>UTS</td>
<td>Teachers</td>
<td>Compulsory</td>
<td>Once every semester</td>
</tr>
<tr>
<td>Ulangan Harian</td>
<td>Teachers</td>
<td>Optional</td>
<td>It depends on teachers, commonly after they have completed two or three Basic Competences (see Figure 5.1 )</td>
</tr>
</tbody>
</table>

Source MoEC Decree Number 23 Year 2016 (Peraturan Menteri Pendidikan nomor 23, 2016)

In addition, the policy states that schools are required to determine the KKM score; KKM stands for Kriteria Ketuntasan Minimal (Minimal Mastery Criteria). The score for KKM is based on the complexity of learning materials, the students’ intake (students’ quality), and teachers and schools’ support. Schools have the autonomy to decide the score, and this score applies across subjects (Amirono & Daryanto, 2016). Students who have achieved the KKM
score are considered to have met the required standard, and all students must surpass the score. In other words, the benchmark of learning is determined by the KKM score. Testing is the most straightforward way of attaining a score, hence its prominent role in the Indonesian assessment system.

A further possible explanation for the prominent role of testing is that testing suits teachers’ teaching contexts. As has been discussed previously, most of the teachers in Indonesia are teaching large classes, and as Taylor and Nolan (2005) argued, testing can accommodate teachers’ need to attain data on student performance simply and quickly. Nevertheless, teachers’ reliance on testing cannot be seen as an obstacle for formative assessment implementation. Although a number of studies have indicated that there has been a misconception among teachers that they consider testing serves only a summative purpose (H. D. Brown & Abeywickrama, 2010; Davison & Leung, 2009; Wiliam, 2011), formative assessment is not confined to a particular set of assessment instruments (Black & Wiliam, 1998a, 1998b; Pinchock & Brandt, 2009). In other words, testing can be part of formative assessment provided that teachers use the data to improve student learning. As Harlen (2006) argued, several assessment instruments, for example the portfolio, function both as summative and formative. But as Harlen (2006) argued, it is unlikely that tests have both of these functions, particularly as tests encourage teachers to teach for the tests (Popham, 2001). Consequently, Harlen (2006) suggested that in order to make tests reliable as formative assessment instruments, additional assessment instruments should be included in the process. In other words, it is recommended that teachers should not only rely on tests when they assess students’ competence for formative purposes.

The current study found that teachers used tests as part of their formative assessment practice. Several participants used summative tests from UN or UAS for formative purposes. For
instance, Nelly readministered the tests in the class and asked students to reconsider their answers from the first test. Students were actually learning because they wanted to make sure that they gave the right answer. Data from the classroom observations showed that by asking the students to do the same tests again, they were learning collaboratively as they negotiated their answers with peers.

Classroom tests were also used for formative purposes, such as when Brenda administered a test at the beginning of a school year in order to identify what students remembered from the previous school year. This enabled her to categorise students based on their competence, and focus her efforts on students who particularly needed help.

Therefore, the present results are significant in at least two major respects. First, they corroborate the idea that any assessment instrument can be both summative and formative, and second, these results further support the idea that summative tests have the potential to improve learning (Black et al., 2003; Broadbent et al., 2017; Carless, 2011; Lam, 2013; Xiao, 2017).

Nevertheless, findings from the document data showed that teachers find writing reliable and valid tests to be a challenge. Detailed analysis of the test documents showed that the majority of the tests administered by the participants still cannot be considered sound assessment instruments, especially in terms of validity. For example, one of the tests measured students’ speaking skill through writing; several reading tests focused on students’ general knowledge instead of students’ reading skills; and most of these tests did not include any listening skills. Furthermore, several grammatical mistakes, misspelling of vocabulary, and ambiguous questions were found in the participants’ tests. Moreover, it is difficult to determine whether the tests actually assess what the students have been learning as some of these participants also encountered problems in establishing learning objectives (see Section 5.5 for details).
Another issue that needs attention is reliability; as Popham (2003, p. 54) said ‘unreliable tests will rarely yields scores from which valid inferences can be drawn’.

A possible explanation for this might be that the participants lack an understanding of what makes a good test. They commonly reflect on the national examination when designing tests; however, several studies document that the test itself is not a sound assessment instrument (OECD, 2015; Rukmini & Jumaroh, 2015). In addition, findings from the study revealed that the majority of the participants had negative attitudes towards writing tests; saying that writing tests is complicated and time consuming (see Section 4.2.2.2.1 for further details).

This finding is rather disappointing considering that the participants had found a way to apply summative tests formatively, but then lacked the capacity to write quality tests. As Taylor and Nolen (2005) point out, in order to arrive at an accurately determined assessment what students have learned, it is essential for test items to be well-written. Similarly, Schneider and Randel (2010) emphasised that one of the requirements for effectively use FUST and classroom tests is that the test should be well-designed.

Furthermore, this study found that despite teachers’ efforts to use the summative and classroom tests for formative purposes, the testing created anxiety amongst the students. Data from classroom observations showed that students’ anxiety was evident even before the tests began, and also when the participants announced the result of the tests. This condition prevailed regardless of the students’ level of competence. For example, in Wendy’s class, whose students were high achievers, they looked anxious when Wendy administered tests, despite the fact that Wendy had assigned several assessment instruments which could be considered as more challenging than answering questions in a test. Cizek and Burg (2006, p. 6) considered anxiety to be a common condition during test administration and stated that ‘test anxiety has existed as long as there have been tests’. On the other hand, it is important
to note that, a number of studies mention that facilitative anxiety may actually support learning (Brown, 2000; Oxford & Ehrman, 1995).

The data from the study also revealed that the participants had a tendency to teach to the tests. The majority of participants would engage students in the lesson by stating that the materials were going to be on the tests, as occurred in Alfred’s class. He introduced the focus of the session as follows:

Alfred: You will have English text in almost every test, UN, UH, UKK, UAS, and also later in college. And the things that you need to understand, one question that always asked is main idea. Jadi akan ada sebuah pertanyaan yang selalu ditanyakan, yaitu [so, there will be a question that commonly asked, that is] main idea ‘what is the idea of the text?’ so today we are learning about main idea, later we will discuss about descriptive text. Raise your hand if you have ever heard main idea, yang pernah dengar kata main idea angkat tangan [those who have heard the word ‘main idea’ raise your hand].

A further drawback of testing, which was observed in this study, is that some participants relied on the test score as feedback for students (see section 4.2.3.2 for further details). They claimed that they employed a score as feedback because it is easily understood by students and even parents. This finding is contrary to Black et al.’s (2003) study which suggested that relying on scores as feedback is unhelpful because a single score does not help students to understand which areas of learning they need to improve. Hence, Brookhart (2013) argued that a score can be effectively used as feedback as long as it is supplemented by a description of what it means, such as in a detailed rubric, so that students can identify their strengths and weaknesses.

In summary, this study found that tests have the potential to be used to improve learning; in fact, summative tests could also be used formatively. Nevertheless, there are several issues that need to be taken into consideration, especially test design, students’ anxiety, teaching to the tests, and using scores as the sole feedback. Consequently, further research should be undertaken to investigate how to reduce students’ anxiety over tests. As has been mentioned
in several studies, students’ anxiety may affect their performance on tests, causing inaccurate assessment. Another issue that needs to be investigated relates to teachers’ understanding of quality tests. Carter (1984) conducted a study on this specific issue but to my knowledge there have been none since.

5.3 Building formative assessment culture

As mentioned in the literature review, formative assessment fits properly within socio-constructivist learning theory, which highlights dialogue or interaction with surroundings as valuable learning resources. As Hickey (2015) and Stobart (2008) maintain, formative assessment occurs when teachers, students, and peers interact. Since every learning situation has features distinct to the students, teachers and schools, their interactions would be different in each context. Likewise, formative assessment practice may be implemented differently from one context to another, building a unique formative assessment culture.

The results of this study show that participants had their own repertoire of formative assessment instruments and strategies. Interestingly, further analysis of the data showed that although the participants applied the same assessment instrument, they often had different procedures, during the implementation. For example, both Anne and Brenda applied portfolio as their formative assessment instrument, but they implemented them differently. Anne preferred to ask for an online portfolio so that it would be accessible at any time, and so that she can show it to other students as an example. In comparison, Brenda divided the portfolio that she applied in her class into three categories: students’ portfolio, teachers’ portfolio, and open access portfolio which she put in the library to provide a model for other students. Thus, the same instrument with the same purpose can be implemented differently in different contexts.
These results are likely to be related to the participants’ beliefs about learning which were reflected in their formative assessment practices. Beliefs are defined by Borg (2001, p. 186) as ‘a proposition which may be consciously or unconsciously held, is evaluative in that it is accepted as true by the individual, and is therefore imbued with emotive commitment; further, it serves as a guide to thought and behaviour’. Thus, Borg (2001, p. 187) defined teachers’ beliefs as ‘teachers’ pedagogical beliefs, or those beliefs of relevance to an individual teaching’. A number of studies point out that teachers’ beliefs about learning have profound effects on their assessment practice (for example, Guskey, 2003; Song & Koh, 2010), and this is evident in the current study.

Brenda and Alfred are two out of the eight participants who indicated that they strongly believed in behaviourist learning theory. Brenda believed that learning mastery is central for students’ successful learning. She gave an example in the interview:

Actually, we expect the students to be at (level) 10 for example, but we cannot immediately give 8-10 because there are some students who were low achievers and needed assistance. To scaffold these students. Treat them equally first, all students required basic capacity, so to this extent we build knowledge of field, up to level 10 later … at first (learning) was simple, ‘you listen, fill in the blank’, then you transform from oral to written text, higher level, because it is more challenging to transform oral text, it is more difficult to ‘you analyse’ (Brenda/Int2/[19:17])

These beliefs were well-translated into her formative assessment practice as shown from the observation and document data. Brenda designed her learning activities from simple to complex; for example, she began with vocabulary level and moved to discourse level. In addition, Brenda always noted a score for each activity in the class; thus, in one session she could have four to six scores for each student. In the interview, she stated that she needed the scores to identify whether students have understood the learning materials before she continued the lesson.
From a similar position, Alfred believed that repetitive activities are central for learning. Further, he pointed out that repetitive activities are quintessential for formative assessment, as he said in the interview:

It’s really, really important, because, honestly formative assessment must be done. If it’s not done, then we can’t control and see students’ competency. And it needs to be done continuously and repeatedly, because I myself understand formative assessment is done repeatedly so students will understand it. I must emphasise that if it’s done repeatedly they’ll understand the concepts, they’ll understand the strategies and will understand the presented material. (Alfred/I1/[07.11])

As in Brenda’s case, Alfred’s beliefs about repetitive learning were also reflected in his teaching. In every session, Alfred had a ritual of giving students a set of learning activities: he began with checking students’ attendance while he was asking about their reading journal. This activity was followed by public speaking and then free writing before Alfred started on the material for that particular day (see Appendix 11 for details about public speaking and free writing activities). He conducted these activities in every session; in fact, he conducted similar activities in every grade with some teaching modifications, according to the students’ level.

Another example of the teaching beliefs which were translated into teaching practice was reflected in Wendy’s case. Wendy believed that fluency in communication is a crucial element of language competence; thus, she designed learning activities that enabled students to interact in the target language. Since technology can facilitate language learning, Wendy commonly integrated educational technology in her class. She recorded video conferencing and team blogging online, so that assessment could be done at any time, and the recordings could also be used as examples for other students.

One of the issues that emerges from these findings is that teachers’ beliefs have a pivotal role in the efficacy of education policy and in particular on the assessment culture. Thus, it is important to ascertain their beliefs prior to releasing new policy. As Mattese, Griesdorn, and
Edelson (2002) pointed out, teachers may be resistant to change due to differences between the new policy and their assessment beliefs. In other words, when an assessment policy is ‘imposed’ on teachers, there are two important notions that need to be taken into account. First, it has to take into account teachers’ beliefs so that there is flexibility to a certain extent and it is comprehensible to the teachers. Second, for teachers to change their practices, they must ‘learn’ and ‘unlearn’; thus, the process may need time to take effect.

It is possible that the government in Indonesia could identify teachers’ beliefs in assessment, commonly known as teachers’ conception (Opre, 2015), by implementing TCoA (Teachers’ Conception of Assessment). This instrument, which was developed by Brown (2006), is an inventory statement of the main purposes in educational assessment. Teachers are asked to state whether they agree or disagree with several statements which are categorised into four sections. First is the concept regarding the role of assessment in improving learning, which emphasises that any assessment-related activity aims to enhance learning. Teachers who have this conception would understand the importance of constructing an accurate portrayal of students’ performance which they would also use to improve their own teaching. The second is the concept of school accountability. Teachers who have the conception that assessment is important for a school’s accountability will have two rationales: demonstrating school and teacher quality instruction, and improving the quality of instruction. The third conception is student accountability. Teachers who hold this conception consider the role of assessment is to determine student accountability; that is whether or not the students have been studying diligently. This is manifested in the form of scores, grades, or certificates. The last conception is that assessment is irrelevant. This concept holds that assessment should not be conducted for various reasons such as intervening with teacher autonomy or narrowing their instruction. Teachers may also argue that assessment is neither valid nor reliable.
Besides teachers’ beliefs and/or conception of assessment, my findings also note that students may contribute to building a formative assessment culture. For example, the findings of the study revealed that peer involvement in formative assessment can be more effective among female students. This is evident based on the data from classroom observation in Evelyn’s class. Evelyn is teaching in a public school under MoRA’s management. This school commonly separates male and female students: they do not sit together and are not put in the same group. Findings from the observation data showed that there were some differences between males and females when they were working in the group. Evelyn assigned the students to work in groups, and they were asked to write about their experience on the last school excursion. In the female group, the girls discussed and divided the assignment among members of the group; afterwards, they revised, clarified, and negotiated what they were going to write on the cardboard. In contrast, the boys divided the assignment amongst group members, and then they did their individual part without discussion or negotiation during the process.

This finding supports the idea that male and female students have different learning styles and, therefore, may contribute differently to the formative assessment culture. In a study by Dundas (2004), involving students from two single-gender schools, data from questionnaires, classroom observations, and teacher interviews showed that female students preferred collaborative work more than male students. In a similar sense, Shuhib and Azizan (2015) reported that male students were more task-oriented, while female students had more interpersonally-oriented learning styles. It is possible that learning style differences between male and female students may influence students’ involvement in formative assessment, as shown by Takeda and Homberg’s (2014) study. This study focused on group work process and performance using the self- and peer assessment results of 1,001 students in British higher education. Results of the study showed that students in gender balanced groups
displayed enhanced collaboration in the group work process. In addition, this research also indicated underperformance by all-male groups and reduced collaborative behaviours by solo males in groups.

Another example of how students may contribute to a culture of formative assessment can be found in Wendy’s case. This study showed that in her high achiever class, these students commonly self-assessed themselves without prompting by the teacher. This was particularly so when the activity involved giving presentations to the class, as shown by the observation data. Students were highly involved in the process of both self-assessment and peer assessment, commenting constructively on each other’s performance. In comparison, in other participants’ classes, students’ involvement was not as abundant as in Wendy’s class. In fact, in some cases, when teachers asked students’ opinions, they remained silent. This finding is in accord with previous studies showing that there are some differences between high-achievers and low-achievers in terms of self-assessment. High-achieving students tend to underrate themselves while low-achieving students tend to overrate themselves (Boud & Falchikov, 1995b; Dunning, Heath, & Suls, 2004; Landrum, 1999; Lejk & Wyvill, 2001).

Overall, the results of the research corroborate the ideas of a number of studies which previously suggested that formative assessment is a context-sensitive, social practice (for example, Black & Wiliam, 2006; Black & Wiliam, 1998a, 1998b; Cowie & Bell, 1999; Crossouard & Pryor, 2012; Earl, 2013; Torrance & Pryor, 1998).

In another study, Dargusch (2010, p. 47) who investigated the nature and function of formative assessment practice in Queensland teachers of year 12 English, concluded that his study supported the idea of ‘formative assessment is a socio-cultural practice informed by a range of interwoven, situated relationships from three interrelated contexts: system, school, and classroom’.
5.4 Overcoming formative assessment barriers

5.4.1 Identifying the actual formative assessment barriers

This study found that teachers considered there to be two barriers to the implementation of formative assessment, that is large class sizes and limited time allocation. According to the MoEC Decree Number 129/a/U/2004 (Peraturan Menteri Pendidikan dan Kebudayaan Nomor 129a/U/2004, 2004), classrooms in Indonesia may consist of 30 to 40 students. This number of students can be challenging for ELT teachers because they are required to provide opportunities for students to practice the target language. The classroom can be the only place where students use the language. According to Richards (2017), the number of students in an ELT classroom should not exceed 15 in order for them to get the maximum benefit from learning the language. The majority of the participants claimed that large class sizes made it difficult for them to provide personal feedback or to manage the classroom during peer learning.

Although this finding seems to be consistent with an OECD (2005) study which revealed that large classes present a barrier for implementing formative assessment, it is important to note that there have been differing opinions about whether large classes disadvantage learning (Locastro, 1989; Schanzenbach, 2014; P. Smith, Molnar, & Zahorik, 2003; L. Xu, 2012; Zyngier, 2014). It can thus be suggested that the finding of this study relates more to the participants’ negative perceptions towards large classes. As Xu (2001) reported, language teachers often have negative perceptions towards large classes although they are a common phenomenon in foreign language education (Chung & Huang, 2009; Tudor, 2001) as in the Indonesian context (Ariatna, 2016).
Furthermore, contrary to the data from interviews, data from classroom observations showed that large classes did not appear to be a barrier in implementing formative assessment. For example, Alfred was more able to give immediate feedback on students’ essays although he has considerably more students than Wendy. A possible explanation for this might be that Alfred could find a way to provide feedback so that it became available for students at any time. Similarly, Brenda’s class was as large as Nelly’s; however, based on observation, Brenda’s class was better organized, and she designed her classroom activities in such a way that by the end of the lesson, she had a variety of evidence of students’ learning achievement.

This study found that teachers considered the time allocation as stipulated by the curriculum as insufficient for them to conduct formative assessment. As a result, one of the participants claimed that she could no longer implement several of the assessment instruments (see Section 4.3.1 for details). This result seemed to be consistent with another study which found that any assessment practice requiring teachers to invest time and materials discouraged them from implementing the practice (Young & Jackman, 2014).

Nevertheless, it seems possible that this result is merely a response to the recently changed curriculum policy which decreases the time allotted for teaching English from 180 minutes to 90 minutes per week. A deeper analysis of this finding indicates that the challenge may not relate to time, but rather to teachers’ reluctance to make teaching adjustments. Studies in other contexts where they have more time allotted compared to the Indonesian context also mentioned that time is one of the barriers to implementing formative assessment (for example, OECD, 2005; Quyen & Khairani, 2016). One possible implication of this is that teachers should find a way to incorporate formative assessment into their practices regardless of how much time is provided.
Another significant challenge to the implementation of formative assessment is negative perceptions towards it. Daisy and Adam stated that formative assessment is complicated and difficult to implement. Further, these participants explained that it was difficult for them to provide descriptions of students’ achievement as stipulated in the current curriculum due to a lack of facilities and, again, the large number of students. In addition, they also stated that the current scoring and reporting systems can be troublesome for them, not to mention that parents may not understand the systems.

This result is likely to be related to the implementation of the Kurikulum 2013 which was hastily disseminated (OECD, 2015; Zuhdi, 2015). Teachers were not ready to implement the curriculum; especially when the curriculum introduced several concepts that were considered to be novel by teachers such as authentic assessment, peer assessment, and self-assessment. Further, these findings confirm Volante and Beckett’s (2011) argument that teachers may develop negative attitudes towards classroom activities which they consider to be complicated, especially when they are required to implement something that is different from their current practice and means that they need to adjust their practice and beliefs (Bullock, 2011; Pham & Renshaw, 2015).

Interestingly, this study did not find UN, the high-stakes standardised national examination, to be one of the factors that hampered formative assessment. In fact, several participants considered UN an essential part of learning, as Brenda said in the interview:

> Because we are constrained by SKL (the graduate competency standard), our tool to measure success is their ability to manage the graduate competencies. For example, the students need to understand several types of texts, their functions, and their purposes. It’s explicitly stated. Because of the orientation on graduate competencies, we automatically steer toward it. It doesn’t mean teachers aren’t creative, but adjust to the requirements of the competency. UN (National Examinations) are the government’s way of measuring whether someone is competent or not. (Brenda/I2/[04:55])
The study’s findings also revealed that the UN played another prominent role, attracting students’ attention and focussing their learning, an effect observed not only for the 12th grade students, who are going to face the UN, but also on students at lower levels.

This finding is contrary to previous studies which have documented the negative backwash effects of high-stakes tests in education in general, in ELT contexts, and also in the context of ELT in Indonesia. For example, cheating, anxiety, and teaching to the test are some of the negative impacts of high-stakes tests (Greenstein, 2010; Popham, 2006; Stiggins et al., 2004). In addition, a number of studies document that high-stakes tests hamper the implementation of formative assessment (for example, Black, 2015; Box et al., 2015; Gu, 2014; Quyen & Khairani, 2016; Umer et al., 2013); as Black (2015) stated high-stakes tests ‘overshadow’ formative assessment. Studies in the ELT context have also documented high-stakes testing as a barrier for formative assessment (for example, Chris Davison & Leung, 2009; Leung & Rea-Dickins, 2007; Rea-Dickins, 2006). In the Indonesian context, the UN, as a high-stakes test, has also been shown to be a barrier for learning (Furaidah et al., 2015; OECD, 2015; Sukyadi & Mardiani, 2011; Zuhdi, 2015; Zulfikar, 2009). These studies do not directly relate to formative assessment.

The findings of this study suggest that high-stakes tests and tests in general are a powerful force for learning even compared to the curriculum. According to a study conducted by Gu (2014), curriculum change may not be significant in the face of high-stakes testing. It can thus be suggested that UN could be used as a potential means to improve learning. The present study raises the question of whether it would be possible in Indonesia to design reliable national standardised testing which is properly aligned with the national curriculum. This would accommodate teachers who have a strong reliance on the national testing system but still benefit learning. Furthermore, a sound national standardised testing system which is
valid and reliable could provide an opportunity for formative and summative assessment to be aligned. Central to the issue of making use of UN as learning media is the idea that high-stakes testing should be relevant to the curriculum.

Taken together, the results of this study suggest that the challenges encountered by participants can be overcome if they have adequate assessment literacy and a positive attitude to formative assessment. For example, formative assessment can be implemented with large classes by applying peer assessment and self-assessment. Simply put, assessment literacy is the knowledge and skills to perform sound assessment, followed by relevant teaching adjustment (Boyles, 2006; Fulcher, 2012; Popham, 2011; Stiggins, 1999; Webb, 2002). This competence is crucial for implementing formative assessment because it is not just a particular set of assessment instruments. Formative assessment requires teachers’ expertise to orchestrate assessment instruments; it requires student involvement, and teaching refinement. They can only perform this role if they have an adequate level of assessment literacy (Lees & Anderson, 2015).

There seems to be a need for measuring teachers’ assessment literacy in Indonesia. A number of studies document instruments that can be applied to measure teachers’ assessment literacy, namely the teacher assessment literacy questionnaire (TALQ, Plake, et al, 1993); classroom assessment literacy Inventory (CALI, Mertler, 2004); assessment literacy inventory (ALI, Mertler & Campbell, 2005); and teachers’ conceptions of assessment (TCoA, Brown, 2006). In addition, Listiani (2014) has developed an instrument to assess teacher assessment literacy, entitled Pedagogy of Science Teaching Test (POSTT): this instrument is designed to measure pre-service teacher assessment orientation. In fact, Arrafii and Sumarni (2018) developed an instrument entitled Teacher Formative Assessment Literacy
Questionnaire (TFALQ) which is specifically designed to measure teachers’ formative assessment literacy.

Further study should be undertaken to investigate the feasibility of administering one of those instruments, or to develop assessment literacy instruments based on the MoEC Decree Number 20 Year 2007 on Teacher Standard Qualification, which includes assessment literacy as one of its requirements (Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007, 2007). Measuring teachers’ assessment literacy is essential because it provides valuable evidence that can be followed up on in professional development programs for teachers.

5.4.2 Providing solutions

The results of this study show that teaching innovation, such as integrating educational technology, was applied by participants to deal with challenges in implementing formative assessment. In addition, technology also addressed teachers’ problems with record keeping because some of the technology-enhanced activities were automatically recorded and stored online; thus, they were always accessible for teachers.

Educational technology fits properly into several formative assessment strategies for the following reasons. First, it can be applied to provide feedback which can move learning forward. For example, Alfred preferred to use storage files online, and asked his students to submit their essays online so that he would be able to assess the essays on his iPad, anywhere and anytime. Anne also made use of WritetoLearn, a software program that can provide feedback for students’ essays in terms of vocabulary and grammar. This enabled Anne to focus more on the essay content, which might not otherwise have been possible due to the large number of students in her class. Wendy also integrated educational technology to provide feedback for her students. She used Quipper, an online learning tool that provides a wide array of tests written by school teachers. Using this website enabled Wendy to avoid
writing tests, a task which is complicated and time-consuming. It also allowed students to
take the test and get immediate feedback based on their learning pace. Nevertheless,
although these tests were written by school teachers, it needs further study whether these tests
are of good quality and in line with the curriculum.

It is possible that integrating educational technology could be used in another assessment
strategy, that is students’ involvement. For example, Wendy applied video conferencing via
the FacetoFaith community which enabled students to communicate with people from
different parts of the world. She recorded the activity and replayed it, asking her students to
assess their own performance and that of their peers during the activity. She also used team
blogging in which students worked in groups to write and post their essays on blogs; thus,
students from other parts of the world were able to comment on each others’ blogs.

These results further support the idea that educational technology has the potential to
overcome problems in learning, including the implementation of formative assessment. A
number of studies document the effective role of technology in overcoming problems in
education (for example, Beatty & Gerace, 2009; Hargreaves, 2003; Russel, 2010). One way
to integrate educational technology is by applying tools which can be used for learning such
as Classroom Response Systems (CRS) (Feldman & Capobianco, 2008; White, Syncox, &
Alter, 2011). This device can accommodate teachers’ needs in eliciting students’ knowledge
and thus enables teachers to create relevant teaching adjustments.

The use of technology is ‘undeniable and widespread’ (Trucano, 2012, p. 101); it has become
part of literally everybody’s life including teachers and students. Further, the Indonesian
education system has also mandated the integration of educational technology in the
Kurikulum 2013. The findings of the study showed that although there were some
participants who made use of technology effectively so that it can enhance learning, several
other participants had not yet implemented it. This was despite their schools having provided devices such as projectors, televisions, or language laboratories. Integrating educational technology requires more than just providing the devices; teachers also need guidance on how to make the best use of the technology (Cuban, 2003; Falchikov & Thompson, 2008; Wang & Reeves, 2008). Wendy who has a strong belief that learning a language means enabling students to communicate in authentic settings had found that technology accommodates her beliefs, and she plans to integrate technology into most of her teaching activities. For example, she would ask students to capture a paragraph using their mobiles, or send the file instead of providing students with a photocopy of the passage. However, while Alfred who asked his students to submit their assignments online, Wendy did not because of her small class size. In contrast, some other participants also integrated technology, but only as a ‘garnish’ for their class. For example, Daisy asked her students to study in the language lab because the session was supposed to focus on practicing listening and speaking skills; however, she focused her teaching on the grammatical point from the materials and only used the recording once during the class.

This study found that participants modified their teaching practice in order to overcome problems associated with implementing formative assessment. For example, Brenda applied ‘three-lined-writing’ in which she asked her students to use different colours as they wrote essays, and to use three lines for their writing. The first line was the original students’ writing, the second line for her feedback correction, and the last one is for the final version after receiving feedback. By doing so, Brenda was able to monitor students’ progress in writing the essay. Another teaching innovation involved categorizing students at the beginning of a school year, allowing participants to know which group of students required further help. Integrating language skills into one learning activity is another of the solutions
that teachers used in when performing formative assessment. By doing so they were able to assess students’ learning progress, across different skills, in one learning activity.

These findings corroborate the ideas of Black and Wiliam (1998b, 2009) who suggested that teachers need to incorporate formative assessment practice into their own teaching contexts. It is interesting to note that participants who came up with ideas for teaching innovations were not the ones who thought that formative assessment was complicated and wanted assessment practice to be simpler. Instead, they were the teachers who believed that they needed to make teaching innovations in order to accommodate students’ needs. As Alfred said ‘I need to find innovations for my teaching, and this cannot be solved at once -it takes time until I’m able to find effective activities for my teaching’.

It is somewhat surprising that participants did not refer to teacher professional development programs as a means to overcome problems in implementing formative assessment although this is a common recommendation in the literature (McGatha et al., 2009; Randel et al., 2016; Trumbull & Lash, 2013). In addition, several professional development programs for teachers have been documented to improve teacher formative assessment literacy and replicated in other contexts, for example keeping learning on track (KLT) and classroom assessment for student learning (CASL) (Stiggins et al., 2006; Thompson & Wiliam, 2008).

A possible explanation for this might be that assessment has not yet become a focus of teacher professional development in Indonesia. Rahman (2016) stated that teacher professional development for teachers in Indonesia is mostly used to disseminate new curricula. Assessment is commonly addressed as part of a bigger program rather than being the main concern. Further, Rahman (2016) characterized teacher professional development in Indonesia as a program that commonly involves a great number of teachers, but is conducted in a limited time. This is not consistent with Lee and Wiliam’s (2005) argument
that to effectively improve teaching, professional development needs to be well-structured to not only enable teacher growth, but should be flexible so that teachers can mold their practice to fit their teaching contexts. Further, professional development that is concerned with improving teacher formative assessment literacy is commonly conducted over a long period (Stiggins et al., 2004; Thompson & Wiliam, 2008), but in Indonesia, professional development programs are usually conducted as short-term workshops (Rahman, 2016).

This section has discussed challenges which are encountered by teachers during the process of implementing formative assessment and how teachers cope with these challenges. This study found that teachers considered large classes and limited time allocation as barriers for formative assessment implementation. Further analysis strongly indicates that these barriers relate to their assessment literacy. Teachers with an insufficient level of assessment literacy can find implementing formative assessment very challenging. Consequently, further work is required to establish the interrelationship between assessment literacy and formative assessment. I also discussed the role of educational technology in overcoming barriers to formative assessment; however, its effectiveness was very much dependent on the particular context.

5.5 Missing formative assessment opportunities

As has been discussed in the previous section, a range of formative assessment strategies were implemented by the participants in this study. The findings from the interview data in this study showed that the participants believed their selection of assessment instruments and the feedback that they provided for their students were effective and contributed positively to students’ learning. Nevertheless, despite the participants’ beliefs about the effectiveness of their formative assessment practice, there were several assessment practices that could have been refined in order to maximise the benefits of their practice.
This section discusses several formative assessment instruments and strategies applied by the participants which had the potential to enhance student learning, but were not exploited to their full potential. Thus, it can be said that the participants might have missed the opportunity to capitalise on formative assessment. The missed opportunities were evident when they were applying the formative assessment instruments, questioning and peer assessment, applying the strategy of clarifying, sharing, and understanding learning intentions and successful criteria. The following part of this section moves on to discuss in greater detail how the participants had missed opportunities in three formative assessment practices: questioning, peer assessment, and clarifying learning intentions.

5.5.1 Missing formative assessment opportunities: questioning

This study found that the participants might have missed the opportunity to improve learning when they applied questioning in their classroom. Data from the observations revealed that all participants applied questioning as part of their formative assessment practice; in fact, in general their questioning was both abundant and fast-paced. Questioning was applied during almost all stages of learning, and it served to check on learning, elicit students’ prior knowledge, probe the process of students’ thinking, pose problems, and seek alternative solutions. In addition, data from the classroom observations showed that several participants made adjustments based on the learning evidence gained from questioning.

This finding is in line with a number of studies which promote questioning as a potential assessment tool. These studies suggest that questioning provides several advantages for learning such as allowing the teachers to clarify and expand learning, helping teachers to identify students’ learning, ability to be implemented at any learning stage, and acting as contingent scaffolding (Afflerbach, 2010; Glasgow & Hicks, 2009; Hammond & Gibbons, 2005; Wiliam & Leahy, 2015; Wilson & Devereux, 2014). Therefore, according to Chaudron
(1988 cited in Tan, 2007) questioning is one of the most popular teaching activities. Heritage and Heritage (2013, p. 178) even referred to questioning as the ‘epicentre’ of formative assessment. Questioning supports formative assessment because it has the potential to provide information for teachers to adjust instruction while learning is still in progress. In addition, questioning may enhance learning in the field of ELT because the process of formulating questions and answers is another way of learning the language (Al-Darwish, 2012; Jiang, 2014; Tan, 2007; van Lier, 1988). Another benefit of questioning is that it can encourage critical thinking (Feng, 2013; Shen & Yodkhumlue, 2010).

Surprisingly, findings from the interview data showed that none of the participants considered questioning as an assessment instrument, let alone as part of their formative assessment practice. In other words, although the participants applied and made use of the data from questioning to adjust learning, they did not seem to realize that it was actually a tool to improve student learning. A possible explanation for this might be that questioning had become an inseparable part of their teaching practice. To a certain extent, these findings are in line with the study conducted by Al-Darwish (2012) which revealed that teachers believe questioning is an integral part of teaching, rather than a teaching or assessment method. Another possible explanation might be that studies of questioning have not yet provided sufficient information on how teachers can use it effectively. Studies in this were mostly focused on the types of questions applied by teachers, or the classification of teachers’ questioning techniques (for example, İnan & Fidan, 2013; Tan, 2007; Wong, 2010). Thus, it remains hard for teachers to translate these studies into classroom practice.

These findings are rather disappointing considering that questioning can make a substantial contribution to learning improvement. The benefits of applying questioning can be maximised if the activity is prepared beforehand, by the teachers. For example, teachers can
prepare their questions in advance in order to ensure that questions are not mainly focused on
the IRE (Initiate-Respond-Evaluate) model (Heritage & Heritage, 2013) as was the case for
the majority of the participants in the current study. Consequently, as Jiang (2014)
recommends, teachers should be trained to formulate good questions which are
comprehensible, simple, and content-oriented (Barrel, cited in Ozuem & Lancaster, 2015),
and to avoid questions which focus on lower order cognition, address a limited number of
students, or are aimed at controlling the classroom (Barnette, Orletsky, & Walsh, 1995). In
addition, teachers need to apply adequate wait time to allow students to answer (Hannel,
2009; Wiliam & Leahy, 2015).

5.5.2 Missing formative assessment opportunities: peer assessment

The current study also found that the participants might have missed formative assessment
opportunities when they involved students in peer assessment. The results of the study
showed that peer involvement in the process of formative assessment was evident; the
majority of the participants put students to work in pairs or in groups to have them assist each
other in achieving learning objectives. During the process, students could remind each other
of the quality of good work. As Alfred said in the interview, ‘They are divided into groups so
that each student contributes, each one is teaching the others, they are helping each other and
being helped, two by two’. The findings of the study revealed that peer involvement during
the formative assessment process was implemented through a number of strategies such as
peer learning, peer tutoring, or peer counselling. This finding is in accord with several
studies that mentioned peer assessment as the most common form of peer involvement
(Panadero, Johnson et al., 2016; Topping, 2005; van Gennip et al., 2010). However, as
discussed in Section 4.2.4, most of the peer assessment techniques applied in the participants’
classrooms were unplanned or unprepared. For example, half of the participants asked vague
questions such as ‘What do you think of your friend’s performance?’ after one of the students
had given a presentation or role play. Only Brenda and Evelyn specifically prepared students for peer assessment, while other participants implemented peer assessment, but did not prepare students for the activity by establishing the criteria or by preparing a peer assessment feedback sheet.

This finding adds further support to previous studies which have suggested that teachers need to prepare students prior to implementing peer assessment (for example, Tsivitanidou, Zacharia, & Hovardas, 2011; Wiliam, 1992). The preparation should include three skills involved in the process of peer assessment, as pointed out by Sluijsmans, Brand-Gruwel, van Merriënboer and Martens (2004). First, students should be able to define the learning criteria; thus, assessment criteria should be discussed prior to assigning students to assess their peers. Second, students should be able to judge the performance of a peer based on the criteria. Third, they should be able to provide feedback for peers. The third skill can be the most important one in peer assessment because peer feedback is considered to be the main benefit of peer assessment. As Carless (2012) points out, peer feedback can be more significant than teachers’ feedback because it offers benefits such as using language which is more comprehensible for students and it is abundant, whereas teachers’ feedback may be limited due to the number of students.

A possible explanation for teachers lack of preparation in applying peer assessment might be that teachers still believe that assessment is the teacher’s full responsibility; thus, they may not consider the activity seriously. As has been discussed in Chapter Two, classrooms in Indonesia are dominated by a teacher-centred approach (OECD, 2015; Zuhdi, 2015; Zulfikar, 2009). Consequently, the data showed that students’ involvement was only peripheral to the process of formative assessment.
Preparing students with the skills mentioned above is quintessential for effective peer assessment, or else the activity may become counterproductive for students. For example, as reported by several studies, peer assessment may be less effective because students have the tendency to disregard their peers’ feedback (Gielen et al., 2010; Yang et al., 2006); also students may lack confidence to assess their peers’ work (Mik, 2011). Although peer assessment is primarily ‘learning by assessing’ (Topping, 1998), the process needs to be carried out in line with several theories of effective peer assessment (Panadero et al., 2016). The findings of the study revealed that the participants understood the importance of student involvement in the process of formative assessment; however, they could still improve the way they administer the assessment instrument so that they do not miss formative assessment opportunities to improve learning.

5.5.3 Missing formative assessment opportunities: clarifying, sharing, and understanding evidence of learning and success criteria

The results of this study also show that the participants missed opportunities to improve learning when they applied one of the strategies of formative assessment; that is, clarifying, sharing, and understanding evidence of learning and success criteria. As elaborated in the previous chapter, all participants shared their learning objectives to the students in every session; some even established assessment criteria or rubrics for their students. Nevertheless, further analysis of the data showed that the learning objectives were generally simply stated, but not clarified or explained. For example, this is the way Daisy stated the learning objective: ‘Today, we are going to learn about descriptive texts’. This learning objective might be ambiguous for students, and possibly, they could interpret it in two different ways: first, that they were going to learn how to understand the structure of descriptive texts, and later be assessed based on their reading comprehension; or second, they were going to learn about descriptive texts because they would be required to write descriptive essays. Wiliam
and Leahy (2015) referred to this type of learning objective as a learning intention; while Tuttle (2014) referred to this as a learning agenda. Learning intentions or a learning agenda may not provide students with sufficient information of what is expected from them. Wiliam and Leahy (2015, p. 32) point out that ‘The only useful learning is that which the student can apply beyond the context of the learning’. Hence, Wiliam and Leahy (2015) suggested that teachers should separate the learning intention from the context of learning, and also highlighted potential contexts in which the learning intention can be applied, as they illustrated in the table below:

### Table 5.2 Confused and clarified learning intentions and success criteria

<table>
<thead>
<tr>
<th>Confused learning intention</th>
<th>Clarified learning intention</th>
<th>Context of learning</th>
<th>Potential new context</th>
</tr>
</thead>
<tbody>
<tr>
<td>To be able to write instructions on how to change a bicycle tyre</td>
<td>To be able to write instructions</td>
<td>Changing a bicycle tyre</td>
<td>Writing instructions on how to make a sandwich</td>
</tr>
<tr>
<td>To be able to construct arguments for or against assisted suicide</td>
<td>To be able to construct arguments for or against emotionally charged propositions</td>
<td>Assisted suicide</td>
<td>Constructing arguments for or against abortion</td>
</tr>
<tr>
<td>To create a written description of a friend</td>
<td>To be able to create an effective characterisation</td>
<td>Describing a friend</td>
<td>Creating an effective characterisation of a relative</td>
</tr>
</tbody>
</table>

Source: Wiliam and Leahy (2015, p. 35)

Learning intention or learning agenda may not provide sufficient information for students of what is expected as a result of learning. Therefore, it is essential that learning objectives are not only shared, but also clarified and understood, because teachers and students may have different perceptions of the expected learning outcomes (Entwistle & Smith, 2002; Simon & Taylor, 2009). Moreover, a lack of understanding of learning objectives may result in
students’ low achievement (Black & Wiliam, 1998b; McDonald & Boud, 2003).

Consequently, learning objectives should be elaborated and made transparent for students in order to enhance learning, and not merely be stated (Leahy et al., 2005; McManus, 2008; Pinchock & Brandt, 2009).

In fact, if teachers do not clarify learning objectives, providing students with assessment criteria may help learners to understand the learning outcomes. Unfortunately, the findings of the current study document that the assessment criteria and/or rubrics established by the majority of the participants were very broad and might not have helped students to understand the expected learning outcomes. For example, data from the interviews and documents showed that Evelyn assessed students’ essays using criteria such as content, accuracy, and coherence. Although the majority of the participants referred to similar criteria when they talked about the rubrics, they did not elaborate the rubrics with specific measurements.

Interestingly, the current study showed that several participants claimed that they purposefully did not apply rubrics for their assessment. These participants argued that if learning criteria are stated in the rubrics with detailed descriptions of quality work, it may be counterproductive for learning because it may decrease students’ confidence by setting the goal too high. Wendy emphasised in the interview that the main goal of language learning is to enable students to communicate, and this process may be hampered by establishing ‘rigid’ assessment criteria in the rubric. This finding is rather disappointing because, as documented by a number of studies, rubrics can help to make learning objectives more tangible for students (for example, Brookhart, 2013; Lipnevich, McCallen, Miles, & Smith, 2014; Luft, 1999; Smit & Birri, 2014; Wiliam, 2011). Nevertheless, it is important to note that rubrics
may not automatically benefit learning, and they require teachers’ expertise and experience in the implementation process (Jeong, 2015).

Besides lack of detail, the assessment criteria and/or rubrics established by some of the participants included several items that were not related to language learning. Handwriting, confidence, effort, attitude, engagement, and participation were several factors mentioned by the participants when performing formative assessment. Among those factors, effort was the factor that was most emphasized by the participants. This finding is contrary to previous studies which have suggested that teachers’ perceived assessment should exclude students’ effort and/or other such items (Adams, 2005). Nevertheless, I argue that further investigation needs to be carried out to explore whether effort should be part of teachers’ formative assessment practice. According to a study conducted by Isnawati and Saukah (2017), effort is often considered to be one of the factors included in teachers’ assessment. Consequently, there are two reasons why investigation in this area is important: first, it is somewhat unlikely that effort can be identified, let alone measured, and second, students’ efforts may not always be in line with the learning outcomes; thus, effort can only determine students’ learning progress but not the outcomes. Teachers should give objective assessments based on the students’ learning outcomes.

It is somewhat surprising that establishing learning objectives was still challenging for the participants considering that these learning objectives were explicitly stated in Kurikulum 2013. In fact, detailed learning objectives are elaborated in the curriculum so teachers only need to specify these objectives into several operative terms, as illustrated by Figure 5.1.
Figure 5.2. Learning objectives
(Adapted from Peraturan Menteri Pendidikan Nomor 22, 2016)

Based on the above figure, learning objectives are formulated into several terms: graduate standard competence, core competence, and basic competence. Teachers were only required to establish the Indikator Pencapaian Kompetensi (Learning Competence Indicator, hereafter IPK). Among other participants, findings from the study showed that Brenda was the only participants who had successfully implemented the Basic Competence to IPK, documented below (see Appendix 9 for details)
Table 5.3 Sample of learning objective

<table>
<thead>
<tr>
<th>Knowledge Basic Competence</th>
<th>Analysing social function, text structure, and language elements from News items in a simple form from newspaper, radio/TV, in line with its usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievement indicator</td>
<td>1. Copying with neat handwriting several simple news items from newspaper/radio/TV from several sources, using legible spelling and correct punctuation.</td>
</tr>
<tr>
<td></td>
<td>2. Identifying social function, text structure and language elements (including main idea and detailed information) from simple news text of newspaper/TV/radio (news items).</td>
</tr>
<tr>
<td></td>
<td>3. Reading and listening to simple news text from radio (news items) to understand the message</td>
</tr>
<tr>
<td></td>
<td>4. Collaboratively, looking for and gathering several light news items from newspaper/radio/TV.</td>
</tr>
<tr>
<td></td>
<td>5. Thoroughly read several light news items from newspaper/TV/radio (news items) that had been compiled, by identifying and mentioning: social function of each text, genre of the elaborated, name, parts, common behaviour that can be found/seen; vocabularies. Grammar, expressions, tone, spelling, punctuation applied in the text.</td>
</tr>
<tr>
<td></td>
<td>6. Comparing social function, text structures, and language elements form several simple texts from newspaper (news items) that had been compiled from the above mentioned sources.</td>
</tr>
<tr>
<td></td>
<td>7. Acquiring feedback from teachers and friends of the result of analysis about social function, text structures, and language elements which is used in the text from the newspaper (news items).</td>
</tr>
</tbody>
</table>

Brenda was probably be the only participant who did not encounter any difficulty in interpreting and translating learning objectives from the curriculum to the class. As she was one of the Instruktur Nasional (National Instructors) whose assignment was to disseminate Kurikulum 2013; she was thoroughly familiar with the curriculum. This finding indicates that understanding the curriculum is still challenging for Indonesian teachers. There is a tendency for teachers to find establishing learning objectives challenging and they are unsure
how the learning objectives might contribute to learning improvement. In addition, the assessment criteria and rubrics established by the majority of the participants were very broad and included several items that may not be directly related to language learning. These results are likely to be related to the teachers’ abilities to comprehend the curriculum and as Stiggins et al. (2006, p. 78) said ‘curriculum is the first assessment guide’.

These findings, then, raise intriguing questions regarding the role of the curriculum in teachers’ teaching practice; specifically, its role as the ‘road map’ (Heritage, 2007) for learning which contains learning objectives that teachers are required to achieve. This is an important issue for future research into how to make the curriculum function as a guide to learning. In future investigations, it might be possible to explore teachers’ assessment literacy and its correlation with their competence in formulating learning objectives. It is also crucial to study whether providing examples as a supplement to the curriculum will enhance teachers’ curriculum literacy. Another possible topic to be explored, concerns teachers’ competence in formulating assessment criteria and rubrics. Studies in this area mostly discuss the benefits of formulating assessment criteria in learning, whereas in ELT these are commonly focused on inter-rater reliability during summative writing assessment, rather than the importance of assessment criteria in formative assessment. Studies into these issues could shed more light on using formative assessment to enhance learning.

Fortunately, despite the strong indication that some participants have not yet obtained adequate competence in establishing learning objectives, assessment criteria, or rubrics, it is important to note that they clarified learning objectives by giving examples of successful learning to the students. Data from the study showed that the participants applied various strategies when giving examples to students: they provided examples of how to respond to the assignments; they used examples from students at the same level in the previous year;
they used examples from the students’ peers and explain why it was a quality product based on the previously established criteria; and, they made the examples accessible for every student by putting them online or in the school library.

In this section, I discussed how the participants missed formative assessment opportunities in three activities. First, the participants did not consider questioning as part of formative assessment practice despite its rapid application; as a consequence, they had not implemented questioning effectively. Second, the study found that students’ involvement had become one of the participants’ formative assessment practices; however, students were not equipped with guidelines on how to assess their friends. Third, the participants always presented their learning intentions to students; however, they did not clarify them sufficiently.

5.6 Summary
This chapter has discussed the key findings arising from this study. Specifically, I discussed several emerging issues from the key findings such as surviving formative assessment fuzziness, applying formative use of summative tests, building a formative assessment culture, overcoming challenges in implementing formative assessment, and missing formative assessment opportunities. There are several points that can be highlighted from this discussion chapter. First, teachers do not have to be experts in defining formative assessment in order to implement it successfully. As shown by the current study, the understanding of formative assessment by teachers is varied; yet, this does not hamper them from applying the practice in their classroom. Second, formative assessment is contextually situated: that is teachers need to incorporate the practice in a manner relevant to their particular teaching context. There are some factors such as teachers, students, and schools which are distinct to each context and contribute to building a formative assessment culture. This section also showed that the government’s assessment policy participated in shaping the
formative assessment culture. This notion is reflected in the Indonesian context, where the government encourages teachers to implement formative assessment, while insisting that teachers administer various summative tests, some compulsory and some optional. This means that teachers are required to strike a balance between the two assessment streams: formative and summative.

In addition, formative assessment culture was also shaped by the way the participants in this study perceived challenges, and overcame them. Although large classes and limited time were mentioned as barriers to implementing formative assessment, this section discussed the idea that the actual challenge may relate to teachers’ competence in establishing, administering, and interpreting the results of assessment. In other words, barriers to implementing formative assessment may prevail due to teachers’ assessment literacy level. Having adequate assessment literacy will enable teachers to come up with solutions to overcome the challenges. This is crucial because some suggested solutions may not successfully overcome the barriers if teachers do not have the capacity in making use of these strategies. This also implies that teachers need to incorporate solutions relevant to their teaching contexts.

The last issue discussed in this section concerned the missed opportunity of formative assessment. Formative assessment strategies occur throughout learning; however, it does not automatically benefit learning. This study showed that the participants have shown best examples of formative assessment which can be applied by other teachers in similar contexts; nevertheless, some formative assessment practice needs to be refined especially when establishing learning objectives and implementing questioning and peer assessment.

In the following chapter I will make recommendations on how to improve the efficacy of the overall formative assessment practice, conclude the research and discuss its implications.
Chapter 6
Embracing formative assessment

Overall, the study has found that teachers in Indonesia have the potential to implement formative assessment and achieve its potential benefits. Despite many teachers’ reluctance to implement it, the teachers in this study successfully embedded formative assessment into their teaching practice. Formative assessment is indeed challenging for teachers in Indonesia due to large classes, limited time, the testing culture, and the responsibility to fulfil the required administrative work. Nevertheless, this study has shown that although the teachers who participated in this study had differing opinions on how formative assessment should be defined, they understand how formative assessment can be used to improve learning. They also comprehend its benefits and its key elements. Results of the study have shown that teachers are able to put their understandings into practice; in fact, some teachers are setting excellent examples of how formative assessment can be implemented and of how its barriers can be overcome. Hence, generally, this study signifies that formative assessment can be successfully implemented in the Indonesian setting and its potential benefits can be achieved.

This final chapter presents the conclusion of this research journey. I will provide a brief summary of the research and present its theoretical and pedagogical contributions, followed by a discussion of the implications arising from the current research. The following section focuses on limitations of the study and recommendations for future researchers who are working in the same area. The last section is a retrospective that closes the journey of the research.

6.1 Research summary
A number of studies have documented the potential benefits gained from formative assessment including, it improves students’ motivation, informs learning progress, and
encourages self-regulated learning (for example, Andersson & Palm, 2017; Black et al., 2003; Broadbent, Panadero, & Boud, 2017; Chappuis & Chappuis, 2008; Earl, 2013). In response to the benefits offered by formative assessment, the Indonesian government has incorporated formative assessment in its assessment systems. Unfortunately, its implementation was criticized by teachers, in particular, who claimed that formative assessment is challenging, and not feasible in the Indonesian context. This criticism may derive from two factors: the testing culture that has been entrenched in Indonesian education systems, and a lack of resources to educate teachers in how to carry out formative assessment.

In light of the great value of formative assessment, this research has provided inspirational examples about how formative assessment can be implemented so that it can encourage other teachers to adopt or adapt the practice. This study is significant in two major respects: first, it contributes an understanding of the practice to the growing body of literature on formative assessment and can serve as a knowledge base for how to use formative assessment in ELT. Second, it will help teachers in Indonesia understand how formative assessment can be implemented, and eventually be used to improve learning.

Three research questions were formulated to guide the study:

1. How do ELT senior high school teachers in Indonesia perceive formative assessment?
   a. What do they understand by the term formative assessment?
   b. What do they believe are its advantages?
   c. Ideally, how do they think formative assessment should be carried out?

2. How are teachers’ perceptions of formative assessment manifested in their practice of it? What instruments and strategies do teachers use to implement formative assessment?
3. What challenges are encountered by teachers when implementing formative assessment? How do they cope with these challenges?

A multiple case study was employed for this study due to its potential to provide comprehensive information on the topic being discussed (Merriam, 1998; Stake, 2006; Yin, 2003). Eight ELT teachers involved in this study were selected based on the following criteria: they represented one of the two schooling contexts in Indonesia, MoRA and MoEC; they had training and experience in applying formative assessment, and they were recommended by their peers. To gain data, three instruments were used in this study, namely semi-structured interviews, classroom observations, and formative assessment artefacts. Ethics clearance was secured from the Ethics committee of the University of Canberra prior to collecting the data. The process of data analysis for this study is adapted from Creswell (2005, p. 230). The analytical framework drew primarily on Thompson and Wiliam’s (2008) formative assessment strategies:

1. Clarifying, sharing and understanding learning intentions and successful criteria;
2. Engineering effective classroom discussion, questions, and learning tasks;
3. Providing feedback that moves learners forward;
4. Activating students as the owners of their own learning; and
5. Activating students as instructional resources for one another.

I also determined whether teachers used the evidence of learning to adjust their teaching style when necessary; this element is the underlying element that differentiates formative assessment from other assessment practices.
6.2 Summary of research findings

6.2.1 Teachers’ understanding of formative assessment

I found that generally the teachers who participated in the study have a general understanding of the notion of formative assessment. Although there are differing opinions among participants on defining formative assessment, this is understandable, because to date, scholars have not reached an agreement on deciding how best to define it (for example, Bennett, 2011; Crossouard & Pryor, 2012; Kingston & Nash, 2011). In addition, this study has shown that formative assessment has a pivotal role in how these participants conduct their assessment practice. This is because it allows them to make adjustments to modify their teaching methods, which might include remedial teaching or making teaching innovations. This finding implies that these participants understand the key concept of formative assessment: using the results of assessment to move learning forward (for example, Black, 2015; Black & Wiliam, 2009; Pinchock & Brandt, 2009; Stiggins, 2005). Other elements mentioned by the participants are continuous process, feedback, and students’ active involvement. Further, this study has identified another key element that is distinctive to the ELT field, that is the integration of the four macro skills: reading, writing, speaking, and listening.

6.2.2 Formative assessment implementation

Another major finding from the study is that teachers’ understanding of the key feature of formative assessment is manifested in their daily teaching practice. In other words, teachers used the result of assessment to improve student learning. This was evident when they gave feedback to students that aimed to information students how they could improve their performance, instead of merely reporting the results of assessment (see Section 2.2.3 for further details).
One of the most significant findings from this study is that a wide array of formative assessment strategies were applied by the participants, and they provided some excellent examples of how formative assessment can be carried out. Moreover, this study has identified some examples of how teachers cope with testing culture that is entrenched in Indonesian educational systems. Despite the domination of the testing culture, the participants were able to make room for formative assessment. In fact, this study has shown that some of these participants were able to use summative tests for formative purposes. These strategies were influenced by factors such as teachers’ beliefs, students’ characteristics, and school contexts. In other words, these factors contribute to the development of a formative assessment culture, and formative assessment instruments are distinctly implemented in each context. Although it is important to note that there are some assessment instruments which were applied in all settings, namely questioning, tests and observation.

Nevertheless, this research has also shown that although formative assessment occurred very widely, some participants missed opportunities to exploit its potential. Some practices which were implemented by the participants could be refined in order to get the utmost benefit from formative assessment. For example, some of these participants found establishing assessment criteria or using effective questioning to be challenging.

### 6.2.3 Formative assessment challenges and solutions

The study has identified that the participants considered large class sizes and limited time allocation as challenges for the implementation of formative assessment. Some of the participants overcame these challenges by integrating educational technology into their teaching practice. For example, some teachers used devices such as iPads, smartphones or projectors; while others applied software such as Writetolearn, Quipper, Screencast-O-matic or used internet-based services such as Google Drive. Some participants also made
modifications to their teaching such as administering tests at the beginning of a school year or semester, in order to identify which students needed further extra assistance.

The findings strongly indicate that teachers’ level of assessment literacy is more likely to be the reason for teachers’ difficulties in applying formative assessment than large class sizes and inadequate time allocation. Hence, this study suggests that enhancing assessment literacy is the key to effective implementation of formative assessment. In addition, this finding can become an information base from which to conduct further analysis on factors that overshadow formative assessment implementation.

Taken together, the results have shown that EFL teachers in Indonesia can embrace the notion of formative assessment, and integrate it seamlessly into their teaching practice. This study has also found that, although some practices could be enhanced, these participants have shown some excellent examples of formative assessment practice that can be adopted or adapted by other teachers in a similar context. Therefore, teachers in Indonesia have the potential to implement formative assessment and achieve its benefits.

6.3 Research contributions

The present study has addressed the paucity of research in formative assessment in ELT, specifically in the Indonesian context. In addition, the research makes several noteworthy contributions:

First, the present study has enhanced our understanding of the ‘fuzziness’ in defining formative assessment. A number of studies have offered to define formative assessment, yet none of the definitions have been widely accepted (for example, Bennett, 2011; Dunn & Mulvenon, 2009; Gulikers et al., 2013; Harlen, 2006; Kingston & Nash, 2011). Nevertheless, this study has shown that although there are differing definitions of formative assessment, teachers’ understandings of the broad notions of formative assessment were in agreement and
they were able to provide some excellent examples of how to implement formative assessment. Most importantly, the teachers who participated in this study agreed that the purpose of formative assessment was to improve learning. In other words, this study has demonstrated that, despite a wide array of the participants’ understanding in defining formative assessment, there were some common notions of the features of formative assessment such as process, feedback, students’ active involvement, and teaching adjustment. Formative assessment can be recognized according to its key elements and in the way that teachers use evidence of learning.

Second, a key strength of this research is that it presents several inspirational examples of how formative assessment strategies are implemented on a day-to-day basis. Most of the studies in formative assessment implementation concern the following topics: suggestions on how the strategies should be implemented (for example, Chappuis & Chappuis, 2008; Pinchock & Brandt, 2009; Stiggins, Arter, Chappuis, & Chappuis, 2006) and focus on particular formative assessment strategies. Although these suggestions may provide valuable information about the application of formative assessment, they may be insufficient to encourage teachers to apply formative assessment. As Fulcher (2012) and Edelenbos and Kubanek-German (2004) maintain, teachers want clear practical, hands-on activities instead of a set of teaching guidelines. Therefore, this study fills a gap in the literature on formative assessment.

Third, this study confirms previous research that argues that teachers need to find their own ways to incorporate formative assessment practice into their own context (Black & Wiliam, 1998, 2009). This study has demonstrated that formative assessment in ELT has distinct features in terms of theory and practice. For instance, the participants mentioned that one of the key elements in formative assessment is integrating language macro skills.
Hence, the current findings add to a growing body of literature on formative assessment in ELT which is still underexplored; there were only a few studies found in this area. This study focuses on how formative assessment strategies are implemented by ELT teachers and provides a framework for the exploration of formative assessment in ELT.

Fourth, a key strength of this research is that it presents several inspirational examples of how formative assessment is carried out by ELT teachers in Indonesia, including how they overcome what they considered to be barriers in implementing formative assessment. The teachers who participated in this study represented a diverse range of contexts in terms of class size, school’s support, region, and ministry management. In addition, the teachers also had a diversity of educational background, teaching experience, and participation in workshop and training courses. Regardless of their diversity, each of these participants has shown that formative assessment is feasible in his or her own context. In addition, a number of these examples of formative assessment practice can serve as a base for other teachers to adopt or adapt in line with their context. Thus, teachers’ criticisms of formative assessment as complicated and burdensome can be allayed.

6.4 Implications of the study

The findings of this study have a number of important implications for future practice. First, this study has demonstrated that despite the differing understandings of formative assessment, the participants have shown several excellent examples of how formative assessment can be carried out, including how to overcome formative assessment challenges. Therefore, the results of this study suggest that the quest of having an exact definition of formative assessment may not be worth following. There does seem to be a need to enhance teachers’ assessment literacy and shift the focus more onto how to support teachers in order to encourage them to implement formative assessment and gain its potential benefits.

According to Black (2015, p. 176), academics and teacher trainers should build good
communication with teachers in order to introduce learning innovations, such as formative assessment, successfully. As Black (2015, p. 176) points out, ‘Academics in education have to find ways to build fruitful interactions between their world and the world of practising teachers if they are ambitious to explore, and to learn how to implement, the potential benefits of their work.’

Second, these findings suggest that more studies need to be carried out in the particular learning area to enrich teachers’ knowledge of formative assessment. The current study confirmed Rea-Dickins’ (2006) findings that formative assessment opportunities occur throughout learning; nevertheless, this study has found that some participants missed such opportunities. An implication of this is that teachers need to refine their practice in order to get the utmost benefit from formative assessment. A number of studies document the canonical role of teacher professional development programs to disseminate the notion of formative assessment (Andersson & Palm, 2017; Black et al., 2003; DiBiase, 2014; McGatha, Bush, & Rakes, 2009; Randel, Apthorp, Beesley, Clark, & Wang, 2016; Thum, Tarasawa, Hegedus, Yun, & Bowe, 2015). In Indonesia, there are two organizations that can perform the role of improving teachers’ formative assessment practice: first is P4TK (Pusat Pengembangan dan Pemberdayaan Pendidik dan Tenaga Kerja, centres for professional development and empowerment of teachers and educational personnel) as the national teacher professional development centre and MGMP/KKM as the regional centres for teacher learning. Reflecting previous studies such as Black et al. (2003), these programs should be conducted over a period of time that is sufficient for teachers to acquire the theory, and to practise it in their context. Further, it should allow teachers some flexibility in implementing the practice because essentially, formative assessment is contextually situated. Besides P4TK and MGMP/KKM, teacher training also has a significant role in enhancing teachers’ formative assessment practice. For example, the institution should place more emphasis on
the theory and practice of formative assessment for its future teachers, instead of only focusing on how to construct quality tests. Furthermore, important changes should also be made by teacher training programs in order to improve teachers’ formative assessment practice. Most of teacher training programs, especially in Indonesia, use the name ‘language testing and assessment’ and teach future teachers how to assess students. The name of the course implies that assessment is given equal weighting with tests; whereas tests are only part of assessment (see Figure 2.1). In addition, the course name may subconsciously impose the legacy of testing and hinder their knowledge of other assessment instruments which can be used to improve student learning.

Third, the study strengthens the idea that tests can be used for formative purposes, as documented by several studies (Black et al., 2003; Carless, 2011; Lam, 2013; Xiao, 2017). There is a paradox in the Indonesian assessment system: teachers are required to implement the notion of formative assessment while at the same time the government administers high-stakes testing and the assessment systems are filled with summative tests (see Table 5.1). Nevertheless, findings from the study have shown that formative assessment can still be implemented despite the high number of tests administered to students. Teachers made use of the tests to categorise students for further instruction, to identify learning difficulties, and to follow up particular students with remedial teaching and/or remedial tests. Considering the potential formative use of tests, this study has raised important questions about the quality of teacher-made tests in terms of their reliability and validity. In addition, the current data highlights an important issue for curriculum designers, that there should be alignment between the curriculum and compulsory tests.

6.5 Limitations of the study

The major limitation of this study is the small sample of the teachers who participated. Indonesia is a vast area that consists of thousands of islands; it was not possible to investigate
teachers’ formative assessment practice throughout the archipelago. Thus, it is not known whether teachers in other regions in Indonesia understand the concept of formative assessment and are able to manifest their understanding in their teaching instruction.

Furthermore, teachers who participated in this study can be categorised as ‘elite’ teachers: they have the access to workshops and training courses; and are supported by some facilities that they need to accomplish learning goals. It is unfortunate that this study did not include other teachers who have not had the privilege to learn or improve their knowledge, and those who do not have sufficient resources in the schools where they teach.

Another issue that was not addressed in this study was the longitudinal impacts of formative assessment on students’ achievement. This study was limited by the absence of students as a research variable; consequently, it remains unknown in what ways formative assessment improves learning, and how to accelerate students’ achievement through formative assessment. In addition, since this study was limited to teachers’ understanding and practice of formative assessment, it was not possible to identify students’ perceptions towards formative assessment. This issue is crucial because students also contribute to the efficacy of formative assessment practice.

6.6 Recommendations

The findings of the study have shed light on the practice and potential of formative assessment in ELT in the Indonesian context. This study has thrown up many questions in need of further investigation. Some areas that can be further explored are as follows:

1. Since this study mainly concerns teachers, future studies should aim to include students as part of the study. More research is required to determine the efficacy of formative assessment to enhance learning. Thus, further work needs to be done to establish in what way formative assessment enhances learning. A further study
comparing teachers’ and students’ perceptions of effective formative assessment strategies would be very valuable.

2. This study documents testing as an inseparable part of teachers’ formative assessment practice, and teachers made use of summative tests for formative purposes. Future studies should therefore concentrate on the investigation of quality tests made both by teachers and/or by the government in Indonesia. It is important to assess the quality of tests prior to using it as a means to improve learning. Future research may focus on teachers’ ability to write tests, or measure their validity and reliability. More broadly, research is also needed to determine the alignment between curriculum and tests, especially the UN which often directs teaching, and to develop tests which achieve a closer alignment with the curriculum.

3. Given the fact that this study is limited in scope, it is recommended that further experimental investigations are needed to estimate teachers’ conception of assessment. As shown by the findings of this study, teachers’ conception of assessment played a crucial role in the implementation of formative assessment. This information could be used to develop targeted interventions through teachers’ professional development programs.

4. This study found that there is a strong indication that teachers’ formative assessment practice may link to their assessment literacy level. Therefore, it is suggested that the association of these factors is investigated in future studies. This would provide valuable information to determine factors that become barriers in implementing formative assessment.

5. Another possible area of future research would be to investigate how teacher training programs design courses on assessment. Further studies regarding the role of teacher training in preparing teachers for formative assessment practice would be
worthwhile. This may improve future teachers’ formative assessment knowledge and overall practice.

6.7 Reflection

This chapter has completed the overall journey of the research. The summary of the whole thesis has led to the conclusion that it is indeed possible to implement formative assessment in an Indonesian context and to reap its potential benefits. This study also highlights several theoretical and pedagogical implications to realize learning enhancement as a result of formative assessment practice. Further, although this journey has shed light on how formative assessment can be carried out, this may be only a starting point in the quest to unlock the advantages of formative assessment. Further studies, as I have recommended in the previous section, need to be undertaken in order to get a more solid understanding of formative assessment in ELT.

As stated at the beginning of this chapter, this study has shown that teachers in Indonesia have vast potential to implement formative assessment and exploit its potential benefits, especially if teachers are supported with relevant training and if formative assessment barriers can be diminished; hence improving learning outcomes in ELT. As for me, the researcher in this study, this research has strengthened my belief that assessment has a central role to play in leading the course of learning, and formative assessment in ELT is a powerful means of enhancing learning.
References


Ash, D., & Levitt, K. (2003). Working within the Zone of Proximal Development:


https://doi.org/10.1080/10627197.2016.1236677


https://doi.org/10.1080/00405841.2016.1148989


https://doi.org/10.1191/0265532204lt284oa


https://doi.org/10.1002/job.


Alexandria, VA: ASCD.


Yulia, Y. (2014). *An Evaluation of English Language Teaching Programs in Indonesian Junior High Schools in the Yogyakarta Province.* RMIT University.


Appendix 1 Approval from Ethics Committee

2 September 2015

Ms Maya Defiant
Faculty of Arts & Design
University of Canberra
Canberra ACT 2601

Dear Maya,

The Human Research Ethics Committee has considered your application to conduct research with human subjects for the project titled *The Practice of Formative Assessment by ELT Teachers in Secondary High Schools in Indonesia*.

Approval is granted until 30 December 2017.

The following general conditions apply to your approval:

These requirements are determined by University policy and the *National Statement on Ethical Conduct In Human Research* (National Health and Medical Research Council, 2007).

<table>
<thead>
<tr>
<th>Monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>You must, in conjunction with your supervisor, assist the Committee to monitor the conduct of approved research by completing and promptly returning project review forms, which will be sent to you at the end of your project and, in the case of extended research, at least annually during the approval period.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discontinuation of research</th>
</tr>
</thead>
<tbody>
<tr>
<td>You must, in conjunction with your supervisor, inform the Committee, giving reasons, if the research is not conducted or is discontinued before the expected date of completion.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extension of approval</th>
</tr>
</thead>
<tbody>
<tr>
<td>If your project will not be complete by the expiry date stated above, you must apply in writing for extension of approval. Application should be made before current approval expires; should specify a new completion date; should include reasons for your request.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Retention and storage of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>University policy states that all research data must be stored securely on University premises; for a minimum of five years. You must ensure that all records are transferred to the University when the project is complete.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contact details and notification of changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>All email contact should use the UC email address. You should advise the Committee of any change of email address or soon after the approval period including, if appropriate, email address(es).</td>
</tr>
</tbody>
</table>

Yours sincerely

Human Research Ethics Committee

Hendrik Flaegeil
Research Ethics & Compliance Officer
Research Services Office
T (02) 6201 5229 F (02) 6201 5466
E hendrik.flaegeil@canberra.edu.au

www.canberra.edu.au

Postal Address:
University of Canberra ACT 2601 Australia
Location:
University Drive Bruce ACT

Australian Government Higher Education Registered
Provider Number HECP0001970K
Appendix 2 Sample of institution approval

LABSCHOOL

Letter of approval
No. 023/YP-UNJ/SMA/f/VII/2015

Dear Human Research Ethics Committee,

This is to confirm that SMA Labschool Jakarta, Indonesia grant’s approval to Maya Defianty to conduct the data collection for her research project entitled The practice of formative assessment by teachers in secondary high schools in Indonesia. Further, Maya has informed us of the design of the study as well as the targeted population, and we agree to support him in obtaining the data she needs.

We will provide assistance for the success of her research study. Should you need further enquiry, please do not hesitate to contact us at drajatinur@gmail.com

Yours truly,

The principal of SMA Labschool Jakarta
Appendix 3 Information sheet

Participant information form

Project title
The practice of formative assessment by ELT teachers in secondary high schools in Indonesia

Researcher
Maya Defianty
TESOL, Faculty of Arts and Design, University of Canberra ACT 2601
Email:

Supervisor
Adjunct Associate Prof Kate Wilson
TESOL, Faculty of Arts and Design, University of Canberra ACT 2601
Ph:
Email:

Project aim
This study investigates how teachers in secondary high schools in Indonesia practice formative assessment, and what are teachers’ perception towards the formative assessment including challenges that teachers may encounter during the process.

Benefits of the project
The benefits of the projects are:
The study seeks to better understanding of how formative assessment is carried out. Specifically, this study focused on some examples of good practice in formative assessment and challenges encountered by teachers in conducting formative assessment. Hence, this study will make a contribution to teachers’ capacity to implement formative assessment, and the practice of ELT in Indonesia

General outline of the project
This study investigates how secondary teachers conduct formative assessment. The instruments teachers use, how feedback is given to students will be investigated thoroughly. Further, teachers’ perception towards the concept of formative assessment and challenges encountered by teachers in implementing formative assessment will also be explored.

Participant involvement
Teachers who agree to participate in the research will be asked to:

1. Participate in two-three interviews with the researcher at a time that is convenient. Each interview will take about approximately 40 minutes, and be audiotaped with teachers’ permission. In the interview I will ask teachers to describe how they conduct formative assessment in terms of instrument and feedback. I will also ask your perception of the implementation of formative assessment and some challenges that you may encounter during the process.
2. Be observed during your teaching process. The observation will take over four weeks, approximately 16 sessions. Researcher will take the role as participant-observer; it means that I will be unobtrusive during the teaching process, or I will have minimal participation.

3. Provide assessment instruments applied during the process of assessment. Participation in the research is completely voluntary. You are free to choose to participate in the study. You may refuse to participate and withdraw anytime from the study without any loss of benefits and giving any reasons. You may also refuse to answer some or all the questions if you do not feel comfortable with those questions. There is no risk involved in this study except your valuable time. There is no direct benefit to you either. However, the study will have benefits for teacher training in Indonesia and similar contexts and add to current knowledge in EFL teaching and theory.

Confidentiality
The information provided by you will remain confidential. Only my supervisors and I will have access to it. Your name and identity will also not be disclosed at any time. However, the data will be published in the final thesis and possibly in a journal but without giving your name or disclosing your identity.

Anonymity
The research data will be presented anonymously, so that no information in the research report will refer to the individual participant. Your original name will be replaced by a code, and I am the only one who can track and identify your personal information.

Data storage
The information collected will be stored securely on a password-protected computer throughout the project, and then stored at the University of Canberra for a five-year period. The information will then be destroyed according to the university protocol.

Ethics committee and clearance
The project has been approved by the University of Canberra’s Human Research Ethics Committee.

Queries and concerns
Queries or concerns regarding the research can be directed to the researcher or my primary supervisor; our contact details are mentioned at the top of this form. We welcome answering any queries
Appendix 4 Consent form for teacher
Informed consent form for teacher-participant

Project title
The practice of formative assessment by ELT teachers in secondary schools in Indonesia

Consent statement
I have read and understood this consent form. I am aware of any condition that would prevent my participation, and I agree to participate in this project. I understand that I will receive a copy of this form, and I have the opportunity to ask questions about my participation in the research.

Please indicate which parts you agree to by putting a thick (V) in the relevant box.

☐ Participate in three interviews with the researcher
☐ Observed by the researcher
☐ Provide assessment tools to the researcher

Name of participant ______________________

Signature ______________________ Date ____________

A summary of the research report can be forwarded to you when published. If you would like to receive a copy of the report, please include your email address below

Email ______________________
Appendix 5 Consent form for students

Informed consent form for student-participants

Project title
The practice of formative assessment by ELT teachers in secondary schools in Indonesia

Consent statement
I have read and understood this consent form. I am aware of any condition that would prevent my participation, and I agree to participate in this project. I understand that I will receive a copy of this form, and I have the opportunity to ask questions about my participation in the research.

Please indicate which parts you agree to by putting a thick (V) in the relevant box.

☐ My class being observed by the researcher
☐ Provide samples of my tasks and/or tests

Name of participant ______________________

Signature ______________________ Date ____________

A summary of the research report can be forwarded to you when published. If you would like to receive a copy of the report, please include your email address below

Email ______________________
Appendix 6 Interview Protocols

Interview protocol

Project title: The practice of formative assessment by teachers in secondary high schools in Indonesia

A. Introduction

There will be three points in this section: First, introduce myself to the interviewees, and thank them for participating in the interview. Second, explain the purpose of the research study. Third, explain the procedure of the interview. A consent form will be given to each participant prior to the interview.

I am Maya Defianty, a Ph.D research student at the Faculty of Arts and Design. The University of Canberra, Australia. I would like to thank you for your time to participate in this study.

This study aims to identify how teachers conduct formative assessment. I would like also to know your perception of formative assessment. Formative assessment defined as activities that teachers use to collect data of students’ performance and competence, and it involves an on-going process, feedback, and teaching adjustment.

Before we start, let me explain briefly the interview process. The interview will take around 30-40 minutes and will be audio-recorded. I will also write some notes during the interview. All your responses will be kept confidential and any information I include in the research project will not be identify you as the respondent. Please remember that you don’t have to talk about anything you don’t want to and you may end the interview at any time. Are there any questions before we begin?

B. Asking background information

1. Can you please tell me about your educational background?
2. Can you please tell me about your teaching experience?
3. Have you ever joined any teachers’ professional development program throughout your career as a teacher?

C. Asking core questions

1. How do you understand formative assessment?
2. What are the roles of formative assessment in teaching and learning?
3. In your opinion, does formative assessment have distinct characteristics?
4. (Probing question: elements of formative assessment e.g. When do you conduct formative assessment?)
5. What are assessment tools/techniques that you commonly used in formative assessment?
6. Do you have any priority in designing assessment tools? If you do, what are your priorities?
7. Do you give feedback to your students? If you do, how do you give feedback?
8. What are the forms of feedback that you usually give to your students?
9. Why do you give feedback to your students?
10. How do you give feedback?
11. Are there any challenges that you encounter in conducting formative assessment? If you do, what are they? Can you elaborate?

D. Closing
By the end of the interview, I will thank the participant and ask interviewee whether they would like the transcript of the interview in case they would like to amend or elaborate.
Appendix 7 Observation protocols

Observation Protocol
The Practice of Formative Assessment by EFL Teachers in Secondary High Schools in Indonesia

Setting : 
Observer : 
Role of observer : 
Time : 
Length : 

<table>
<thead>
<tr>
<th>Time</th>
<th>Description of Object</th>
<th>Reflective Notes</th>
</tr>
</thead>
</table>

(Adapted from Creswell, 2005, p.213)
# Observation Protocol

**The Practice of Formative Assessment by Teachers in Secondary High Schools in Indonesia**

Observer : 
Date : 
Class :

## Formative assessment in general

<table>
<thead>
<tr>
<th>Question</th>
<th>Description of activities</th>
<th>Reflection</th>
</tr>
</thead>
<tbody>
<tr>
<td>When does teacher conduct assessment?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does teacher establish criteria of quality task?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does teacher adjust instruction after assessment?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Assessment tools used by teacher

<table>
<thead>
<tr>
<th>Question</th>
<th>Description of activities</th>
<th>Reflection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does teacher use assessment tool that elicit evidence of learning?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does teacher apply questioning?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does teacher implement peer assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does teacher implement self-assessment?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>What other assessment tools implemented by teachers?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Feedback given by teacher

<table>
<thead>
<tr>
<th>Question</th>
<th>Description of activities</th>
<th>Reflection</th>
</tr>
</thead>
<tbody>
<tr>
<td>How do teachers scaffold learners?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>What types of feedback given by teachers?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>What are the format of feedback given by teachers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do teachers provide individual feedback?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do teachers use summative feedback formatively?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 8 Sample of assessment criteria applied by one of the participants

ENGLISH WORKSHEET

Write a field trip report.

In your group write field trip reports on a big size of papers (that have been provided by the teacher) by documenting a highly organized account of your experiences during the trip to Banten. You should also use outlines to help you uncover and express the main points of the trip.

Your report will consist of some paragraphs (parts). The first paragraph of a field trip report should include background information related to the trip. The location, date, number of people present, names of guest speakers and other relevant information should also be included in this paragraph.

The second paragraph of the field trip report should expand on the objective of the trip, such as to learn how to produce electricity, explore the nature and learn the ecosystem by exploring the nature, etc. Any anticipated outcomes of the trip should also be mentioned in this paragraph.

The third paragraph of the field trip report should discuss the results of the trip and go into detail about how the results matched or were not in line with the pre-trip expectations.

Good ways for group to conclude field trip reports are to write about how you were effected by the trip and how you would do things differently if you are invited to go on a similar trip in the future. If applicable, the report should include diagrams or charts or photos or other types of visual data that allow others to better understand the scope of the report.

Use your art skill to make an interesting and eye catching of report. The rubric for your report will include the following:

1. Content
2. Generic structure (opening, body and closing paragraph as mentioned above).
3. Grammar
4. Spelling
5. Choice of words (diction)
6. Lay out (including picture, color, etc)

Good luck and have a nice trip!
Appendix 9 Sample of lesson plan

RENCANA PELAKSANAAN PEMBELAJARAN

<table>
<thead>
<tr>
<th>Nama satuan pendidikan</th>
<th>: Nama Lengkap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mata Pelajaran</td>
<td>: Bahasa Inggris</td>
</tr>
<tr>
<td>Kelas/semester</td>
<td>: XII/2</td>
</tr>
<tr>
<td>Materi Pokok</td>
<td>: Newt Item</td>
</tr>
<tr>
<td>Alokasi Waktu</td>
<td>: 2 x 2 JP (dari 6 x 2JP)</td>
</tr>
</tbody>
</table>

A. Kompetensi Inti (KI)

1. Menghayati dan mengamalkan ajaran ajaran yang disumatnya.
2. Menghayati dan mengamalkan perilaku jujur, disiplin, tanggung jawab, peduli (gotong royong, kerjasama, toleran, damai), santun, responsif dan proaktif dan menunjukan sikap sebagai bagian dari suatu atas berbagai permasalahan dalam berinteraksi secara efektif dengan lingkungan sosial dan alam serta dalam menerapkan diri sebagai cermat bangsa dalam pergaulan dunia.
3. Memahami, menerapkan, menganalisis dan mengelola pengetahuan fakta, konseptual, prosedural dan metakognitif berdasarkan rasa ingin tahu serta tentang ilmu pengetahuan, teknologi, seni, budaya, dan humaniora dengan wawasan kemanusiaan, kebangsaan, keagamaan dan keadilan serta menurunkan pengetahuan prosedur pada bidang kajian yang spesifik sesuai dengan bakat dan minatnya untuk memecahkan masalah.
4. Mengolah, menalar, menyajikan, dan mencipta dalam ranah konkret dan ranah abstrak terkait dengan pengembangan dari yang dipelajarnya di sekolah secara mandiri serta bertindak secara efektif dan kreatif, dan mampu menggunakan metoda sesuai keadaan kelimuan.

B. Kompetensi Dasar:

1.1. Menyukai kesempatan dalam mempelajari bahasa Inggris sebagai bahasa pengantar komunikasi internasional yang diwujudkan dalam semangat belajar
2.3. Menunjukkan perilaku tanggung jawab, peduli, kerjasama, dan cinta damai, dalam melaksanakan komunikasi fungsional

3.9 Menganalisis fungsi sosial, struktur teks, dan unsur kebahasaan dari teks News item berbentuk berita sederhana dari koran/radio/TV, sesuai dengan konteks penggunaannya

4.14 Menangkap makna dalam teks berita sederhana dari koran/radio/TV.

C. Indikator Pencapaian Kompetensi

3.9.1 Menyalin dengan tulisan tangan yang rapi beberapa berita sederhana dari koran/radio/TV [news item] dari berbagai sumber, dengan menggunakan ejaan dan tanda baca dengan benar. (pertemuan 1)

3.9.2 Mengidentifikasi fungsi sosial, struktur teks dan unsur kebahasaan (termasuk a.l. gaggasan utama dan informasi rinc) dari teks berita sederhana dari koran/radio/TV [news item] dengan penuh percaya diri bertanggung jawab, dan kerjasama yang baik (pertemuan 1)

3.9.3 Membaca dan mendengarkan teks berita sederhana dari radio [news item] tersebut untuk memahami isi pesannya. (pertemuan 1)

3.9.4 Secara kolaboratif, mencari dan mengumpulkan beberapa berita sederhana dari koran/radio/TV [news item] dari berbagai sumber, termasuk dari internet, film, koran, majalah, buku teks, dsb.

3.9.5 Membaca berbagai berita sederhana dari koran/radio/TV [news item] yang telah terkumpul, secara lebih cermat dengan cara mengidentifikasi dan menyebutkan:
   - fungsi sosial soiops teks
   - jenis atau golongan dari obyek yang dipaparkan
   - nama, bagian-bagian, sifat dan perilaku yang umum ditemukan/ dilihat
   - kosa kata, tata bahasa, ucapan, tekanan kata, ejaan, tanda baca yang digunakan

3.9.6 Membandingkan fungsi sosial, struktur teks, dan unsur kebahasaan dari beberapa teks berita sederhana dari koran [news item] yang telah dikumpulkan dari berbagai sumber tersebut di atas.

3.9.7 Memperoleh balikan (feedback) dari guru dan teman tentang hasil analisis mereka tentang fungsi sosial, struktur teks, dan unsur
kebahasaan yang digunakan dalam teks berita sederhana dari koran (news item) yang dibaca.

4.14.1 Menyampaikan teks berita sederhana yang dibuat sendiri dengan menggunakan program Screenomatic (Screen Casting), dengan cara antara lain
- Membuat draft naskah news item untuk SOM (Screenomatic).
- Mengubah draft menjadi naskah/skrip SOM.
- Melakukan perekaman dan/atau pengambilan gambar untuk memenuhi kebutuhan naskah/skrip SOM.
- Menyiapkan video, power point, atau program tambahan untuk SOM.
- Melaksanakan pembuatan news item dalam bentuk berita TV dengan menggunakan program SOM.

4.14.2 Berupaya membaca secara lancar dengan ucapan, tekanan kata, intonasi yang benar dan menulis dengan ejaan dan tanda baca yang benar, serta tulisan yang jelas dan rapi pada saat membuat naskah dan menyajikannya menjadi berita dengan program SOM.

4.14.3 Membicarakan permasalahan yang dialami dalam memahami teks berita sederhana dari koran/radio/TV (news item) dan menulisnya dalam jurnal belajar sederhana dalam bahasa Inggris.

D. Materi Pembelajaran

<table>
<thead>
<tr>
<th>Peremuan</th>
<th>Materi Pembelajaran Reguler</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A. Let's talk</td>
</tr>
<tr>
<td></td>
<td>B. Active Listening (Newsworthy event from Moscow)</td>
</tr>
<tr>
<td></td>
<td>C. Find out</td>
</tr>
<tr>
<td>2</td>
<td>A. News Item Text</td>
</tr>
<tr>
<td></td>
<td>B. Three Column News Item Text Analysis</td>
</tr>
<tr>
<td>3</td>
<td>A. Passive sentence</td>
</tr>
<tr>
<td></td>
<td>B. Active sentence</td>
</tr>
<tr>
<td>4</td>
<td>Direct and indirect sentence</td>
</tr>
<tr>
<td>5</td>
<td>SOM (Screenomatic) stage 1</td>
</tr>
<tr>
<td>6</td>
<td>SOM (Screenomatic) stage 2</td>
</tr>
</tbody>
</table>
1. Fungsi sosial

Membentuk kegiatan, kejadian, dan peristiwa yang menarik dan layak menjadi berita kepada pembaca, pemirs, pendengar.

2. Struktur teks

a. Menguraikan secara ringkas dan padat tentang kejadian, kejadian, peristiwa yang menarik menjadi berita.

b. Menguraikan secara lebih rinci tentang kejadian/peristiwa/ kegiatan, pihak yang terlibat atau terkait, waktu, tempat, dsb.

c. Menyebutkan sumber berita: komentar, pemyataan, pendapat dari pihak terlibat/terkait, saksi, pihak berwenang, ahli, dsb., tentang yang diberitakan.

3. Unsur kebahasaan

(1) Kata dan ungkapan yang menarik perhatian tentang isi berita pada judul (headline)

(2) Penghilangan kata-kata funsional atau gramatikal (the, a, kata kerja bantu, dll.)

(3) Kalimat langsung dengan kutipan, dan kalimat tidak langsung.

(4) Kalimat pasif

(5) Adverbial dan frasa preposisional tentang waktu, tempat, cara, dll.

(6) Penggunaan nominal singular dan plural secara tepat, dengan atau tanpa a, the, this, those, my, their, dsb secara tepat dalam frasa nominal

(7) Ucapan, tekanan kata, intonasi

(8) Ejaan dan tanda baca

4. Topik

Kejadian, peristiwa, kegiatan yang menarik dan layak menjadi berita

1. Integrasi Muatan Lokal

Membuat berita dengan durasi 2-2.30 menit menggunakan Screenomatic yang memuat kejadian/peristiwa/ kegiatan yang menarik yang terdapat di lingungan sekolah atau lingkungan tempat peserta didik tinggal.

2. Sumber lain dengan sudut pandang yang berbeda.

- Sumber: www.Youtube.com/ABC News

E. Kegiatan Pembelajaran

Pertemuan pertama (2 JP)
<table>
<thead>
<tr>
<th>Kegiatan</th>
<th>Uraian Kegiatan</th>
</tr>
</thead>
</table>
| Pendahuluan (8')              | - Menyiapkan peserta didik untuk mengikuti proses pembelajaran seperti berdak, absensi, menyiapkan buku pelajaran, menyiapkan speaker, memasangkan LCD pada laptop, dan menyiapkan lembaran kerja siswa.  
  - Memotivasi peserta didik secara kontekstual sesuai dengan manfaat pembelajaran teks news item yaitu memberitakan kegiatan, kejadian, dan peristiwa yang menarik dan layak menjadi berta kepada pembaca, peminat, pendengar,  
  - Menjelaskan tujuan pembelajaran atau kompetensi dasar yang akan dicapai dan menyampaikan cakupan materi dan penjelasan uraian kegiatan sesuai silabus.                                                                                                                                 |
| Inti (75')                    | - Peserta didik mengamati layangan LCD dan memusikartikan perhaluan pada liga jenis sumber berta yang disajikan pada power point.  
  - Peserta didik menjelaskan kejadian/pengalaman lucu ketika menonton TV.  
  - Peserta didik menjelaskan fungsi sosial teks new item.  
  - Peserta didik menerima 25 kosa kata yang berkaitan dengan newsworthy event in Moscow.  
  - Peserta didik menemukan arti kata-kata yang disediakan dalam waktu 3 menit secara berpasangan dengan cara menjadokan (mencari dari kamus).  
  - Peserta didik menyimak berita dari radio (recorded) mengenai newsworthy event in Moscow sebanyak 1 kali.  
  - Peserta didik menerima teks rumpang newsworthy event in Moscow dan mengisi bagian rumpang sambil menyimak rekaman berita radio sebanyak 3 kali.  
  - Peserta didik menyimak kembali teks lisan (recorded) sambil mencocokan jawaban yang benar dengan dibimbing guru.  
  - Peserta didik mendapatkan penjelasan singkat mengenai ciri news item dalam ragam tulis.  
  - Peserta didik mengubah teks lisan menjadi teks tulis dengan cara menyalin dengan tulisan tangan yang rapi dengan ejaan dan tanda baca yang benar.  
  - Peserta didik menjawab pertanyaan untuk mengidentifikasi dan menyebutkan: fungsi social teks, gagasan utama dan informasi rinci teks berita.  
  - 5 Soal dibacakan guru dan setiap soal memiliki skor 70 dan
Penutup (7')
- Peserta didik menyampaikan pembelajaran (dibimbing oleh guru)
- Peserta didik mendapatkan umpan balik terhadap proses dan hasil pembelajaran: Thank you very much for your participation. You did a good job today, I'm very happy with your activity in the class. How about you, did you enjoy my class?
- melakukan kegiatan findak lanjut dalam bentuk pemberian tugas individual untuk membaca dan membuat kipling berupa artikel dari berbagai koran (off line atau on line).
- menginformasikan rencana kegiatan pembelajaran untuk pertemuan berikutnya yaitu menganalisis teks news item diambil dari salah satu teks dari kipling.
- Peserta didik Melakukan refleksi terhadap kegiatan yang sudah dilaksanakan dengan cara menulis pada jurnal belajar.

F. Penilaian Pembelajaran

<table>
<thead>
<tr>
<th>Teknik Penilaian</th>
<th>Instrumen Penilaian</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sikap</td>
<td></td>
</tr>
<tr>
<td>a. Jurnal guru</td>
<td>Lembar Jurnal guru:</td>
</tr>
<tr>
<td>Hari/Tanggal</td>
<td>Nama siswa</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>b. Jurnal belajar</td>
</tr>
</tbody>
</table>
|                  | At the end of this session, ask yourself the following questions to know how effective your learning process is:
|                  | 1) What have you learned today? |
|                  | 2) What have you learned from this session? |
|                  | 3) Can you do all the exercises given to you today? Explain and give reasons for that. |
|                  | 4) What is your plan to improve your ability in understanding news item from radio? |
| 3. Pengertian     | Tes isian: What is the communicative purpose of news item? |
|                  | Tes tulis: Match the words in column A with its meaning in column B |
|                  | A           | B    |
Listen to the text and fill the blank.

Good morning listeners. It’s ten o’clock in the morning and you’re listening to Headline News in 007 Archid FM. I’m Mike Tjokro. The first newsworthy event today comes from Moscow. A Russian journalist has uncovered (1) ______ of another Soviet nuclear catastrophe, which killed ten (2) ______ and contaminated an (3) ______ town.

Yelena Vazhavskaya is the first (4) ______ to speak to people who witnessed the (5) ______ of a nuclear submarine at the naval base of Shikotan 22, near Vladivostock.

The (6) ______ which occurred 13 months before the Chernobyl (7) ______, spread radioactive fallout. The base and nearby town, were covered up by officials of the Soviet Union. Residents were told that the explosion was in the reactor of the Victor submarine. It was said to be a “thermal” and not (8) ______ explosion. Those involved in the clean up operation to (9) ______ more than 600 tons of contaminated material were sworn to (10) ______.

A board of investigators were later to (11) ______ it as the worst (12) ______ in the history of the Soviet Navy.

OK listeners, that’s the first news for today.

I am coming back soon with some other important news. So, stay tuned to 007 FM Archid Radio Station.

Find out:
1. Who found evidence of another Soviet nuclear catastrophe?
<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. How many sailors were killed in the accident?</td>
<td>To inform readers, listeners, or viewers about events of the day which are considered newsworthy or important</td>
</tr>
<tr>
<td>3. What contaminated an entire town?</td>
<td>Reactor</td>
</tr>
<tr>
<td>4. Who interviewed the witness of the accident?</td>
<td>Nuklir</td>
</tr>
<tr>
<td>5. Where did the accident happen?</td>
<td>Keelakaan</td>
</tr>
<tr>
<td>6. When the accident did happened, what did the Soviet Union do?</td>
<td>Saksi</td>
</tr>
<tr>
<td>7. Do you know any reason why people involved in the clean-up operation were sworn to secrecy?</td>
<td>Ledakan</td>
</tr>
<tr>
<td>8. How much contaminated material was there in the clean-up operation?</td>
<td>Resmi</td>
</tr>
<tr>
<td>What does ‘catastrophe’ mean?</td>
<td>Menyebar</td>
</tr>
</tbody>
</table>

**Kunci jawaban**

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Jawaban</th>
<th>Skor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reactor</td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Nuclear</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accident</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Witnessed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explosion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Officials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spread</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sailor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Material</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investigators</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catastrophe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fallout</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entire</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covered</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contaminated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newsworthy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secrecy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journalist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Submarine</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. evidence
2. sailors
3. entire
4. journalist
5. explosion
6. accident
7. disaster
8. explosion
9. remove
Pembelajaran Remedial dan Pengayaan
Pembelajaran remedial dilakukan segera setelah kegiatan penilaian

G. Media, Alat, dan Sumber Belajar:
1. Media/alat: Rekaman berita dari radio
   Power Point Presentation
   Speaker active
   Laptop, LCD

2. Sumber Belajar: www.dailyenglish.com,
   http://americanenglish.state.gov/files/ae/resource_files
   Grace and Sudarwati, 2006, Look Ahead for senior High School, Unit 7 Page 173, Erlangga Jakarta

Mengelahui:
Kepala SMAN 2 Cianjur

Cianjur, 22 Januari 2016
Guru Mata Pelajaran

Agam Supriyantia, S.Pd, M.M.Pd
NIP. 196912271994129001

Badriah
NIP. 196804061995122001
Appendix 10 Assessment instruments applied by participants

Participants’ formative assessment practice

<table>
<thead>
<tr>
<th>Nelly</th>
<th>Adam</th>
<th>Wendy</th>
<th>Alfred</th>
<th>Anne</th>
<th>Brenda</th>
<th>Evelyn</th>
<th>Daisy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Questioning Observation Early Diagnostic</td>
<td>Test Questioning Observation Early Diagnostic Assessing students while checking attendance list sampling</td>
<td>Test Questioning Observation Early Diagnostic</td>
<td>Test Questioning Observation Early Diagnostic</td>
<td>Test Questioning Observation Early Diagnostic</td>
<td>Test Questioning Observation Early Diagnostic</td>
<td>Test Questioning Observation Early Diagnostic</td>
<td>Test Questioning Observation Early Diagnostic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 11 Some examples of formative assessment practice applied by only one participant

1. Public speaking
   Basically, this activity assigned individual students to deliver a speech on a topic that they have selected. After the presentation, other students were encouraged to ask questions. During the presentation, the teacher took notes assessing the presenter and the students who asked questions. The teacher gave feedback to the presenter, and students who asked questions, after the presentation; then, he gave suggestions for the next presenter.
   This is an example of formative assessment where the teacher provided immediate formative feedback to improve the performance of the following presenter. Furthermore, the teacher provided an opportunity for other students to actively participate, and he gave feedback to them.

2. Free writing
   In free writing, the teacher gave two key words. He asked students whether they knew the meaning of the words being presented. Afterwards, he defined the words, and gave examples of how to use the words in a sentence. Next, he assigned students to write a short essay that contained the keywords. The students were asked to send their writing via Google Drive so that the teacher could give immediate feedback.
   The teacher was using a formative assessment approach in which learning is a process. He scaffolded students instead of doing direct teaching. He applied questioning to check on students’ previous understanding; then, he asked students to use the words in their essay after he explained the meaning of the words and how they could be employed in a sentence. In addition, students were asked to submit their essay via Google drive so that the teacher could provide immediate feedback.

3. Three-line writing
   This technique is specifically used to improve students’ writing skills. Students were assigned to write essays, but they must leave spaces between sentences. Thus three-line technique applied the following procedures: the first line is the student’s sentence; the second line is the teacher’s feedback; and the last line is the student’s sentence based on feedback from the teacher.
In this example of formative assessment, the students were provided with teacher’s feedback, and the teacher could trace the students’ personal writing development.

4. Video conferencing

Video conference in Wendy’s class was facilitated by Facetofaith website. This website enables students to communicate in English with students from different countries. Prior to the session, a topic was selected, and a moderator to facilitate the session. Besides facilitating, the moderator also gives feedback about the session. Wendy usually recorded the session and asked students to assess their peers during video conference.

This example demonstrates students’ active involvement in assessment. The recorded video conference session enabled students to assess their peer’s performance as well as their own performance. Furthermore, it also enables teachers to return to the collected learning evidence in order to ensure that they have made an accurate assessment. By doing so, the teacher was able to identify students’ learning progress, and adjust her teaching when necessary.

5. Team blogging

In team blogging, students were put in groups; each group was assigned to write an essay about their environment. Topics may vary, but mostly they talked about tourism or local culture. The essay was posted in a blog in which students from other countries also participated. Students were assigned to give comments towards each essay posted by the other groups.

This formative assessment example focuses primarily on improving students’ communicative competence in written form. There are two formative assessment strategies applied in this activity: self-assessment and peer assessment. Students tended to check their work carefully prior to uploading it on the blog, because they realized that it wasn’t just going to be their teacher who read their essay. In addition, students were also given an opportunity to assess their peers’ work.