DEVELOPMENT AND VALIDATION OF COGNITIVE DIAGNOSTIC ASSESSMENT (CDA) FOR PRIMARY MATHEMATICS LEARNING OF TIME

CAROLYN SIA JIA LING

UNIVERSITI SAINS MALAYSIA

2017

DEVELOPMENT AND VALIDATION OF COGNITIVE DIAGNOSTIC ASSESSMENT (CDA) FOR PRIMARY MATHEMATICS LEARNING OF TIME

by

CAROLYN SIA JIA LING

Thesis submitted in fulfillment of requirements for degree of

Master of Arts (Education)

February 2017

ACKNOWLEDGEMENT

First of all, I would like to express my sincere gratefulness to my supervisor, Prof. Dr. Lim Chap Sam. She has guided and supported me endlessly not only throughout the preparation of this thesis but also my two years of postgraduate life. I would also like to thank Assoc. Prof. Dr. Chew Cheng Meng and Dr. Wun Thiam Yew for their time and advices in completing this thesis.

I wish to thank all the teachers and schools who participated in this study. This study could not be completed without the commitment and cooperation from them.

Besides, I would like to express my gratitude to my dearest family, especially my late father. He had read and given me advices on my writing in every stage of my postgraduate life. Even though he could not make it for this very last stage of my study, I really appreciate him for supporting me all the while. Of course, my gratitude also goes to my mother and my siblings for their love and support.

Last but not least, I would like to thank all of my friends who gave me encouragements, spiritually supports and helps. It is so good to have all of you with me for these two years.

What shall I render to the Lord for all His benefits toward me? (Psalm116)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PEMBENTUKAN DAN PENGESAHAN PENILAIAN DIAGNOSTIK KOGNITIF DALAM PEMBELAJARAN TOPIK MASA UNTUK MATEMATIK SEKOLAH RENDAH

ABSTRAK

Kajian ini bertujuan untuk membentuk dan mengesahkan satu Penilaian Diagnostik Kognitif (PDK) dalam topik Masa untuk matematik sekolah rendah. PDK merupakan sejenis penilaian yang boleh menghasilkan maklum balas yang berguna mengenai kekuatan dan kelemahan kognitif pelajar. Dalam kajian ini, PDK dibentuk berdasarkan *Assessment Triangle* dan *Attribute Hierarchy Method (AHM)*. Kajian ini melibatkan tujuh orang guru matematik sekolah rendah yang berpengalaman dan 1123 orang murid sekolah rendah. Data kuantitatif telah dikutip dengan menggunakan ujian kertas dan pensel manakala data kualitatif telah dikutip melalui sesi temu bual dengan murid. Pembentukan penilaian ini melibatkan dua peringkat iaitu pembentukan model kognitif dari segi pakar dan penghasilan modal kognitif dari segi pelajar. Terdapat tiga langkah dalam peringkat pertama, iaitu identifikasi atribut, pembinaan hierarki atribut dan pembentukan soalan. Di samping itu, peringkat kedua melibatkan dua langkah, iaitu pelaksanaan penilaian dan analisis data mengikut kaedah statistik. Pengesahan instrumen telah dilaksanakan melalui dua cara. Cara pertama ialah proses pembentukan penilaian tersebut telah diulangi sebanyak empat kitaran untuk memastikan kesahihan penilaian yang dibentuk. Seterusnya, *Hierarchical Consistency Index (HCI)*, kebarangkalian atribut dan korelasi antara kebarangkalian atribut juga digunakan menyokong struktur hierarki model kognitif. Hasil kajian membuktikan kesahihan dan kebolehpercayaan penilaian yang telah dibentuk. Guru boleh mengenal pasti kesukaran dan salah faham konsep murid dalam pembelajaran Masa memandangkan atribut dalam PDK ini

disusun mengikut tahap kesusahannya. Selanjutnya, guru boleh merancang pengajarannya dan membuat keputusan mengenai pengajaran yang bersesuaian dengan tahap murid. Kesimpulannya, memang terbukti bahawa CDA merupakan kaedah penilaian yang efektif untuk membantu guru menilai muridnya supaya meningkatkan kualiti pengajaran dan pembelajaran di dalam kelas.

DEVELOPMENT AND VALIDATION OF COGNITIVE DIAGNOSTIC ASSESSMENT (CDA) FOR PRIMARY MATHEMATICS LEARNING OF TIME

ABSTRACT

This study aims to develop and validate the cognitive diagnostic assessment (CDA) for primary mathematics learning in the topic of "Time". The CDA is an assessment which can provide meaningful feedback on learner's cognitive strengths and weaknesses during teaching and learning. In this study, the CDA was developed based on the Assessment Triangle and Attribute Hierarchy Method (AHM). This study involved seven experienced primary mathematics teachers and 1123 primary school pupils. Quantitative data was collected through paper and pencil test while qualitative data was collected via think-aloud interview session. There were two phases included in the development of CDA: a) the development of the expert-based cognitive model; and b) generation of the student-based cognitive model. Generally, three stages were involved in Phase I: identification of attributes; development of attribute hierarchy; and item design. Meanwhile, Phase II involved two stages, namely administration of the assessment and statistical analysis of the data collected. Validation of the instrument was carried out in two different ways: a) the usage of development process involving four cycles in order to validate the instrument, and b) the application of Hierarchical Consistency Index (HCI), attribute probabilities and correlations among attribute probabilities to support the hierarchical structure of the cognitive model. The results showed that the developed CDA was valid and highly reliable (Cronbach's Alpha > 0.8). Since the attributes in the CDA were arranged hierarchically from the most basic to the most complex level, teachers were able to identify pupils' learning difficulties and misconceptions about "Time". Based on

these inferences, teachers could plan their lessons and make instructional decisions appropriately. Hence, the CDA could be an effective assessment technique for teachers to assess their pupils in order to improve the quality of teaching and learning in the classroom.

CHAPTER 1

INTRODUCTION

1.0 Introduction

As a matter of fact, assessment plays an important role in teaching and learning. As defined by White (2007), "Assessment is seen as a process of gathering evidence and making judgments about students' needs, strengths, abilities and achievements" (p.46). However, the overemphasis on centralized national examinations or high-stakes assessments in the Malaysian previous assessment system was criticised by some of the experts. These assessments are criticised as they are (a) unable to provide a "realistic estimate of overall achievement" (Ong, 2010, p.99); (b) unable to provide an accurate result of a particular individual's real competencies (Che Noraini, Adlina & Nurhidayah, 2013; Ong, 2010); and (c) they might put pressure on teachers to produce high score performers in the assessments (Ong 2010).

Hence, in 2011, the Malaysian Ministry of Education (MOE) came out with the School Based Assessment System (SBA, or better known as *PBS* in the Malay language) which proposes an assessment system to assess every student fairly and equally, based on what he/she has acquired in the end of the lesson. This educational reform is a shift from the major focus on central examination to a more holistic educational assessment system which evaluates students according to their cognitive, affective and psychometric aspects as prescribed in the National Education Philosophy (MOE, 2011). In the SBA system, assessments are administered continuously throughout all grades. This is in view of the previous assessment system failing to provide meaningful diagnostic feedback on pupil's level of understanding and their learning progress. Hence, the change is necessary in order to evaluate the learners more effectively and ensure that the assessment system is educationally sufficient in this globalized era (Mansor, Ong, Rasul, Raof & Yusoff,

2013).

In the SBA system, teachers are expected to evaluate their students by using various methods of assessment (Ong, 2011). However, most teachers are still unfamiliar with their roles in the latest assessment system, not to mention the role of item analyses, the most basic type of diagnostic test. Hence, there is a need to develop a better assessment framework that will help teachers to design more effective and efficient assessment items.

1.1 Background of the study

School Based Assessment [SBA]

The Malaysian primary school system consists of six years of formal education. In 2011, the existing curriculum (Integrated Curriculum for Primary Schools, or better known as *KBSR* in Malay language) was replaced by the Primary School Standard Curriculum (PSSC, or better known as *KSSR* in the Malay language) for Year 1 pupils. In both syllabi, all primary pupils begin learning about “Time” since Year 1. The curriculum of "Time" is designed according to the manner of increasing complexity, for example, pupils will learn the name of days in a week in the beginning, followed by the conversional relationship between minutes and hours, and then, that of years and decades. In short, throughout the six years of primary education, pupils are expected to be familiar in (a) reading the clock faces and calendars, and applying the knowledge gained in daily life; (b) knowing the relationship between two different time units, such as hour, minute, second, day and month, and also (c) performing all four basic mathematical operations (addition, subtraction, division, and multiplication) involving "Time" confidently. These curriculum specifications of the topic of “Time” in the primary school Mathematics

syllabi are shown in Appendix A.

As the consequence of the implementation of KSSR in the Malaysian primary education system, students are now assessed continuously from Year 1 to Year 6 through School-Based Assessment (SBA). The rationales of the implementation of SBA are that it aims to (a) achieve the aspiration of National Philosophy of Education towards learners' physical, emotional, spiritual and intellectual abilities; (b) reduce the elements of exam-oriented education in the national education system; (c) assess learners' progress of learning (with more meaningful assessments); and (d) improve teachers' integrity in assessing, recording and reporting learners' learning (MOE, 2011).

The implementation of the SBA aims to produce excellent human resources, which is parallel with the aspirations and objectives of the National Education Philosophy by maximizing an individual's potential (Jaba, Hamzah, Bakar & Rasid, 2013). Furthermore, it also intends to evaluate pupils conforming to the three domains of learning development: cognitive, affective and psychomotor. The SBA system consists of four different assessments, namely school assessment, physical, sports and co-curricular, psychometric/psychological assessment, and centralized assessment. The structure of the SBA system is summarized in Figure 1.1.

In the SBA system, teachers will have to shoulder full responsibility, from the selection and design of assessment tasks, followed by the administration of assessment tasks during teaching and learning process, and then, the evaluation of students' learning progress based on their responses, and lastly, the conclusive report about their academic progress. In SBA, pupils' performances will be categorized into bands accordingly, with Band 1 indicating the lowest band, and Band 6 as the

highest band. The purpose of implementing school assessment is to collect information about students' learning.

Another essential part of the assessment is the centralized assessment. The centralized assessment is prepared by the Malaysian Examination Syndicate (*Lembaga Peperiksaan)*, while the administration will need to be carried out by the teachers. In Physical Activities, Sports and Co-curricular (PASK) assessment, students' physical endurance and body mass index will be measured. Students' ~~i~~nvolvement in physical activities, sports and co-curricular activities will also be a part of the PASK assessment as well (MOE Circular, 2014).

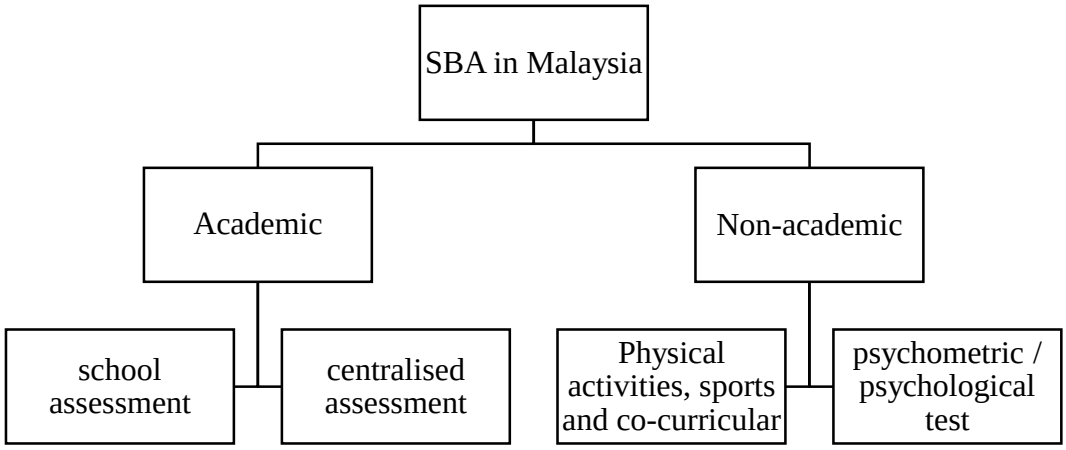

Figure 1.1: The structure of SBA in Malaysia (adapted from MOE circular, 2014)

Cognitive Diagnostic Assessment [CDA]

Cognitive Diagnostic Assessment [CDA] is an assessment that provides fine-grained analysis of students' skill mastery profiles and their cognitive knowledge state (Jang, 2009). According to Alves (2012), CDA is a combination of theories of learning and psychometric model, thus, students' mastery levels on a defined set of domain could be reported instead of only the numerical score. Understanding on students' learning

progress helps educational stakeholders to make effective instructional decision (Ketterlin-Geller & Yovanoff, 2009). Besides, students' cognitive processes and skills involved in solving the tasks could also be investigated by using CDA. Therefore, CDA could be a helpful tool for teachers in fulfilling their role in SBA system.

1.2 Problem Statement

Teachers play a very significant role in the effective implementation of the SBA system (Hamzah & Sinnasamy, 2009; Jaba et al., 2013; Majid, 2011). In fact, in the SBA system, teachers are expected to assess pupils on all three domains: cognitive, affective and psychomotor. They have to choose and design tasks which are able to provide diagnostic information about pupils' learning progress in the end of the analyses. Besides, pupils are expected to acquire the skill of completing the self-assessment themselves for they have to be responsible for their own learning after all. SBA emphasizes both formative and summative assessments. The formative part is carried out alongside with the teaching and learning process to monitor pupils' learning progress and gather feedback, whereas the summative portion is carried out at the end of a lesson unit, several units in a specific theme or a particular semester to obtain an overall picture of the pupils' learning performance. The latter certainly requires teachers to master various assessment methods in order to gather sufficient evidence on pupil's learning based on different aspects as aforementioned.

Several studies have been carried out to investigate teachers' acceptance and perception towards SBA (e.g. Che Norliza et al., 2013; Hamzah & Sinnasamy, 2009; Jaba et al., 2013), and concerns about the implementation of SBA (e.g. Majid, 2011) in national schools. Majid (2011) conducted a study on 40 English teachers to find

out their levels of concern about SBA. His finding showed that the teachers involved in the study mentioned that they were not able to understand about the demands of SBA entirely. Teachers were worried that they might not be able to fulfil the demands as well as execute their roles successfully in SBA. This finding concurred with that of Che Norliza et.al. (2013) which stated that in the initial stage, teachers perceived positively towards SBA as the result of SBA assessment was more meaningful for it assessed various learning domains of students. However, after two years of SBA implementation, teachers found out that they faced several challenges in carrying out SBA on the learners. One of the challenges was that they were not sure if SBA could really reflect pupils' learning competencies in the classroom. There was also an issue of insufficient knowledge in the management and usage of the SBA. All five teachers involved in this study mentioned that they were not familiar with the operation of SBA system as there were only limited numbers of formal training courses about SBA available around them since it was introduced. Therefore, these challenges must be solved to ensure effective and efficient implementation of the SBA system. As stated by Stern (1983) in his study, teachers' understanding on the demands of new curriculum (assessment inclusive) is very important in order to ensure full acceptance towards the newest curriculum.

Similar concerns were brought up by two international teams, namely the Centre of Public Policy Studies and Asian Strategy and Leadership Institute (CPPS-ASLI, 2012), and the Centre of British Teacher (CfBT) Education Trust (http://www.cfbt.com) when they reviewed the Malaysian Education Blueprint 2013-2025. Both reports raised the concern about whether teachers were receptive towards SBA, and whether they could cope well with the shift of role in the SBA system. In these reports, the teams proposed that primary school teachers need to be

given more support in terms of understanding about the concepts of SBA and aspects to be taken into consideration in designing appropriate tasks. As mentioned by the EFA Global Monitoring Report team (2014), teachers could only maximize their potential “in the right context, with well-designed curricula and assessment strategies to improve teaching and learning” (p. i). Hence, there is a need to develop more helpful guidelines and provide supportive training courses for teachers so that the SBA can be effectively implemented in order to associate teaching and learning in the classroom.

Additionally, effective assessment strategies are needed in order to help teachers to assess their pupils with accordance to the newly introduced curriculum. As suggested by Lee and Sawaki (2009), the Cognitive Diagnostic Assessment (CDA) might be useful as it could assist teachers in the furnishing of detailed diagnostic information about a leaner's learning process in terms of mastery in various sub-skills in a specific skill of interest. Although the CDA is still a relatively new educational technique used in the formative assessment, it has recently retrofitted to norm-referenced tests (Jang, 2008). One of the main advantages of the CDA is that it does not require uni-dimensionality. In other words, the CDA could tap into an array of attributes or sub-skills, each of which could create a statistically separable dimension (Aryadoust, Akbarzadeh, & Akbarzadeh, 2011; Buck, 1994).

According to Alves (2012), there are three major features of CDA. First, formative inference obtained from the CDA helps teachers to monitor students’ learning progress. In addition, the CDA also encourages students to be engaged in their own learning by using the prepared assessment as a learning tool so that they will be aware of their own weaknesses and thus, overcome them for betterment. Second, diagnostic score from the CDA provides a diagnosis profile that includes each

student's mastery level on certain skills. Knowing about one's strength(s) helps in efficient and effective learning, whereas knowing about one's weakness(es) enables learners to be more mindful throughout the learning process. Lastly, the CDA also provides more detailed and in some way, more precise information about the knowledge/skills/processes than other forms of educational tests. Therefore, it is believed that the CDA could help teachers in selecting appropriate tasks in order to measure students' cognitive knowledge more accurately. Furthermore, the developed model can be used as an example for teachers' reference while designing tasks, not only for assessment purpose/s, but also for diagnostic purpose(s).

Nevertheless, the features of the CDA might also help in providing teachers valuable feedback about pupils' learning progress at the end of the tasks as it assesses pupils formatively and in minute details. Furthermore, it is believed that the CDA will be able to initiate and enhance pupils' learning interest conforming to different knowledge aspects included in the tasks (Alves, 2012). In short, the CDA can act as a catalyst as well as a motivational factor in improving teaching and learning quality. Moreover, the CDA is a dual-purpose instrument which gathers evidence from both teachers and pupils. On the part of teachers, the CDA will help to provide them with information about pupils' learning patterns which will be beneficial in future instructional planning and implementations. On the part of students, they will be informed about their weaknesses and strengths. Subsequently, they will have the opportunities to take pragmatic steps to overcome their weaknesses and at the same time, improve their strengths. Thus, CDA is potential to be used as a tool to help teachers implement the SBA by providing a framework to assess tasks design.

As stated by Tan Sri Dr Murad Mohammad Nor (21 January 2007, The Star Online), the former Education Director-General of the new national education system, "The

most important part of the implementation of any plan, is the teachers. However good the plan is, it will still be of no use if the teachers do not implement it well". Thus, efforts must be done in order to lighten teachers' burden in implementing the new system. CDA might act as a suitable measure to help teachers carry out the SBA in teaching and learning. However, to the best knowledge and review of the researcher, only few studies on CDA especially in mathematics had been done in Malaysia. Therefore, there is a need to develop a valid CDA to help teachers in assessing their students effectively.

Learning of time

"Time" is an important concept in our daily life. This concept basically is linked to mathematics, and is related to the representations, and symbol systems, such as clocks, calendars, lunar cycles, time-tables, etc. (Kelly, Miller, Fang & Feng, 1999).

In fact, children start to learn about the topic of time since they enter kindergarten. However, some primary school teachers reflect that their students are still weak in the comprehension and mastery of "Time", especially when it involves the calculation of duration and conversion of time units. To illustrate the statement above, a professor described an incident she met previously. There was once when she created an online quiz which was valid until October 12, 2014, 12:00pm. However, one of her students tried to do the quiz on October 12, 2014, 11:00pm, and eventually found out that she could not access the quiz. This incident shows that at times, even an undergraduate might be confused with the concept of *am* and *pm* especially when it comes to 12:00 am (midnight) and 12:00 pm (noon).

Since 1920s, studies on children's temporal concepts have been carried out (Burny, Valcke & Desoete, 2009). Researchers have concluded that "Time" is a complex

concept which is not easy to teach to children (Burny et al., 2009). McGuire (2007, p. 30, as cited in Harris, 2008) cited the Queensland Studies Authority by stating that "time does not easily link with the other topics of measurement due to its abstract nature: you can't go into a shop and buy a dozen minutes, or stub your toe on mid-day or an hour". Hence, "Time" could be a problematic learning content area for students.

In the Malaysian educational system, "Time" is only included in the syllabus of preschool and primary school, and very briefly covered in the Secondary One (Grade 7) mathematics syllabus. This indicates that if students do not master time-relevant skills in primary education and the beginning of Secondary One, they will not be taught and learn about it later in higher grades. Such situation might influence students' learning in other subjects too, as time is integrated in several subjects such as history (chronology), mathematics (mechanical time), geography (deep time) and literacy (time-related vocabulary) (Burny et al., 2009).

Since "Time" is a part of our important basic knowledge in life, teachers must apply appropriate methods to teach about "Time" and at the same time, identify students' problems/difficulties in mastering the concept of "Time". Yet, Burny et al. (2009, p. 488) states that "despite the fact that authors continuously point at the complexity of time conceptions and the difficulties involved for children, the available research does hardly help to understand what makes time so complex for children to grasp". Hence, this study introduces the application of CDA in teaching and learning which may serve as a tool to diagnose students' strengths and weaknesses in the classroom as well as help teacher to gain better understanding on students' cognitive processes in learning about "Time".

1.3 Research Objectives

This study aims to develop a cognitive diagnostic assessment (CDA) for the topic of "Time" in primary school learning. This assessment tool will be developed based on both experts' (teachers') and students' perspectives.

Specifically, the objectives include:

1. To develop a Cognitive Diagnostic Assessment (CDA) for primary mathematics learning of "Time";
2. To determine whether the list of time-related attributes is arranged in hierarchical structures based on students' responses; and
3. To determine the consistency between the expert-based cognitive model and student-based cognitive model.

1.4 Research Questions

Based on the objectives stated, this research intends to answer the following research questions:

1. (a) To what extent is the developed Cognitive Diagnostic Assessment valid?

   (b) To what extent is the developed Cognitive Diagnostic Assessment reliable?

2. (a) Are the attribute probabilities ordered from easy to difficult for each of the attribute hierarchies?

   (b) Does the correlation among attributes support the hierarchical structure of the attributes?

3. Is the expert-based cognitive model consistent with the student-based cognitive model?

1.5 Significance of the study

The findings of this study will benefit both primary school teachers and students in the following way.

First, by conducting assessments using the CDA, teachers will be able to identify students' strengths and weaknesses in learning about "Time". In other words, this assessment helps teacher to have a better insight on students' learning progress. Subsequently, it will contribute to more appropriate instructional decision-making on the remedial works and effective instructional planning. In addition to that, the diagnostic feature of the CDA could also fulfil the characteristics of school assessment as directed by the Malaysian MOE (2014).

Second, this assessment might be a tool to help teachers adapt themselves better to the new education system in Malaysia. In the School Based Assessment, teachers are required to design tasks for students' assessment purpose. At this initial stage of the educational transformation, teachers might need some guidelines to construct the tasks and at the same time, avoid the practice of assessment-oriented teaching, which immensely defeats the purpose of assessments. Furthermore, the CDA could act as a reference for teachers to learn about designing meaningful tasks to assess students' knowledge or skills mastery level instead of assessments of achievement that merely focus on students' scores.

Lastly, students can do self-evaluation by using the CDA assessment method for self-improvement purpose. It helps students to identify their weaknesses and thus, students will focus and take initiatives to fix their problematic learning areas/aspects. This encourages and motivates students to initiate a self-learning process, which is also one of the main purposes of the implementation of SBA in national schools.

Therefore, Cognitive Diagnostic Assessment (CDA) is proposed as a framework for teachers to come out with more effective and better assessment items or tasks. The CDA also helps to expose teachers to a more effective way of diagnosing pupils' learning difficulties as well as improving their mastery levels on skills and sub-skills via the understanding of their cognitive processes.

1.6 Limitations of the study

There are some limitations in this study.

First, time constraint during the administration of instrument is the major limitation of the study. Before the instrument was carried out, pupils were briefed on the purpose of the assessment which emphasizes on the diagnostic result rather than the total score. As the instrument is a type of diagnostic assessment, it is expected to be carried out without any time limitation. In fact, pupils are allowed to have sufficient time to complete the assessment. However, in this study, the designed instrument could only be administered during school hours. Consequently, the duration of administration was set to only 60 to 90 minutes to avoid any interruption on normal school activities. Due to the time constraint, pupils who were involved in this study might not be able to complete the assessment on time, which might cause anxiety among the participating pupils. Hence, in order to reduce the negative impacts of this limitation, extra time was given to pupils who could not complete the assessment within the given period.

Second, the sample of this study was only limited to Year Six pupils from 11 primary schools in Penang and thus, the result might not be able to generalize to all primary school pupils in Malaysia.

Last but not least, the original idea of think-aloud session could not be carried out as pupils were not familiar with this type of method, and sufficient training was not given to pupils. Researchers only managed to give a simple briefing about the session and demonstrate the method of think-aloud before the pupils commence the session. As a result, some pupils were quite passive during the session and shared their thoughts only when they were prompted by the researchers. Instead of performing the think-aloud session, a think-aloud interview session was carried out in this study. Researchers prompted the pupils with some questions such as, "Can you tell me what you are thinking now?", "Why are you doing addition?" when they did not tell their thoughts while solving the items.

1.7 Definition of terms and operational definitions

Panel of experts: Panel of experts refers to the group of experienced primary mathematics teachers who take part in this study.

Attribute(s): The knowledge, skills or processes that needed in order to solve a given task (Alves, 2012; Roberts, Alves, Chu, Bahry, Gotzmann & Thompson, 2012; Tatsuoka & Tatsuoka, 1997).

Cognitive model: "A simplified description of human problem-solving on standardized tasks at some convenient precisions or level of details in order to facilitate explanation and prediction of students' performances, including their strengths and weaknesses" (Leighton & Gierl, 2007b, p. 6)

Expert-based cognitive model: The cognitive model formed based on the perception of panel of experts.

Student-based cognitive model: The cognitive model formed based on the perception of pupils.

Attribute probability: A statistical numeral that shows the probability of an examinee's mastery level on a specific attribute (de la Torre & Minchen, 2014).

Primary pupils: Children who are receiving primary education in Malaysian primary schools.

The topic of "Time": Concepts of "Time" as which included in Malaysian Primary Mathematics Curriculum Specifications.

Cognitive Diagnostic Assessment: Assessment which identify specific attributes exhibited by learners to provide learners' cognitive strengths and weaknesses (Leighton & Gierl, 2007a).

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

This chapter presents the literature reviews on the definition of an assessment, Cognitive Diagnostic Assessment and its past-related studies as well as the Cognitive Diagnostic Model or Psychometric Model. Based on these reviews, a theoretical framework and conceptual framework of this study are developed.

2.1 Assessment

An assessment is a crucial process to make judgment on students' learning progress and it is widely used in education as a tool to categorize and evaluate students. According to The Assessment Standards for School Mathematics (NCTM, 1995), an assessment should be seen as a process to obtain a valid and relevant evidence on the students' competencies in a subject domain. This is supported by White (2007) who defined assessment as the process of gathering information and making inferences on students' competencies, needs, strengths and weaknesses. Besides, assessment is also a tool to evaluate the quality of the current education system (Wu & An, 2007).

There are many methods of developing an assessment. Examples of traditional assessment methodology are paper and pencil tests, multiple-choice tests, true/false tests and essays. These days, an assessment can also be carried out in many different ways such as authentic assessment, performance-based assessment and constructivist assessment (Dikki, 2003). These methodologies of assessments have received numerous views and opinions on developing a thorough evaluation of the student. Buhagiar (2007) argued that traditional forms of assessment or commonly known as examination and test are 'unfriendly' in four aspects, namely: learning unfriendly,

curriculum unfriendly, students unfriendly and teacher unfriendly. These traditional kind of assessment were quoted as learning and curriculum unfriendly because they revealed the things that the learners remembered rather than providing the whole picture of the mastery level of their knowledge and skills. Thus, the results of the assessment are non-informative in nature.

In addition, most traditional examinations and tests focused more on rote learning rather than conceptual understanding and higher order thinking skills. Thus, it was said to be teacher unfriendly as teachers were making assumptions based on the scores given. The purpose of the assessments was merely for an improvement in the scores but not tracking students' learning progress. Moreover, these examinations and tests were said to be student unfriendly as scores that they obtained were use to classify them into "good students" and "weak students'. If they obtained a higher score then they would be labeled as "good students", conversely, they would be labeled as "weak students".

This phenomenon has twisted the original purpose of assessment which is to gather their information to determine the students' abilities in a subject domain. Broadfoot (1996) stated that these traditional assessments demotivated the students from moving forward in learning and is pushing most of them away from the learning system. It is certain that the information from the assessment actually plays a vital role in teaching and learning process. A teacher should have the feedback and reviews about their students to find out their learning progress, their talents and their weaknesses. A profile of student's performances can reveal many things about a student. Without all these information, a teacher would not be able to monitor their student's learning progress.

There is a phenomenon of overemphasizing on a student's achievement in the centralized public examinations that was set by many developed countries in their education system. Malaysia is one of the countries that implemented the centralized public examination. In these examinations, the focus of the teachers, parents and students will be on the student's performances and results in the public examination (Ong, 2010), by publicizing the number of "As" attained by the students. Che Noraini et al. (2013) commented that this system was a burden to the students and are very impractical to reveal the students' competencies effectively. Therefore, the School Based Assessment System (SBA, or known as *PBS* in the Malay language) was introduced by the Ministry of Education Malaysia (MOE Malaysia) in the year 2011 to challenge the issue.

2.1.1 School examination system in Malaysia

There were four public examinations in Malaysian education system. The first public examination was the *Ujian Penilaian Sekolah Rendah* [*UPSR*, Primary School Achievement Test] which will be taken in the final year of primary education (the 6th year). This centralized examination was responsible to evaluate the achievement of the primary school pupils after they undergo a total of six years of education, and to some extent it acted as the criterion for a placement in secondary schools.

After three years of secondary school education, the students will take *Penilaian Menengah Rendah* [*PMR*, Lower Secondary Examination], another centralized examination, and according to their achievements, they will be streamed in their higher secondary school education.

At the end of two years higher secondary school education, students will take *Sijil Pelajaran Menengah* [*SPM*, Malaysian Certificate Examination] as their last public

examination in secondary school. After completing a total of 11 years of education, as stated in the Education Act 1992, the students have the option to choose their pre-university program by studying Form 6. They will take the next public examination, *Sijil Tinggi Pelajaran Malaysia* [*STPM*, Malaysian Higher Certification Examination] to proceed to a higher level of education (Ong, 2010).

Apart from the centralized examinations, continuously there were also other school-based assessments such as science practical for science subjects and oral reading tests for language subjects which were introduced in 1997. Malaysia education system however were putting focus on the students' result from public examination instead to determine their progression to higher level of education or job opportunities (Ong, 2010). Thus, the scoring report from these exams usually consisted of the exam grade for each of the subjects which did not reflect the students' competencies throughout the years of study effectively.

In the year of 2011, School-Based Assessment [SBA] system was implemented in the Malaysian education system as a part of the education reform. Compared to the central public examination, it is a better approach and also a holistic assessment system to evaluate student's achievement and learning progress. Centralized assessment is still a part of the new system, but it will be administered by the school administration and exists for the purpose of a summative assessment. SBA is an on-going school assessment which will be carried out throughout the schooling years. Teachers are empowered to evaluate their students in the teaching and learning process (Circular from MOE, dated 31 March 2014). They will in-charge of the planning, test items construction, and the assessment instrument, administration, scoring, recording and reporting.

According to MOE (2014), SBA consists of the following characteristics: (a) holistic, the assessment is able to provide complete information on the knowledge and skills achieved by a student; (b) continuous, whereby the assessment is aligned with the teaching and learning process; (c) flexible, in terms of the way to access is based on student's readiness and suitability; and (d) reference to the achievement standards stipulated in the curriculum standard.

In this system, students' results will not be compared with their peers, but rather for the objective of monitoring their learning progress. There were guidelines prepared by the Malaysian Examination Syndicate to monitor student's learning progress, identify student's strengths and weaknesses, obtain information on student's progress and plan for effective teaching as well as to report student's progress to educational stakeholder effective feedback on a teaching. Guidelines to plan and to set the target at the beginning of year/lesson were also provided. Hence, teachers have to equip themselves with a variety of assessment methods and techniques to gain a much more constructive feedback from the assessment (MOE, 2014).

2.2 Cognitive Diagnostic Assessment (CDA)

Snow and Lohman (1989, as cited in Gierl & Alves, 2010) stated that the basic of educational tests is a cognitive problem-solving task. Later, researchers have suggested integrating cognitive sciences and educational assessment. The integration is aimed to produce inferences on students learning progress and to improve the teaching and learning experience (Nichols, 1994; Nichols, Chipman & Brennan, 1995).

According to Jang (2009), CDA is an alternative assessment that aims to provide fine-grained analysis of learners' skill mastery profiles and their cognitive

knowledge state. This minute result is crucial as two examinees who score the same marks in a test might have different levels of mastery skill (or cognitive levels). It helps the examinees to take essential actions to fulfill the gap between their current competency levels and their desired learning goals (Black & William, 1998). Learners' strengths and weaknesses in a subject domain can be diagnosed by using CDA (Wu, Chen, Sung & Chang, 2012; Ye, 2005). CDA can also be used to measure learners' cognitive structures and processing skills (Leighton & Gierl, 2007b).

According to Nichols (1994), there are five steps in developing a CDA.

(1) Substantive theory construction: A development of a model which involved the knowledge and skills that are intended to be measured. The construction can be done based on the original research and best understanding on learners' learning process. Then, test developer will design the items based on the constructed theory.

(2) Design selection: Selection of the observation design (the characteristics of the task) and measurement design (the psychometric model used in analysis) will be done at this step. The selection will ensure assessment activities are designed and arranged in a way to reveal the cognitive processes involved when learners are responding towards the items.

(3) Test administration: The administration aims to evaluate the designed items and, to revise the constructed substantive theory.

(4) Response scoring: After applying the chosen measurement design on the observed responses, a score report regarding learners' mastery level on different attributes is generated.

(5) Design revision: A process of obtaining evidences from the observed responses about the chosen observation design and measurement design in step 2 will be done. For example, the consistency between the expected and observed responses can be used as an indicator.

However, the development of CDA is time consuming and requires a lot of human resources. As such, not every research develops a new CDA. Some researchers chose to make use of the existing large-scale assessments such as SAT (Jang, 2009; Wang & Gierl, 2011) and TIMSS (Birenbaum, Tatsuoka & Yamada, 2004; Su, Choi, Lee, Choi, & McAninch, 2013) for cognitive diagnostic purposes. This process is known as retrofitting of CDA. In developing a new CDA, a cognitive model is first constructed then followed by item design, while in retrofitting of CDA, a cognitive model is established based on the review of the existing items. In other words, a psychometric model is applied to an existing data set which is collected from an existing test (Gierl & Alves, 2010).

Retrofitting of CDA does not involve the lengthy procedures as in developing a new CDA. However, experts claimed that retrofitting of CDA yields only a few successful applications due to the specific requirement of the proposed cognitive model (Gierl & Cui, 2008). Cognitive model was constructed by relating the defined set of knowledge/skills/processes, thus, the existing test items might not be enough to examine all the knowledge and skills specified in the cognitive model.

2.2.1 Advantages of using CDA

CDA emphasizes more on identifying students' response patterns which can be used to classify them into different meaningful categories. Alves (2012) has summarized the strengths of using CDA as educational assessment in four aspects. First, CDA

can be used for both summative and formative purposes. The purpose of assessment must be clearly identified before an assessment is administered. Different purpose in assessment will lead to a different evaluation method. CDA can be used summatively at the end of an academic year to capture the overall learning situation of students throughout the year. Meanwhile, it can also be used as formative assessment to investigate the knowledge and skills that are required to emphasize in the next instructional planning. It provides a thorough feedback for teacher to make an instructional decision. (Ketterlin-Geller & Yovanoff, 2009).

Second, CDA provides education stakeholders a detailed score on learners' performance. The scoring report that was generated from CDA is known as the cognitive profile. This profile shows learners' mastery level on each of the attributes that was measured through the assessment. Thus, learners' strengths and weaknesses in an interested domain will be shown as the results of the report. This is an imperative and meaningful feedback on learners' mastery level which will help learners to have a clearer picture about the level of their competencies and their intended learning goals. The acquired information can greatly motivate them to put pragmatic action(s) towards achieving their learning goals.

Third, more detailed and meaningful feedback on the measurement of knowledge and skills can be obtained by using CDA. Besides revealing the learners' mastery level on each attribute, CDA also allows education stakeholders to fully understand learners' learning process and cognitive structure. In other words, through the responses in CDA, teachers can easily understand the students' ways of thinking and their learning sequences. Lastly, CDA is possible to provide quality diagnostic feedback to learners. The feedback provided is constructive in nature to both

educational stakeholders and learners as it depicts the strengths of the learners, at the same time, presenting the areas of improvement.

2.2.2 Attributes

Attribute are defined as any knowledge, skills or cognitive processes that a student requires to solve a test item (Alves, 2012; Roberts et al., 2012; Tatsuoka & Tatsuoka, 1997). Attributes are very specific, in other words, they are very detailed in nature. Attributes that are too coarse, or not well described, might display many potential problems to the test items and, might result in the absence of precise information and conclusion at the end of the test. As stated by Ketterlin-Geller, Jung, Geller and Yovanoff (2008), attributes can be identified by using four methods: item analysis, mathematics textbook review, expert review and verbal protocols. The selection of methods would be depending on the availability of time and human resources of the studies.

The defined set of attributes is linked together to form a cognitive model, which served as a framework for the items development later (Alves, 2012). For example, in mathematics, the ability of a student to convert an hour into minutes correctly is considered as an attribute. The ability to calculate correctly the time taken to do a specific task within a specific time frame is another attribute. In languages subject, the ability to use tenses correctly is also an attribute. The same goes to the ability to use prepositions. In short, an attribute is the knowledge/skills/processes required to give the correct response to any problem or given task related to a specific task in a specific domain of interest.