DEVELOPMENT OF ESL READING COMPREHENSION ASSESSMENT SYSTEM FOR MALAYSIAN YEAR FIVE STUDENTS

CHANG KUAN LIM

UNIVERSITI SAINS MALAYSIA

2018

DEVELOPMENT OF ESL READING COMPREHENSION ASSESSMENT SYSTEM FOR MALAYSIAN YEAR FIVE STUDENTS

by

CHANG KUAN LIM

**Thesis submitted in fulfillment of the requirements
for the degree of
Doctor of Philosophy**

August 2018

ACKNOWLEDGEMENT

I am lucky to have a group of Chang's family members. My thanks from the bottom of my heart to grandmother, Mdm Seow Choon Kee, Father, and Mother who are understanding and keep reminding me to relax myself. Ah Ma, Mum and Dad I made it!

My Godson, Henry Ng Yong Kang, who always gave me the support, lent me his ears whenever I was in the loss of direction. My Godsisters, Tan Geok Cheng, Ooi Seok Ngoh and Loh Ling Ling. Their endless support was important to me. Last but not least, my deepest gratitude to all my 'fans', my students, who have been eager to know the progress.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

PEMBANGUNAN SISTEM PENILAIAN PEMAHAMAN ESL UNTUK

PELAJAR TAHUN LIMA DI MALAYSIA

ABSTRAK

Kajian ini tertumpu pada pembangunan sistem penilaian kebolehfahaman membaca ESL untuk pelajar Tahun Lima di Malaysia. Ujian Pencapaian Sekolah Rendah di Malaysia hanya memberikan gred komposit untuk merumuskan prestasi bahasa Inggeris pelajar. Gred komposit tidak memberikan maklumat lanjut mengenai kebolehfahaman pelajar dalam menjawab soalan-soalan pemahaman. Dalam kajian ini, pembangunan sistem penilaian pemahaman ESL termasuk pembinaan satu ujian pemahaman yang generik dan seragam, pembangunan matriks membaca dan diskriptor kebolehan membaca. Ujian pemahaman yang generik dan seragam tersebut terdiri daripada tiga bahagian dengan soalan pemahaman pada tahap rendah, pertengahan dan lanjutan berdasarkan Barrett's and Bloom's Taxonomy selaras dengan Huraian Sukatan Pelajaran Bahasa Inggeris dan Kurikulum Standard Sekolah Rendah. Setiap bahagian dalam ujian pemahaman terdiri daripada teks linear dan bukan linear dengan soalan-soalan pemahaman literal, pengolahan semula dan inferensi. Kajian rintis telah dijalankan pada ujian pemahaman yang generik and seragam tersebut dan penyelidik meneruskan usaha dalam pembinaan matriks membaca melalui pengenalpastian skor sempadan (cut scores) untuk setiap tahap (band) dan menentukan bilangan tahap (bands). Matriks membaca merupakan sebuah carta di mana guru ESL boleh menyelaraskan skor ujian dengan tahap pendidikan. Pelaksaan ujian pemahaman yang generik dan seragam melibatkan murid Tahun Lima dari sekolah rendah di Larut, Matang dan Selama. Data yang dikumpul dianalisiskan untuk pembangunan diskriptor kebolehan membaca yang

menggambarkan kebolehan penguasaan kemahiran membaca pelajar-pelajar. Dapatan kajian menunjukkan responden Tahun Lima pada Band 1, Band 2, Band 3, Band 4 dan band 5 telah menguasai sub-kemahiran pemahaman literal, pengolahan semula dan inferensi pada tahap yang tertentu. Sistem penilaian kebolehfahaman membaca ESL menyediakan maklumat mengenai kebolehan membaca pelajar di peringkat mikro and makro. Dengan mengetahui maklumat pada peringkat mikro, guru ESL boleh merancang cara pengajaran mereka untuk memenuhi keperluan pelajar mereka dan ibu bapa akan lebih memahami keupayaan membaca pelajar. Kemungkinan besar, maklumat pada peringkat makro boleh membantu pihak Jabatan Pendidikan Negeri ataupun daerah di Malaysia untuk merancang program membaca dalam meningkatkan keupayaan membaca pelajar.

DEVELOPMENT OF ESL READING COMRPEHENSION ASSESSMENT SYSTEM FOR MALAYSIAN YEAR FIVE STUDENTS

ABSTRACT

The present study focused on the development of an ESL reading comprehension assessment system for Malaysian Year Five students. The current Malaysian Primary school achievement test assigns a composite grade to summarise students' performance. The composite grades do not provide any specific information with regards to each student's reading ability. In this study, the ESL reading comprehension assessment system includes the development of standardised generic reading comprehension test, reading matrix and reading performance descriptors. The standardised generic reading comprehension test consists of three sections with reading comprehension questions at elementary, intermediate and advanced levels based on Barrett's and Bloom's Taxonomy in line with the Malaysian English Language Syllabus and Standard Curriculum Document and Assessment. Each section of the test contains linear and non-linear texts with literal, reorganisation and inferential comprehension questions. The standardised generic reading comprehension test was piloted and the researcher proceeded to develop the reading matrix by determining the cut score for each band and the number of bands. A Reading Matrix is a chart that ESL teachers match their students' test scores with educational levels. The administration of the standardised generic reading comprehension test involved Year Five respondents from primary schools located in Larut, Matang and Selama. The data gathered was analysed for the development of reading performance descriptors illustrating the reading sub-skills the students have and have not acquired. The findings revealed that the Year Five

respondents at Band 1, Band 2, Band 3, Band 4 and Band 5 have acquired the literal, reorganisation and inferential reading sub-skills to a certain extent. The ESL reading comprehension assessment system provides information on students' reading abilities at micro and macro levels. By knowing the information at micro level, ESL teachers can plan their teaching instructions to meet their students' needs and parents can better understand students' reading capacity. Perhaps, the information at macro level can assist the District or State Education Departments in Malaysia to plan reading programmes to improve the students' reading ability.

CHAPTER 1

INTRODUCTION

1.1   Introduction

The study intends to develop an ESL (English as a Second Language) reading comprehension assessment system for Malaysian primary schools. It begins with a background of study on the common practice of reading comprehension which allows the researcher to clearly identify the objectives of the study. The research questions that determine readers' performance are formed. The discussion of the research is followed by an explanation of the significance of the study as well as its limitation. By the end of this chapter, terms used in this study are highlighted and operationalised to ensure analytic clarity.

1.2   Background of the Study

Reading is a cognitive process. It occurs when the reader interacts with the text. Reading ability is important to ensure the achievement of educational objectives (Harris and Sipay, 1979). Without a robust ability to read, individuals are at serious disadvantage with respect to educational and vocational opportunities. Individuals that cannot read are unable to comprehend any material for obtaining information. In English language examinations, reading comprehension involves abstracting the main ideas, understanding the sequence of events, guessing the meaning of words and drawing conclusions.

Assessment is an ongoing process and whenever a candidate responds to a question, the teacher makes an assessment of the students' performance (Brown, 2004). Assessment is essential in classroom teaching and learning process as it allows teachers to determine a student's strengths and weaknesses. It is crucial for a teacher to know how a student interprets a reading text so that the student can be assisted through additional learning instrument and approaches if the particular student's difficulties were found (Popham, 1999).

In Malaysia, English is a compulsory subject and is assessed at all levels of public examinations namely: Primary School Achievement Test (also known as Ujian Penilaian Sekolah Rendah), Lower Secondary Assessment (also known as Pentaksiran Tingkatan Tiga), Malaysian Certificate of Education (also known as Sijil Pelajaran Malaysia) and MUET (Malaysian University English Test). Malaysian primary school students are taught reading comprehension and assessed using the Primary School Achievement Test. The students' performance in English language is reported using grades ranging from A to F. Primary school students lack opportunities to apply English in and out of the classroom (Mohd Sofi Ali, 2003) even though students are taught reading comprehension and assessed in the Primary School Achievement Test at the end of Year 6. The effectiveness of English language teaching at primary level is evaluated based on the achievement of the examination results.

The Malaysian Ministry of Education introduced the School Based Assessment (SBA) in 2010. Hwa and Lim (2008) noted that school-based assessment not only helps to boost classroom teaching and learning but also the assessment itself. Students are

assessed and graded in line with the criteria and standards stated in the syllabus. Faizah A Majid (2011) concluded that the new assessment system is a combination of centralised and school-based assessment. The administration of school-based assessment involves all students in a school. The Malaysian Standard Curriculum Document and Assessment (2014) (also known as Dokumen Standard Kurikulum dan Pentaksiran, 2014), in reading section, aims to produce primary school students with the ability to read for information indepedently. Students' performance is categorised into six bands. Teachers are provided with a set of descriptors.

1.3   The Problem

Reading skills are crucial throughout our lifespan but Malaysian pupils prefer to read for their preparation of examination than read for acquiring knowledge and previous research has found out Malaysians students are poor in reading (Inderjit, 2014; Lee, 2014).

The Primary School Achievement Test assesses primary school students' English proficiency. However, according to some primary school teachers and parents who were interviewed during this study, primary school ESL teachers and students are not provided with a standardised instrument for assessment especially in reading comprehension. ESL teachers in primary schools have to adopt English workbooks from private publishers. According to primary school ESL teachers, different schools use different English workbooks from different publishers. This reveals the weakness of the daily formative assessment required by the Malaysian Standard Curriculum Document and Assessment (2014). The worksheets adopted for assessing the students are done

without investigating their validity and reliability. Wiersma and Jurs (1990) expressed that if a test is not valid, the inferences and conclusions made are not convincing. If a test is not reliable, the information about the performance is not consistent (Gay and Airasian, 2000). Swanson and Watson (1989) in Lipson and Wixson (2009) expressed that students who are taking a standardised test should perform the same tasks under uniform directions. Therefore, the question rises about how are teachers able to know exactly the specific reading performance of the students to answer reading comprehension if they do not have a standardised instrument to evaluate the students' reading ability even though they teach with the guidance of the syllabus?

Because of this disparity and for this study, the researcher intended to develop a set of valid and reliable standardised generic reading comprehension test to gauge Year Five students' reading performance. The word 'generic' refers to 'shared by or relating to a whole group of similar things' (Cambridge Dictionary). The standardised generic reading comprehension test developed is applicable to all respondents of upper primary school students. The study does not focus on lower primary students because the teaching of reading skills focuses only on basic literacy with the use of phonics (Ministry of Education, 2014).

Malaysian Primary School Achievement Test is a summative assessment that aims to assess students' ability in reading and writing. Currently, the English language paper consists of two papers that comprise the assessment of vocabulary, language functions, grammar, reading comprehension, sentence construction and note expansion. The English language paper tests different types of skills within a paper. Reading is only

one of the sections in the test paper. However, students' performance is reported by using composite grades that do not further describe students' ESL reading ability accurately. Abdul Rashid Mohamed et al. (2010) remarked the disadvantages of the current assessment system that is, the test scores or grades are the solely information relating to students' reading abilities the ESL teachers possess. However, such results could hardly inform ESL teachers to what extent their students have achieved especially in reading. Consequently, teachers have limited knowledge of students' ESL reading abilities because there are no specific bands and cut scores developed to categorise the students' specific performance in reading.

As stated, grades do not describe how proficient a student is with respect to the material covered (Hammons and Barnsley, 1992). Kubiszyn and Borich (2003) highlighted that parents and students are not provided a detailed description in terms of strength and weakness based on the grades. Therefore, the scores obtained may not be able to serve as a predictor of success in a programme (Abd Samad Arshad et al., 2008). In addition, Sapon-Shevin (1999, 2001, 2003) stated that normal classrooms have always served students who possess different performance or ability. Santhi (2011) agreed that there is mixed ability in each class in which learners have different skills progress at different rates and possess different strengths and weakness in reading. The Malaysian Standard Curriculum Document and Assessment (2014) provides descriptors of performance standards for teachers. The syllabus includes the descriptors of reading but the worksheets of reading comprehension taken from the workbooks of private publishers do not offer detailed descriptors of reading performance. Therefore, ESL

teachers can hardly identify students' strength and weakness in answering reading comprehension questions.

Currently, the Ministry of Education Malaysia (2014) categorises schools using a performance scale from Band 1 to Band 7. Each band is based on a composite score. Benchmarking secondary school students' reading ability was carried out by Abdul Rashid et al. (2010). However, there is still a lack of research on developing a reliable reading assessment system for primary schools in Malaysia. Thus, in this study, not only does the researcher attempts to provide teachers with a detailed description of what the students can and cannot do in reading but also provides a reliable ESL reading comprehension assessment system that deals with ESL teachers' problems. Due to the lack of study on assessing primary school students' reading ability, the researcher intends to conduct this study with several specific objectives.

1.4   Objectives of the Study

The researcher attempts to develop a standardised ESL reading comprehension assessment system for Malaysian primary schools. The objectives and research questions of the study were based on the conceptual framework of this study (Pleaase refer to section 2.12, page 69). The specific objectives were to:

1. compile a standardised generic reading comprehension test by ascertaining its:

   a.   content validity

   b.   construct validity

   c.   reliability

d. discrimination index

e. difficulty index

f. amount of time taken

2. structure a reading matrix by ascertaining its:

 a. the number of bands

 b. the cut-score for each band

3. establish a reading descriptor to diagnose pupils' reading comprehension ability by describing the reading skills which they have acquired

4. establish the reading comprehension assessment system by determining the reliability of the:

 a. generic reading comprehension test

 b. reading matrix

 c. reading descriptors

1.5 Research Questions

The following research questions guide the study:

1. What are the processes involved in developing a standardised generic reading comprehension test?

 a. What could be the best possible combination of reading comprehension questions selected for the standardised generic reading comprehension test?

b. What are the content validity and construct validity of the standardised generic reading comprehension test?

c. What is the reliability of the prototype reading comprehension test?

d. What is the discrimination index of the test?

e. What is the difficulty index of the test?

f. What is the optimum amount of time taken for students to answer the reading comprehension test?

2. What is the appropriate process to develop a reading matrix for the reading comprehension assessment system?

a. What are the most suitable cut scores for each band?

b. What are the appropriate number of bands?

3. What are the reading sub-skills they have acquired?

4. What is the reliability of the reading comprehension assessment system?

a. What is the reliability of the standardised generic reading comprehension test?

b. What is the reliability of the reading matrix?

c. What is the reliability of the reading descriptors?

1.6   Rationale of the Study

In Malaysia, literacy skill is the most fundamental goal of the national system of education (Fong, 2012). Gehrer et al. (2013) pointed out that the National Educational

Psychological Service (NEPS) measures participants' reading progress by collecting the data in standard scores as it is the most statistically valid way to further illustrate the rates of progress made by the participants. Qualitative data is used to indicate that students have developed a more positive attitude toward reading. Hamidah Yamat et al. (2014) noted that the Malaysian English Language syllabus includes all language skills but the 'literacy' aspect is given more focus as these skills are assessed in the national examinations. However, language learning involves more than just decoding printed words (Hamidah Yamat et al., 2014).

Therefore, it is of significant to develop an ESL reading comprehension assessment system for Year Five students to identify their reading ability in terms of what they can and cannot do in answering reading comprehension questions at elementary, intermediate and advanced levels. The results will inform ESL teachers, parents, schools and education departments whether or not to take action or allocate funds to help those students who are not performing well in reading comprehension.

1.7 Significance of the Study

It is essential for ESL teachers to know students' progress in reading. The standardised generic reading comprehension test is able to assess the upper primary students' reading abilities. This research attempts to provide primary school ESL teachers a set of detailed reading descriptors of students' ESL reading ability. McMahon (2003) expressed that the classroom-based assessment process encourages teachers to examine students' specific needs, thus, design teaching instructions accordingly. This study will benefit the students because they will know their reading ability specifically.

9

For ESL teachers, by knowing the students' strength and weakness in reading, they can teach one level above students' reading proficiency. As claimed by Guth and Pettengill (2005), reading assessment allows teachers to examine and discuss the reading assessments used in the schools. They provide specific information obtained from each test and clarify commonly used terms such as decoding skills and instructional reading level.

As reported by Khoo (2014), the implementation of school based assessment has made teachers frustrated because it requires teachers to key-in data on each student daily. In terms of practicality, the standardised generic reading comprehension test just needs to be carried out thrice every year. Teachers only have to conduct standard tests at the beginning, middle and the end of the school terms so that they can identify the weakness of the students in reading. By identifying the students' weakness, the teachers can prepare their instructional materials to teach the students so that the students can answer the reading comprehension questions effectively. Teachers do not need to spend too much time with data entry every day. Instead, the time commonly used for keying-in data can be used to design or develop practical teaching materials for their students.

Moreover, this research also intends to provide a set of reading performance descriptors for each band (Band 1, Band 2, Band 3, Band 4 and Band 5). Each band describes in detail what reading skills the students have and have not acquired.

With the information at macro level, the education department can allocate funds to organise workshops for the students who are not performing well in reading

comprehension. For students who are performing well, the education department may provide enrichment reading programmes to strengthen their reading skills.

1.8   Limitations of the Study

Although this research is carefully prepared, there were unavoidable limitations. First of all, this study was only conducted in primary schools in Larut, Matang and Selama in Perak. Therefore, to generalise the results for larger groups, the study should have involved students from different states in Malaysia. Secondly, the study does not include the technique of teaching reading comprehension and prepare teaching materials to suit the needs of the students because these issues would entail another research project. Thirdly, the results obtained through the standardised generic reading comprehension test cannot be generalised to the Year 5 students in other states of Malaysia.

1.9   Operational Definition of Terms

Defining what is analytically meant by a term, or operationalising terms, to avoid confusion multiple interpretations of a term might have is a necessary part of defining the scope of this research. The terms in need of operationalising that are used in this study include:

Assessment for learning:

Assessment for learning is not a different form or class of assessment. It utilises assessment information to guide decision making to improve learning outcomes (Masters, 2014). Marking is not used to make comparative judgement but is done to

highlight each student's strengths and weaknesses, thus provide feedback to the students. This study attempts to design a standardised generic reading comprehension test system to identify what the students can and cannot do.

Barrett's Taxonomy of Reading Comprehension:

Airasian (2001) defined taxonomy as a system of classification. Barrett's Taxonomy of Reading Comprehension is organised into five levels namely, literal, reorganisation, inferential, evaluation and appreciation. In this study, the sub-skills involved in the development of reading comprehension questions are literal, reorganisation and inferential comprehensions.

Cut Score:

Cut scores are the selected points on the score scale of a test (Salvia et al., 2010; Zieky and Perie, 2006). Zieky and Perie (2006) further explained that the points are used to determine whether a particular test score is sufficient for some educational purposes. For example, a group of students' performance on a test may be classified into 'basic', 'proficient', or 'advanced' on the basis of cut scores. In this study, cut scores were used to categorise the students into five bands (Band 1 to 5). The researcher used the z-scores to develop the cut scores.

Reading Comprehension:

Bormuth (1969) cited in Harris and Sipay (1980), defined comprehension ability as the generalised knowledge-acquisition skills that permit a person to acquire and exhibit information. The core reason for reading is to figure out the information in the text

12

(Tannenbaum et al., 2006). Dechant (1970) stated that readers use contextual setting in interpreting words. In this study, reading comprehension involves the three major skills categorised in Barrett's Taxonomy of Reading Comprehension (literal, reorganisation, inferential) to answer the standardised generic reading comprehension questions.

Reading Descriptors:

Reading descriptors refer to the descriptive schemes that describe the learners' reading progress (Cambridge University Press, 2001). In this study, a set of reading performance descriptors will be developed. They describe in detail a learner's reading ability at each reading performance band (Band 1 to Band 5). Each reading performance band explains the learners' strengths and weaknesses in answering reading comprehension questions at elementary, intermediate and advanced levels. In this study, the reading performance descriptors are developed based on the test scores obtained from the standardised written reading comprehension test and structured interview.

Reading Matrix:

A reading matrix refers to a chart that acts as a reading indicator. It indicates a learner's reading ability at a particular educational level (Abdul Rashid, et al.2010). In this study, the reading matrix is developed to identify whether the respondents are Band 1, Band 2, Band 3, Band 4 or Band 5.

Revised Bloom's Taxonomy of Cognitive Domain:

According to Kubiszyn and Borich (1996), Bloom et al. devised a method of categorising objectives according to cognitive complexity. The taxonomy explains six levels of cognitive complexity ranging from 'remember', 'understand', 'apply',

'analyse', 'evaluate' and 'create'. In this study, the sub-skills involved in the development of reading comprehension questions are 'remember' and 'understand'.

1.10 Conclusion

To conclude, there are various types of reading assessment conducted in overseas countries and by local researchers in Malaysia. However, research on developing an ESL reading comprehension assessment system for primary schools is still lacking. Therefore, the study develops a set of standardised generic reading comprehension tests to assess the Year Five students' reading ability. With the help of the test scores, the researcher able to determine the Year Five students' reading ability and the reliability of the standardised generic reading comprehension test.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter provides an overview of related studies found in the literature and a detailed elaboration of the definition of reading, reading comprehension, various kinds of assessments, and validation of an assessment. A discussion then follows regarding the concept of reading matrix and reading performance descriptors.

2.2 Reading

Reading refers to the overlapping steps and techniques that readers use to help understand the printed page. Basically, three major processes are involved in reading that help readers to read effectively namely: before reading, while reading and after reading (Wiener and Bazerman, 2006). Cunningham and Stanovich (2001) stated that cognitive process and task of lifting meaning from a passage takes place in reading.

Dechant (1982) defined reading as the making sense of experience as well as graphic symbols. Similarly, Basaran (2013) defined reading as a process that primarily contains cognitive aspects that include the perception of written symbols, to know letter voice, the comprehension of information, and relating relating this information both with interlocutors and other prior knowledge.

Alderson (1984) (cited in Madhumathi and Arijit Ghosh, 2012) expressed that reading is vital because it ensures success in academic learning as the ability to read will help students to excel academically (Carrell, 1991; Clarke, 1979; Cziko, 1978). According to second language reading research, reading is an interactive and meaning building process in which readers apply their strategies to understand information from available resources. The reading strategies involve skimming, scanning and inferring.

Shazila Adbullah et al. (2012) noted that reading is the most integral part in language learning because it enables students to open a window to the outside world as readers with strengthened reading skills will be able to progress and attain greater development in all academic areas. However, Shazila Abdullah et al. (2012) also pointed out that poor readers are not only reluctant to read but also they tend to perform poorly in reading tests. Generally, readers who failed in answering reading comprehension questions correctly means that they failed to comprehend the given texts.

By and large, there are many definitions of reading expressed by researchers and it is evident that reading is an important skill that all learners must possess to comprehend a text.

2.3 Reading Comprehension

Wallace (1992) in Morales (2010) expressed that reading is a medium for social interaction and a means to access general knowledge of the world besides being a tool for survival. Therefore, reading and understaning a written text is one of the most important skills; this is referred to as reading comprehension (Tannenbaum et al., 2006).

A learner must first be able to make sense of the smaller word units before being able to comprehend paragraphs or stories. Burt et al. (2005) in August (2011) expressed reading comprehension means the ability to make meaning from a written text.

However, Cain and Oakhill (2006) in Bellinger and Diperna (2011) defined reading comprehension as a complex cognitive process which involves the integration of information, making inferences and constructing meaning from the texts. Children will face educational obstacles if they are poor at comprehension skills. Undoubtedly, reading comprehension is an essential element of gaining knowledge, improving one's learning, and communicating information when one reads (Bellinger and Diperna, 2011). Similarly, Cain and Oakhill (2006) found that children who were poor at reading comprehension made fewer academic gains than those who have good comprehension skills especially children aged 7 to 8 years old.

According to Nader Assadi Aidinlou and Ambigapathy (2011), reading comprehension involves lower- and higher-level processing skills that are coordinated in a very efficient combination. However, many ESL learners never practise reading comprehension as an active process (Shazila Abdullah et al., 2012). Most of the syllabi for ESL reading outline the sub-skills of reading comprehension such as making inferences, predicting and making conclusions. Learners should integrate the sub-skills in order to comprehend a reading text. To improve learners' reading comprehension, ESL learners need sufficient practice and participate in the reading process actively. Pressley (2000) in Shazila et al. (2012) expressed that reading comprehension involves decoding the text into words. It also involves readers' behaviours but it is a multi-

dimensional process which includes the reader, the text and factors associated with the activity of reading. Reading Study Group (RAND) (2002) (cited in Lipka and Siegel, 2012) defined reading comprehension as a process of extracting and constructing meaning simultaneously through interaction with written texts.

In short, reading comprehension involves a complex set of skills (Andreassen and Braten, 2010; Sweet and Snow, 2003). Bormuth (1969) in Harris and Sipay (1980), stated that the ability to comprehend allows learners to acquire and exhibit information gained.

The next section will discuss the definition of assessment and what kind of assessment is suitable for the purpose of this study.

2.4   Assessment

Salvia et al. (2010) defined assessment as a process of data collection done for making decisions about what students have learned and, what and where they should be taught. Students' competence is measured during the assessment. Specifically, teachers measure the students' progress toward their attainment in schools. The assessment information allows parents and the educators to know the extent to which students gain benefits from their school experiences. The federal education policy contains specific expectations for states to develop high educational standards and utilise tests to measure the extent to which students meet the standards.

Similarly, Stiggins (2008) defined assessment as the procedure of assembling evidence of learners' learning progress to inform instructional decisions. The test can function effectively when teachers gather accurate information about the achievement of the students. Teachers not only grade the students but also help to enhance both the students' passion to learn and their accomplishment.

Generally, in education, assessment is conducted to improve student learning, monitoring their progress and certificating their level of performance. The following sub-sections will discuss various types of assessments: assessment of learning, assessment for learning, summative assessment, formative assessment, norm-referenced assessment and criterion-referenced assessments.

2.4.1   Assessment of Learning

According to Earl (2003), the main kind of assessment happening in schools is assessment of learning. It is summative, intends to certify learning and report to parents and students about progress in school. Assessment of learning is typically administered at the end of a unit, a course or a programme. Earl (2003) stated that the examination questions are developed from the materials adopted in classroom teaching. In assessment of learning, results are reported symbolically with marks or letter grades assigned to summarise student performance. However, grades provide little direction or advice for improvement. Typically, the test content itself has the limitation and the test scores are too simplistic to represent the wide range of skills and knowledge (Earl, 2003).

Stiggins et al. (2007) defined assessment of learning as assessments are carried out after learning has occurred. The assessment aims to make statements of students' learning progress at that point in time. Assessment of learning done within the classroom allows the teacher to gather evidence to determine a student's report card grade.

Assessment of learning does not imply a different class of assessments. It is simply the use of assessment information to draw conclusions about progress. The progress can be at the level of groups or individual growth (Masters, 2014).

2.4.2   Assessment for Learning

In conducting assessment for learning, teachers must collect a wide range of data that will then allow them to modify the learning activities for their students. To conduct the assessment, teachers observe, use worksheets and question in the class. Marking is not used to make comparative judgment. Instead, it intends to highlight each student's strengths and weaknesses thus, provide feedback to the students. In assessment for learning, teachers are central characters but they play different roles. They use their personal knowledge of their students and their understanding of the context in the assessment. Assessment for learning is interactive and helps to provide teachers to scalffold the process of learning (Earl, 2003).

Stiggins et al. (2007) stated that assessments for learning takes place when learning is still underway. They are being conducted by teachers throughout teaching and learning to diagnose students' needs. Teachers can plan their teaching instructions

and provide students with feedback to improve the quality of their work (Stiggins et al., 2007; Stiggins, 2008).

Stiggins (2008) believed that teachers, with the help of evidence of students' achievement, can develop an assessment map that parallels a continuous-progress curriculum map. By doing so, students can anticipate when every assessment will take place. This will help them to learn better next time. With the descriptive details in assessment for learning, teachers can focus on critical improvements in students' achievement.

Assessment for learning is not a different form or class of assessments. The information obtained is used to help teachers in decision making on how to improve learning outcomes (Masters, 2014).

2.4.3   Summative Assessment

Summative assessment provides information that can help to appraise the work of teachers and systems of education. It is a kind of high-stake assessment (Knight, 2001).

Summative assessment intends to record the overall achievement of a student in a systematic way (Horton, 1990). Banks (2005) noted that summative assessment is a type of formal assessment assessing the outcome of learning after the instructional programme. Summative assessment aims to determine if the student achieved the

objective of an academic programme. Usually, school final examinations and competency tests are examples of summative assessment.

According to Masters (2014), summative assessments are formal and externally developed tests and examinations. They are used by teachers to identify a learner's learning progress and thus, make a report on learning achievement.

(a)   The Benefits of Summative Assessment

Standardised tests generate a performance ethos in the classroom. They become the rationale for teachers to make all classroom decisions besides shaping students with strong extrinsic orientations toward performance. Students obtain the benefits as the descriptions received help them to better understand the criteria of assessment as well as what is expected of them. Summative assessments are used for internal purposes and the non-judgmental feedback from tests motivates students to put further effort into their work. The dynamic classroom assessment environment concerns what is valued. This helps to establish a learning centric culture in the classroom. It is also an influence on students' learning enthusiasm and achievement goals (Ames, 1992; Brookhart, 1997; Harlen & Crick, 2003).

Teachers benefit from being exposed to assessment strategies that require students to think more deeply (Black et al., 2010). With summative assessment, teachers made significant changes to their assessment practices (McMillan, 2008).

(b)   The Disadvantages of Summative Assessment

In summative assessment, grades are calculated without defining the criteria of success for different levels of performance (Griswold, 1993; Hills, 1991; Stiggins et al., 1989). According to Goldberg and Roswell (2000), a student's performance occurs consistently but teachers tend to record their judgments only after a learner's performance. Thus, the accuracy of the students' performance is weakened. In reality, teachers stay focussed on test items, provide exam tips and even extend the time frames (Hall & Kleine, 1992; Nolen, Haladyna, & Haas, 1992).

Assigning grades as rewards and punishments will both decrease students' motivation to learn and harms the students' passion in learning. It could be hinderance for teachers to obtain specific information when the assessment is subjected to close external control, Black et al., (2010) revealed that assessment purposes could rarely match the assessment goals due to the inconsistency of teachers' assessment validity;. McMillan (2008) concluded that higher ability students are motivated when the atmosphere is engaging. However, low achieving students experience rote learning.

2.4.4   Formative Assessment

Assessments are useful in identifying what learners need to do to improve their work. Formative assessment intends to inform or provide students some suggestions about how to do better besides offering feedback to students about their achievements. It has been perceived that formative assessment stresses on providing useful feedback when learners do not attempt to conceal their mistakes. It is not a high-stake assessment

and learners cannot fully rely on the advice about continuing to improve the particular work (Knight, 2001).

The purpose of formative assessment is to identify the positive achievements of a student so that the appropriate steps may be planned next (Horton, 1990). Banks (2005) defined formative assessment as a planned assessment that offers a guide for both teachers and students. It can also serve as a self-assessment and might not be used to determine grades.

According to Caldwell (2014), formative assessment aims to identify the learning needs of students so that adjustment can be done on teaching instructions. Therefore, formative assessment involves frequent checks of students' understanding and skills.

Masters (2014) noted that formative assessments are based on classroom observations done by teachers in detail every day. They allow teachers to have ideas on starting lessons in the classroom.

(a)   The Benefits of Formative Assessment

Chappuis and Chappuis (2008) highlighted that formative assessment made the style of teaching becomes lively and interactive instead of demonstrating to the students the ways to look for solutions. In formative assessment, usually, a question is asked and pupils are given time to look for answers with their classmates. Students are expected to be able to answer at any time. Pupils are comfortable even if they give a wrong answer. Thus, pupils become active participants in the classroom.

Educators found that pupils are productive as they engage themselves in improving their work. Implementation of such reforms can change the attitudes of teachers and pupils when assessment is less competitive; they take the summative judgement as a process of learning.

Pupils can only achieve a learning goal if they understand what they need to accomplish. Therefore, self-assessment is important in learning (Sadler, 1989). Peer-assessment is extraordinarily valuable because pupils may receive commnets of their work from teachers. Teachers are free to observe, reflect on what is happening and then frame helpful interventions.

(b)   The Disadvantages of Formative Assessment

In formative assessment, more effort must be spent in framing questions that are worth asking and follow-up activities should be adequate to ensure learners' comprehension.

The research by Butler (1988) established that giving feedback through comments could help pupils' learning; however, marks or grades could have negative effect on the pupils especially those who ignore comments. Improvement on comments requires more work because teachers are responsible for the quality of the comments that they wrote on pupils' work. This is because a teacher's judgment can directly influence a student in terms of his or her achievement, study patterns, perceptions, attitudes, effort, and motivation to learn (Black & Wiliam, 1998; Brookhart, 1997; Rodriguez, 2004).