DEVELOPMENT AND EVALUATION OF ONLINE COGNITIVE DIAGNOSTIC ASSESSMENT WITH FEEDBACK ON THE CONCEPT OF TIME IN YEAR FOUR AND FIVE MATHEMATICS

CHIN HUAN

UNIVERSITI SAINS MALAYSIA

2020

DEVELOPMENT AND EVALUATION OF ONLINE COGNITIVE DIAGNOSTIC ASSESSMENT WITH FEEDBACK ON THE CONCEPT OF TIME IN YEAR FOUR AND FIVE MATHEMATICS

by

CHIN HUAN

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

December 2020

ACKNOWLEDGEMENT

The completion of this thesis would not have been possible without the support and encouragement of several special people. Hence, I would like to take this opportunity to show my gratitude to those who have assisted me in a myriad of ways.

First and foremost, I would like to express the deepest appreciation to my supervisor, Associate Professor Dr. Chew Cheng Meng. Based on the classic Chinese culture, he would be considered as my respected *shī fù* (师父) in my journey of doctoral study. I have been extremely lucky to have him accepting me as his master student as well as his graduate research assistant. He guided me with big doses of patience for converting my candidature level from a master's degree to a PhD. His ongoing support and guidance turned this tough journey into a pleasant experience. He always prepared to sit and listen to my troubles and always made me feel as if I am progressing well. With his close supervision and prompt responses, I never felt lost even facing the challenges in data collection. His willingness to offer me so much of his time and intellect is the major reason this thesis was completed on time.

Besides that, I would like to express my sincere gratitude to the former Dean of School of Educational Studies, Associate Professor Dr. Shaik Abdul Malik Mohamed Ismail and the former Deputy Dean (Research, Innovation and Industry-Community Network) of School of Educational Studies, Associate Professor Dr. Salmiza Saleh for their supports and kind assistance given regarding my PhD upgrading. I am particularly grateful for the recommendation given by the PhD Upgrading Evaluation Committee and the approval of the PhD upgrading granted by the University Board of Graduate Studies, USM.

I would like to give special thanks to Dr. Wun Thiam Yew and Professor Dr. Munirah Ghazali for their effort in reviewing my work and the constructive comments and suggestions given to improve my research work. I would like to express my gratitude to Associate Professor Dr. Lim Hooi Lian and Dr. Thien Lei Mee, as well as Dr. Mastura Baharudin and Dr. Nur Syibrah Muhamad Naim for giving me insightful comments and suggestions during item development process, and the web application development process respectively. I would like to thank Mr. Muhamad Azam Haris for his enormous contribution in developing the online Cognitive Diagnostic Assessment web application within a short period of time.

Special thanks to the experienced teachers and the psychometric experts who were involved in the instrument validation process. I am particularly grateful for the assistance given by Mr. Ng Kiang Wah, the officer in district education office, in identifying the school with suitable facility for conducting the study. I would also like to express my gratitude to the headmasters of the school for giving me permission to conduct my study in their school. I would like to give special thanks to the teachers and pupils for giving me support during the data collection process.

Taking this opportunity, I also would also thank the faculty staff in School of Educational Studies for their helpful and warm behaviours. Not forgetting to all my friends in postgraduate lounge, School of Educational Studies, for bringing me so much joy and fun in my PhD journey, besides motivating me when I faced various challenges. Lastly, I would like to offer my special thanks for my parents and my siblings for their unconditional and infinite love, support, patience, and sacrifice.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

A matrix	Adjacency matrix
AE	Assessment Engineering
AHM	Attribute Hierarchy Method
ANCOVA	Analysis of Covariance
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
BEAR	Berkeley Evaluation and Assessment Research Centre
CDA	Cognitive Diagnostic Assessment
CDM	Cognitive Diagnostic Model
CDS	Cognitive Design System
CMS	Content Management System
CSS	Cascading Style Sheets
CTT	Classical Test Theory
CVI	Content Validity Index
DE	Distractor Efficiency
DNS	Domain Name System
DSKP	*Dokumen Standard Kurikulum dan Penilaian* (Curriculum Standard and Assessment Document)
e-ATLMS	Web-based Assessment in Teaching and Learning Management System
ECD	Evidence-Centred Design
EF	Elaborated Feedback
EM2	Eliciting Mathematics Misconceptions
EPRD	Education Planning and Research Division
FD	Functioning Distractor

GDM	General Diagnostic Model
HCI	Hierarchy Consistency Index
HOTS	Higher Order Thinking Skills
HTML	Hypertext Mark-up Language
I matrix	Identity matrix
ICT	Information and communication technology
I-CVI	Item-level Content Validity Index
IRT	Item Response Theory
JePEM	*Jawatankuasa Etika Penyelidikan Manusia* (Human Research Ethics Committee)
JISC	Joint Information Systems Committee
KBSM	*Kurikulum Bersepadu Sekolah Menengah* (Intergrated Curriculum of Secondary School)
KCR	Knowledge of Correct Response
KR	Knowledge of Result
KSPK	*Kurikulum Standard Prasekolah Kebangsaan* (National Preschool Standard Curriculum)
KSSM	*Kurikulum Standard Sekolah Menengah* (Secondary School Standard Curriculum)
KSSR	*Kurikulum Standard Sekolah Rendah (*Primary School Standard Curriculum)
KWSK	Know What Student Know
LOTS	Lower Order Thinking Skills
MCQ	Multiple Choice Question
MLP	Multilayer Perceptron Neural Network
MOE	Ministry of Education Malaysia
NCTM	National Council of Teachers of Mathematics

NEA	Neuman Error Analysis
OECD	Organisation for Economic Co-operation Development
OMC	Ordered Multiple Choice
OMR	Optical Mark Recognition
PHP	Hypertext Pre-processor
PISA	Programme for International Student Assessment
PPKT	*Pusat Pengetahuan, Komunikasi dan Teknologi* (Centre for Knowledge, Communication and Technology)
Q matrix	Incident matrix
Q_r matrix	Reduced Q matrix
R matrix	Reachable matrix
RCI	Response Consistency Index
RDBMS	Relational Database Management System
RMSE	Root-Mean-Squared Error
RSM	Rule Space Method
S-CVI	Scale level Content Validity Index
SDLC	System Development Lifecycle
SFTP	Secure File Transfer Protocol (SFTP)
SJKC	National Type Chinese Primary School
SJKT	National Type Tamil Primary School
SK	National Primary School
SMART	Specific Mathematics Assessments that Reveal Thinking
SQL	Structured Query Language
SSE	Sum Squared Error
TIMSS	Trend in Mathematics and Science Study
XAMPP	Cross-platform (represent as ‘X’), Apache, MySQL, PHP and Perl

LIST OF APPENDICES

PEMBANGUNAN DAN PENILAIAN PENTAKSIRAN DIAGNOSTIK KOGNITIF ATAS TALIAN DENGAN MAKLUM BALAS UNTUK KONSEP MASA DALAM MATEMATIK TAHUN EMPAT DAN LIMA

ABSTRAK

Masa merupakan konsep matematik yang susah untuk murid Tahun Empat dan Lima. Tujuan kajian ini adalah untuk membina dan menilai Pentaksiran Diagnostik Kognitif (PDK) atas talian dengan maklum balas untuk mentaksir pencapaian murid Tahun Empat dan Lima dalam topik '*Masa*'. Pembinaan PDK atas talian dengan maklum balas dijalankan pada fasa pertama kajian dan penilaian PDK atas talian dengan maklum balas dijalankan pada fasa kedua kajian. Reka bentuk kajian berkepelbagaian kaedah telah digunakan dalam kajian ini. Sejumlah 28 model kognitif Tahun Empat berasaskan pakar dan 21 model kognitif Tahun Lima berasaskan pakar telah dikenal pasti oleh pakar pendidikan matematik. Kesahan model kognitif berasaskan pakar tersebut telah dinilai oleh enam pakar isi kandungan daripada Sekolah Kebangsaan (SK), Sekolah Jenis Kebangsaan Cina (SJKC) dan Sekolah Jenis Kebangsaan Tamil (SJKT), bagi aspek (a) perkaitan antara atribut dengan standard kandungan dan standard pembelajaran Matematik Tahun Empat dan Tahun Lima; dan (b) susunan atribut secara hieraki. Sejumlah 28 peta konstruk Tahun Empat dan 21 peta konstruk Tahun Lima telah diperolehi daripada model-model kognitif berasaskan pakar. Kesahan peta-peta konstruk tersebut telah dinilai oleh panel pakar isi kandungan yang sama. Kesahan isi kandungan PDK atas talian dengan soalan Aneka Pilihan Tersusun (APT) dalam bahasa Melayu, Cina dan Tamil juga telah dinilai oleh pakar isi kandungan bagi aspek: (a) klasifikasi atribut berdasarkan Taksonomi Bloom dengan persetujuan peratus mudah 100 peratus; (b) perkaitan item dengan atribut yang

diukur dengan S-CVI lebih daripada .93; dan (c) liputan item berdasarkan standard kandungan dan standard pembelajaran Matematik Tahun Empat dan Tahun Lima dengan S-CVI 1.00 kecuali soalan dalam subtopik '*Bahagi Masa*'. Selain itu, kesesuaian tahap yang diperuntukkan bagi setiap pilihan soalan APT juga dinilai oleh dua pakar psikometrik. Selepas proses penilaian kesahan, PDK dengan soalan APT telah dirintiskan kepada 90 orang murid Tahun Empat (SK: 30; SJKC: 48; SJKT: 12) dan 129 murid Tahun Lima (SK: 48; SJKC: 66; SJKT: 15) dari negeri Pulau Pinang, Malaysia yang dipilih dengan menggunakan persampelan mudah. Hasil analisis item menunjukkan bahawa ketiga-tiga versi PDK dengan soalan APT terdiri daripada soalan dengan indeks diskriminasi yang memuaskan walaupun sebahagian daripadanya adalah terlalu mudah bagi murid dari SK, SJKC dan SJKT. Hasil analisis distraktor menunjukkan bahawa setiap soalan atau sub-soalan APT terdiri daripada sekurang-kurangnya satu distraktor berfungsi secara purata dan dapatan ini adalah dapat diterima. Pada peringkat pentaksiran, ketiga-tiga versi PDK dengan soalan APT adalah boleh dipercayai dengan nilai KR-20 dalam lingkungan .57 hingga .97, di mana .50 adalah nilai miminum yang boleh diterima bagi pentaksiran yang mengandungi kurang daripada 15 soalan. Analisis respons murid untuk setiap soalan menunjukkan bahawa atribut dalam model kognitif berdasarkan murid disusun megikut peningkatan kerumitan seperti yang ditunjukkan oleh penurunan pola min atau median kebarangkalian atribut. Di samping itu, susunan hieraki atribut bagi model kognitif berdasarkan murid juga disokong oleh korelasi antara kebarangkalian atribut seperti yang ditunjukkan oleh penurunan pola Koefisien Kolerasi Rank Spearman. Konsistensi model kognitif berasaskan pakar dengan model kognitif berasaskan murid mencatat tahap sederhana-tinggi dengan nilai keseluruhan Indeks Konsistensi Hierarki (IKH) dalam lingkungan .60 hingga 1.00. Kandungan maklum balas ringkas dan

terperinci juga dinilai sah oleh pakar isi kandungan. Fasa pertama kajian ini diakhiri dengan pembinaan CDA atas talian dengan maklum balas sebagai sebuah aplikasi web. Kemudian, keberkesanan CDA atas talian dengan maklum balas telah dinilai pada fasa dua kajian dengan menjalankan kajian kuasi-ekperimen. Sampel kajian tersebut mengandungi 97 murid Tahun Empat (SK: 39; SJKC: 41; SJKT: 17) dan murid Tahun Lima (SK: 56; SJKC: 42; SJKT: 27) dari negeri Pulau Pinang, Malaysia yang dipilih dengan menggunakan persampelan bertujuan. Perbezaan yang signifikan antara kumpulan kawalan dan kumpulan eksperimen menunjukkan bahawa maklum balas terperinci dalam PDK atas talian Tahun Empat dan Tahun Lima adalah lebih berkesan berbanding dengan maklum balas ringkas. Berdasarkan dapatan analisis tematik dan analisis kes bersilang, kebergunaan maklum balas dan kelengkapan kandungan maklum balas adalah dua faktor utama yang menyumbang kepada perbezaan keberkesanan bagi kedua-dua jenis maklum balas tersebut. Kajian ini telah menunjukkan kelebihan dalam pengintegrasian maklum balas dengan PDK atas talian. Oleh itu, guru-guru akan digalakkan untuk menggunakan PDK atas talian sebagai alat pentaksiran bilik darjah untuk memberi maklum balas terperinci segera dalam menyokong pembelajaran murid untuk topik ‘Masa’.

DEVELOPMENT AND EVALUATION OF ONLINE COGNITIVE DIAGNOSTIC ASSESSMENT WITH FEEDBACK ON CONCEPT OF TIME IN YEAR FOUR AND FIVE MATHEMATICS

ABSTRACT

Time is a difficult mathematical concept for Year Four and Five pupils. This study aimed to develop and evaluate an online Cognitive Diagnostic Assessment (CDA) with feedback for assessing Year Four and Five pupils' achievement in the topic of '*Time*'. The development of the online CDA with feedback was conducted in the first phase of the study while the evaluation of the online CDA with feedback was conducted in the second phase of the study. Multimethod research design was employed in the study. The 28 Year Four expert-based cognitive models and 21 Year Five expert-based cognitive models were specified by mathematics education experts and judged to be valid by six subject matter experts from National School (SK), National-Type Chinese School (SJKC) and National-Type Tamil School (SJKT) respectively, in terms of (a) relevance of the attributes to Year Four and Year Five Mathematics content and learning standards; and (b) hierarchical arrangement of attributes. The 28 Year Four construct maps and 21 Year Five construct maps derived from the expert-based cognitive models were then judged to be valid by the same panel of subject matter experts. The contents of the online CDA with Ordered Multiple-Choice (OMC) items in Malay, Mandarin and Tamil respectively were also judged to be valid by the subject matter experts in terms of: (a) classification of attributes based on the revised Bloom's Taxonomy with simple percent agreement of 100 percent; (b) relevance of the items to the attributes intended to be measured with a S-CVI of more than .93; and (c) coverage of the items based on the Year Four and Year Five

Mathematics content and learning standards with a S-CVI of 1.00 except the items in subtopic of '*Division of Time*'. Further, the assignments of level to each option of the OMC items were judged to be appropriate by the two psychometric experts. After the validation process, the CDA with OMC items was piloted to 90 Year Four pupils (SK: 30; SJKC: 48; SJKT: 12) and 129 Year Five pupils (SK: 48; SJKC: 66; SJKT: 15) in Penang state, Malaysia, which was selected by convenience sampling. The findings of item analysis indicated that the three versions of CDA with OMC items consisted of items with satisfactory discrimination power although some of them were too easy for the pupils from SK, SJKC and SJKT. The distractor analysis indicated that each OMC item or sub-item consisted of at least one functional distractor on average and is accepted. At the assessment level, the three versions of the CDA with OMC items were reliable with the values of KR-20 ranging from .57 to .97 whereby .50 is the minimum threshold for an assessment with less than 15 items. Analysis of the pupils' item responses indicated that the attributes in the pupil-based cognitive models were arranged in an increasing order of complexity as shown by the decreasing pattern of the mean or median of attribute probability. In addition, the hierarchical arrangements of attributes of the pupil-based cognitive models were supported by the correlation among the attributes probabilities as shown by the decreasing pattern of Spearman's Rank Order Coefficients. The expert-based cognitive models were also moderately high consistent with the pupil-based cognitive models with the values of the overall Hierarchical Consistency Index (HCI) ranging from .60 to 1.00. The simple and detailed feedback content were also judged to be valid by the subject matter experts. The first phase of the study ended with the development of the online CDA with feedback as a web application. The effectiveness the online CDA with feedback was then evaluated through conducting a quasi-experimental study in phase two. The

sample of the study consisted of 97 Year Four pupils (SK: 39; SJKC: 41; SJKT: 17) and 125 Year Five pupils (SK: 56; SJKC: 42; SJKT: 27) in Penang state, Malaysia which was selected by using purposive sampling. The significant difference between the control groups and experimental groups indicated that the detailed feedback in Year Four and Year Five online CDA is more effective than the simple feedback. Based on the findings of thematic analysis and cross-case analysis, the usefulness of the feedback and the comprehensiveness of the feedback are the two main factors which contributed to the difference of effectiveness for the two types of feedback. . This study sheds light on the advantages of incorporating feedback in online CDA. Thus, the teachers might be encouraged to use it as classroom assessment tool for providing instant detailed feedback in supporting students' learning for the topic of 'Time'.

CHAPTER 1

INTRODUCTION

1.1 Introduction

'Time' is an underestimated topic which is considered as a difficult topic to the primary school children (McGuire, 2007). Regardless of the complexity of this topic, the problems pupils encountered when learning this topic are rarely explained in the available research (Burny, Valcke, & Desoete, 2009). Moreover, the available diagnostic assessment tools which can be used to gauge pupils' cognitive strengths and weaknesses in this particular topic are considered scarce (Sia, 2017). Consequently, typical formative classroom assessments are used to measure pupils' understanding and diagnose their problems in learning this topic. Unfortunately, the formative classroom assessments have been criticised for unable to provide adequate diagnostic information (Herrera, Murry, & Cabral, 2012) which is meant for making instructional decision to tailor pupils' need. To date, a relative new form of diagnostic assessment, named Cognitive Diagnostic Assessment (CDA) might serve as an alternative assessment which could provide the education stakeholders more detailed information about pupils' learning progress for the topic of 'Time' (Sia & Lim, 2018).

According to Nichols (1994), CDA appears as the integration of cognitive psychology and assessment practices for making inferences about pupils' specific knowledge structure. The valid cognitive models which hypothesise pupils' cognitive process for correctly solved the tested problem are used to guide the item development and infer pupils' performance in terms of skill mastery. Thus, CDA could provide richer and meaningful information which can be used directly to support teaching and learning. Specifically, this diagnostic information can direct the classroom instruction such as planning of remedial instruction and placement of pupils into related intervention programme (Ketterlin-Geller & Yovanoff, 2009). In addition, the comprehensive CDA results could be a good prompt for the student to reconsider their learning strategy (Alves, 2012). In short, CDA could be a worthwhile assessment which would benefit the education stakeholders.

Despite the strength of CDA in supporting the classroom assessment practice, the CDA is rarely being implemented in the classroom. The usability issues such as tedious operation procedures and the significant delay between assessment administration and the receiving of the scoring report and corrective feedback reduce the potential use of CDA in a real classroom setting (Huff & Goodman, 2007). Nevertheless, Pellegrino, Chudowsky, and Glaser (2001) asserted that these practical constraints can be minimized by recent and upcoming technology advancement. Thus, an online Cognitive Diagnostic Assessment with feedback would be developed in this study to provide immediate support to the needs of educational stakeholders for the betterment of the learning of 'Time'in primary school.

1.2 Background of the Study

This section provides information which establishes the context of the study. Specifically, the brief introduction about cognitive diagnostic assessment and the learning of mathematical concept of times are presented in the following subsections.

1.2.1 Cognitive Diagnostic Assessment (CDA)

Cognitive diagnostic assessment (CDA) is an alternative form of assessment which emerges due to the fusion of cognitive psychology and psychometrics in educational measurement in the mid-1980s (Leighton & Gierl, 2007b). Particularly, CDA falls under the umbrella of diagnostic assessment and it is primarily used to “measure the specific knowledge structures and processing skills of the pupils so as to provide information about their cognitive strengths and weaknesses” (Leighton & Gierl, 2007b, p. 3).

In contrast with other assessment, the construction of CDA is not relying solely on the test specification tables which often being criticised for only specifying the content requirement and rarely giving consideration to the types of cognitive skills underlying the curriculum (Huff & Goodman, 2007). The item development of CDA is guided by the cognitive models which specify the incremental development of pupils’ understanding in the tested domain (Ketterlin-Geller, 2016). In mathematics, cognitive models usually refer to the learning trajectories (Ketterlin-Geller, 2016).

Cognitive models are the key component which makes CDA different from other diagnostic assessment. The test developers may either adopt the existing cognitive models or develop new cognitive models. Nonetheless, Gierl and Leighton (2007) highlighted that the test developers must ensure the cognitive models are supported by empirical evidence as the premature cognitive models may have limited explanatory power. After applying the psychometric model to analyse pupils’ item response pattern, these cognitive models might hardly to be linked their performance to the inferences made about their cognitive strengths and weaknesses (Leighton & Gierl, 2007b). Thus, it believed that CDA could provide more valid and meaningful diagnostic information compared to other assessment.

Unlike typical assessment, in which the pupil’s score is usually identified by determining his or her location along a single proficiency continuum (de la Torre & Minchen, 2014), CDA can be tapped into an array of interrelated attributes or sub-skills (Boora, Pasiphol, & Tangdhanakanond, 2015; Ketterlin-Geller & Yovanoff,

2009) and yields a scoring report with specific information about pupil's mastery of attributes in the tested domain (Roberts & Gierl, 2010). The detailed diagnostic information obtained from the CDA can be directly translated into pragmatic action that can be taken by teachers to adapt and respond to pupils' needs (de la Torre & Minchen, 2014). Moreover, detailed diagnostic information yielded in CDA will also prompt pupils to overcome their weaknesses and develop their strength and potential at the same time. Thus, CDA can be developed for classroom use.

To increase the potential use of the CDA, a more instructional relevant score report should be provided to the educational stakeholders. The CDA scoring report is fundamentally different from the ordinary scoring report of school examinations and national assessment. Instead of percentage correct score and the grade, the attribute-based diagnostic information such as attribute properties and classification of attribute mastery are reported to the educational stakeholders. This information is essential for promoting pupils' learning. Meanwhile, the corrective feedback should be given to students in a timely manner to help them overcome their cognitive weaknesses. In short, there is a need to develop a comprehensive and coherent cognitive diagnostic assessment with feedback to support teachers' classroom assessment practice.

1.2.2 The Learning of Mathematical Concept of Time

'Time' is a basic mathematics concept that is closely related to daily life. Acquisition of time concept enables us to schedule our daily life activities such as when to have our breakfast, when to go to school and so on. More particularly, the point of time (labelled by clocks, and calendars), and durations (measuring elapsed time) are the two main mathematical concepts of time (Harris, 2008). 'Time' is addressed in the Principles and Standards for School Mathematics (National Council of Teachers of Mathematics [NCTM], 2000) for both the pre-K–2 and Grade 3–5 band under the domain of Measurement. In other words, pupils learn about time concept from preschool until Grade 5.

In Malaysia, the topic of ‘Time’ is covered in both preschool and primary school mathematics syllabi. Preschool children are exposed to the concept of time in their early numeracy learning as stated in the Curriculum Standard and Assessment Document (*Dokumen Standard Kurikulum dan Penilaian* [DSKP]) of National Preschool Curriculum Standard (*Kurikulum Standard Prasekolah Kebangsaan* [KSPK]) (Ministry of Education Malaysia [MOE], 2017a). At five years old, the children will learn to sort the event according to the time stated and state the time in terms of morning, afternoon or night according to the event given (Harun, Ghazali, Hamid, & Nasir, 2017). When they reach six years old, they will learn about the time in a day such as 7 a.m., 11.00 a.m. and so on (Harun et al., 2017). Moreover, they will also learn to state the day (Sunday, Monday and so on) as well as state the month (January, February and so on) (Harun et al., 2017).

At the primary school level, the concept of time will be taught under learning area of Measurement and Geometry as outlined in DSKP of Primary School Standard Curriculum (*Kurikulum Standard Sekolah Rendah* [KSSR]) (MOE, 2016c, 2016d, 2016e, 2016f, 2016g, 2017b). Throughout six years of primary education, pupils are expected to equip with knowledge about time concept such as : (i) determine the appropriate unit of time measurement for different events; (ii) conversion between units of time; (iii) estimate the interval of time for certain events; (iv) use calendar and timetable in solving word problems; (v) calculate the duration between two given time in any units using the four operations; (vi) read both analogue and digital clock and write time in 12-hour system and 24-hour system; (vii) know the relationship and convert the time in 12-hour system to 24-hour system and vice versa; (viii) solve daily problems involving time, including time zone (MOE, 2016c, 2016d, 2016e, 2016f, 2016g, 2017b). The detailed Content Standards and Learning Standards for the topic of 'Time' in primary mathematics syllabi are attached in Appendix A.

Along with the implementation of Secondary School Standard Curriculum (*Kurikulum Standard Sekolah Menengah* [KSSM]) to replace the existing Integrated Curriculum of Secondary School (*Kurikulum Bersekutu Sekolah Menengah* [KBSM]), the concept of time is no longer being covered in secondary school Mathematics Syllabus (MOE, 2016b). In other words, pupils will not be taught about the concept of time after Year 6. Hence, it is crucial to help pupils to grasp the concept of time at the primary school level.

1.3 Problem Statement

'Time' is a basic mathematics concept that is very important in daily life. Although the time concept has been introduced to the children since preschool, they still struggled to grasp the concept in primary school (McGuire, 2007). Based on the findings of the Trend in International Mathematics and Science Study (TIMSS) 2015, the mean score of Grade Four participants in the 26 out of 49 participated countries fell below international average for the item involving the concept of 'Time' (Mullis, Martin, Foy, & Hooper, 2016). This indicates that the Grade Four participants in these countries have not mastered the concept of 'Time'.

Likewise, Van Steenbrugge, Valcke, and Desoete (2010) found that more than fifty per cent of the pupils in Grade 1 until Grade 5 had difficulties in learning the concept of 'Time'. Several studies (e.g. Burny, 2012; Burny et al., 2009; Earnest, 2015; Harris, 2008; Kamii & Russell, 2012) have been done to identify pupils' difficulties in learning the concept of time. For instance, Earnest (2015) found that pupils had difficulties in performing arithmetic operations involving ‘Time’ because they tend to treat time notation in terms of base ten. Under current base-60 time system (McGuire, 2007), each hour can be divided into 60 minutes and each minute can be subdivided into 60 seconds. The grouping of these hierarchical units is different from the base-10 system underlying place value and standard algorithms in elementary mathematics (Earnest, 2015). Thus, it could be mathematically challenging for the pupils.

In the Malaysian context, pupils’ cognitive strengths and weaknesses in the learning of clock time and calendar time have been identified by the researchers such as Sia, Lim, Chew, and Kor (2019) and Tan, Lim, and Kor (2017). According to Sia et al. (2019), knowing the relationship between the hour and minutes should be the most basic attribute in learning of clock time. Surprisingly, some of the pupils from National Primary School [Sekolah Kebangsaan, (SK)], Chinese National Type Primary School [*Sekolah Jenis Kebangsaan Cina*, (SJKC)], and Tamil National Type Primary School [*Sekolah Jenis Kebangsaan Tamil*, (SJKT)], who can perform addition or subtraction of the unit of time involving hour and minute might not master this prerequisite sub-skill (Sia et al., 2019). Similarly, Tan et al. (2017) revealed that 18.91% of Year 6 pupils in SK, SJKC and SJKT did not master any of the eight attributes about the concept of ‘after’ in calendar time. In short, the studies constantly indicated that the topic of ‘Time’ is also a problematic concept for Malaysian pupils from all school types.

Even though ‘Time’ has been reported as a problematic concept, there were an increase in the amount of content related to ‘Time’ in the Year Four mathematics curriculum under the Revised KSSR. Specifically, the skills covered in Year Four and Five mathematics curriculum under KSSR was included in a single topic in Year Four mathematics curriculum under the Revised KSSR (MOE, 2018a). Since the Revised KSSR mathematics curriculum would be used in SK, SJKC and SJKT, it might be more challenging for the pupils to learn such huge amount of content related to 'Time'. Thus, there is a crucial need to develop a cognitive diagnostic assessment which can diagnose the pupils' mastery of the attributes for the topic of 'Time', in order to support them in the learning of 'Time' concept.

In fact, CDA could provide more detailed and important diagnostic information to the education stakeholders compared to the other assessment. Yet, CDA is still in its infancy at the present moment (Leighton & Gierl, 2007b; Leighton, Gokiert, Cor, & Heffernan, 2010). It is still under research and not yet ready to be used for practical purposes (Leighton et al., 2010). Despite the strength of CDA in providing detailed information about pupils' cognitive strengths and weaknesses (Leighton & Gierl, 2007b), several drawbacks might discourage the practical use of the CDA as classroom assessment.

Firstly, the tedious operational procedures (Leighton & Gierl, 2007b) might hinder the utilisation of the CDA in the classroom. Since the CDA was administered manually, teachers have to mark pupils' responses, key in the responses in the binary pattern, perform psychometric analysis based on the cognitive diagnostic model, in order to compute pupils' mastery of the attributes. Nevertheless, the software used for the computation of pupils' mastery of attributes might be new to the teachers. The learning and mastering of new software might be time-consuming. Thus, the use of CDA is burdensome for the teachers.

Secondly, the granularity of the cognitive models of CDA does not fit the diagnostic purpose. According to Leighton and Gierl (2011), granularity or grain size can be defined as “both the depth and breadth of knowledge and skills being measured” (p. 62). For any of the diagnostic assessment, the attributes must be specified to fine grain size in order to make specific diagnostic inferences about examinees’ cognitive skills (Gierl & Leighton, 2007; Leighton & Gierl, 2011). The attributes are not fine-grained enough to make valid and precise diagnostic inferences about pupils’ cognitive strengths and weaknesses if the cognitive models of the CDA span the syllabi from Year 1 to Year 6. The course-grained cognitive models might limit the diagnostic power of the CDA.

Thirdly, the feedback was not given to the pupils after they had finished answering each item of the CDA. According to Hattie and Timperley (2007), feedback is given to the pupils so that the discrepancy between their current understanding and the desired understanding could be reduced. In other words, feedback given would consolidate pupils’ understanding and strengthen their skill mastery. However, Dihoff, Brosvic, Epstein, and Cook (2004) found that pupils demonstrated less recall, less confidence on their answer and more persistent to their original incorrect responses on the tested items when delayed feedback is provided. In other words, pupils tend to forget their mistakes if they are not given any feedback instantly. As a result, they will make the same mistakes again when they are answering similar questions.

Lastly, the scoring report was not generated after the pupils had finished answering the CDA. It is important to realise that score report is a critical component for any assessment programme (Gotch & Roberts, 2018) as it acts as the vehicle which transmits assessment information to the educational stakeholders for making inferences about examinees’ performance (Hattie, 2009). Although teachers could obtain data analysis report after performing the Artificial Neural Network (ANN) pattern recognition analysis, the report might be too technical and hardly understood by the teachers. Thus, teachers can barely make any diagnostic inference about pupils' knowledge states if the scoring report is not generated.

The advancement of technology promises the way to address the drawbacks of the CDA. Specifically, a coherent and comprehensive online assessment system named online Cognitive Diagnostic Assessment (CDA) is proposed to be developed in this study. The online CDA includes the CDA instrument, feedback and pupil's cognitive profiles which to date no research on online CDA with these components have been reported in the literature. By integrating the CDA with the online assessment system, the operational procedures of CDA can be simplified. Teachers can obtain pupils' diagnostic information immediately after their pupils have finished answering the items in online CDA, without going through the process of pupils' answer script marking and assessment data analysing. Thus, the online CDA might be more user-friendly to the teachers as it can support the crucial needs of the teachers in implementing the classroom assessment which is formative in nature.

Unlike the common online assessment system that tends to employ multiple-choice questions (Boitshwarelo, Reedy, & Billany, 2017; Lin & Dwyer, 2006; Nicol, 2007; Oldfield, Broadfoot, Sutherland, & Timmis, 2012; Westhuizen, 2016), this study opted to use ordered multiple-choice (OMC) items which are introduced by Briggs, Alonzo, Schwab, and Wilson (2006). Since each response option in the OMC items is connected to a specific level of understanding, the OMC-based instrument will provide "a more detailed picture of pupils' understanding compared to multiple-choice items" (Hadenfeldt, Bernholt, Liu, Neumann, & Parchmann, 2013, p. 1602) in a simple and efficient way. The ability of OMC in providing detailed insight into pupils' understanding of core concepts has coincided with the requirement of CDA. It is worth to note that, OMC items have been widely used for diagnosing pupils' learning progression in Science (e.g. Alonzo & Steedle, 2009; Briggs et al., 2006; Hadenfeldt et al., 2013; Hadenfeldt, Neumann, Bernholt, Liu, & Parchmann, 2016). However, there is a paucity of literature about mathematics OMC items. To fill this research gap, OMC items would be developed in this study.

Similar to other diagnostic assessment, the online CDA primarily focuses on assessing basic mathematics procedural skills. Nevertheless, Stevens (2009) argued that the problems that pupils encountered when solving problems involving HOTS should be identified, once they have mastered the basic procedural skills. Thus, problem-solving cognitive models would also be developed in this study. Particularly, the problem-solving cognitive models are incorporated with Newman Error Analysis (Clements, 1980; White, 2010). Newman's hierarchy which specifies the literacy and numerical skills needed in solving mathematical word problems (White, 2010) would be adapted to form the problem-solving cognitive models. As such, pupils' strengths and weaknesses in solving word problems for the topic of 'Time' can be diagnosed through the online CDA.

In order to fit the diagnostic purpose, the fine-grained cognitive models would be developed to guide the item development of the online CDA. Instead of spanning the topic of "Time" for the whole primary mathematics syllabus, the cognitive models of this study would only focus on the topic of 'Time' in Year 4 and Year 5 mathematics syllabi. With the small content coverage cognitive models, pupils' cognitive learning can be penetrated deeper (Gierl & Leighton, 2007). Thus, a more precise diagnostic inference about pupils' cognitive strengths and weaknesses can be obtained by using the online CDA which would be developed in this study.

The development of the fine-grained cognitive models began with the identification of attributes. Leighton and Gierl (2011) highlighted the need for attributes to align with the curriculum for ensuring the meaningful use of the diagnostic inference to guide remediation and instruction. Similarly, the accurate ordering of attribute hierarchy is also crucial because the attribute hierarchy will serve as the cognitive models which guide item development and test performance interpretation (Gierl, Alves, & Taylor-Majaeu, 2010). Thus, the validation of the attributes and the attribute hierarchies were conducted in this study.

According to Briggs et al. (2006), the construct map is the centrepiece of OMC items which enables pupils' cognitive development to be linked with the assessment. The error made in the derivation of the construct map will cause the inference made about pupils' cognitive development to be invalid. Besides that, the validity is a property of the prescribed interpretation and uses of the assessment score (Kane, 2015). Without the support of the evidence about content validity, the interpretations and uses of the scoring report generated in online CDA with OMC items cannot claim to be valid. As illustrated by Briggs et al. (2006), counting the frequency of the distractors chosen by the student for each level is the simplest way for determining pupils' cognitive level based on their responses. However, the diagnostic inference made about pupils' responses might be invalid if the level associated with the option of OMC item is incorrect. Thus, it is crucial for validating the construct map, the content of the CDA with OMC items and the level associated with each option of the OMC items to ensure the valid interpretation of the diagnostic inferences.

It is worth to note that, the psychometric properties of the CDA items often remain unknown or undocumented (Gierl et al., 2016), even though item analysis and distractor analysis serves as a quality-control step in the development of any assessment (Livingston, 2006). The psychometric properties which are evaluated in the item analysis including item difficulty and item discrimination. The results of the item analysis inform the decision in the item filtering process. The items which do not function well would be either revised or removed.

Similarly, the efficiency of the distractors is crucial to be evaluated to ensure the quality of the distractor is on par. This is because too many non-functioning distractors will decrease the discrimination power of the item and thereby decreasing the quality of the item at the same time. Thus, there is a crucial need to evaluate the psychometric properties of the items and the efficiency of distractors to ensure the online CDA with OMC items developed is of high quality (Livingston, 2006).

According to Downing (2004), reliability is another imperative quality assurance indicator of any assessment. The interpretation of inconsistent result yields in a low-reliability test might not be meaningful as it might consist of a high composition of random errors. Thus, the reliability of the online CDA with OMC items should be evaluated because it may affect the interpretations of the results and the academic decisions being made based on the results (Bardhoshi & Erford, 2017).

Rather than solely relying on the quality of the assessment itself, Keehner, Gorin, Feng, and Katz (2017) asserted that the validity claims about the pupils who take the CDA could be informed by the quality of the cognitive model. Even though Leighton, Cui, and Cor (2009) claimed that the cognitive models which are developed using expert analysis could be more credible, discrepancies between the expert-based cognitive models and pupil-based cognitive models could decrease the validity of cognitive models and cause the diagnostic inferences made based on the cognitive models to be less reliable. In other words, wrong diagnostic inferences could be made if the cognitive process for solving the task which is hypothesised by the experts as shown in expert-based cognitive models is not really matched with the pupil-based cognitive models which illustrated the pupils' actual cognitive process for solving the task.

To address this issue, cognitive models should be validated with empirical data (Keehner et al., 2017). Specifically, the hierarchical structure of the cognitive models can be confirmed by comparing the attribute probabilities pattern (Gierl et al., 2010) and analysing the correlation among the attributes (Keehner, et al., 2017). In addition, the overall model-data fit could provide evidence about the consistency of expert-based cognitive model and pupil-based cognitive model (Cui & Leighton, 2009). This evidence will eventually support the construct validity argument of the cognitive models (Keehner et al., 2017). With this, the diagnostic inferences made based on the cognitive models can be claimed as reliable.

With the advancement of technology, the feedback can be generated in a real-time manner in the online CDA developed in this study. According to Dann (2016), feedback is the integral component of teaching, learning and assessment. The feedback is given to the pupils after the assessment with the intention to help them to reduce the gap between what is known and what needs to be known (Dann, 2016). Following this, the attention should be given to the content of the feedback to ensure the meaningful use of the feedback (Dann, 2019). Thus, there is a need to validate the feedback content.

According to Van der Kleij, Feskens, and Eggen (2015), the feedback is considered effective if it could support learning by helping pupils to understand their performance for the task and giving actionable suggestions for the improvement. The effectiveness of the immediate feedback in mitigating pupils' misconceptions has been demonstrated in a study conducted by Chin and Lim (2018). Nonetheless, the currently available literature does not provide clear evidence on the effectiveness of the feedback generated in the online CDA on improving pupils' mastery of skill. Thus, the effectiveness of the feedback generated in online CDA would be investigated in this study.

In order to help pupils to reduce the discrepancy among the current learning goals and the desired learning goals, the feedback provided to the pupils should be of high quality. According to Small and Attree (2016), high-quality feedback which is clear, unambiguous, instructional and directive will be useful for the students for enhancing their learning. Nevertheless, Evans (2013) warned the extensive use of this type of feedback as it may limit pupils' thinking and even result in pupils' dependence. In other words, this type of feedback might lead to some unwanted consequence even though Small and Attree (2016) claimed that it is good feedback. In short, the factors which contribute to the effectiveness of feedback are inconsistent over the literature being reviewed (Evans, 2013). Thus, this study sought to determine the contributing factors of the effectiveness of the feedback generated in online CDA.

1.4 Research Objectives

This study aimed to develop and evaluate an online Cognitive Diagnostic Assessment (CDA) with feedback for assessing Year Four and Year Five pupils' achievement on the topic of 'Time'. The development of the online CDA with feedback was conducted in the first phase of the study while the evaluation of the online CDA with feedback was conducted in the second phase of the study.

Specifically, the objectives of the first phase of this study were:

1. To determine the validity of the expert-based cognitive models.

2. To determine the validity of the construct map.
3. To determine the content validity of the CDA with OMC items.
4. To determine the appropriateness of the level assigned to each option of OMC items for Year Four and Year Five topic of ‘Time’.
5. To evaluate the quality of the item for the Year Four and Year Five CDA with OMC items
6. To evaluate the quality of the distractors for the Year Four and Year Five CDA with OMC items
7. To evaluate the reliability of the Year Four and Year Five CDA with OMC items.
8. To determine whether the Year Four and Year Five ‘Time’ attributes are arranged in hierarchical structure based on pupils’ responses.
9. To evaluate the consistency between expert-based cognitive model and pupil-based cognitive model for Year Four and Year Five topic of ‘Time’.
10. To determine the content validity of the feedback of online CDA for Year Four and Year Five topic of ‘Time’.

Meanwhile, the objectives of the second phase of this study were:

11. To investigate the effectiveness of the feedback generated in the online CDA on the achievement of Year Four and Year Five pupils in the topic of ‘Time’.
12. To identify the factors that contribute to the effectiveness of the feedback generated in the Year Four and Year Five online CDA.

1.5 Research Questions

Based on the research objectives, the first phase of the study intended to answer the following research questions:

1. To what extent are the Year Four and Year Five expert-based cognitive models valid in terms of (i) attribute relevance and (ii) attribute hierarchy arrangement?
2. To what extent are the construct maps of Year Four and Year Five topic of ‘Time’ appropriate?
3. To what extent are the content of Year Four and Year Five CDA with OMC items valid in terms of (i) appropriateness of attribute classification in test specification table based on Revised Bloom’s Taxonomy, (ii) item relevance, and (iii) content coverage?
4. To what extent is the level appropriately assigned to each option of OMC items for Year Four and Year Five topic of ‘Time’?
5. To what extent do the item difficulty and item discrimination of Year Four and Year Five CDA with OMC items appropriate?
6. To what extent are the distractors for the Year Four and Year Five CDA with OMC items efficient?
7. To what extent are the Year Four and Year Five CDA with OMC items reliable?
8. To what extent are the Year Four and Year Five ‘Time’ attributes arranged in hierarchical structure based on pupils’ responses?
9. To what extent is the expert-based cognitive model consistent with the pupil-based cognitive model for Year Four and Year Five topic of ‘Time’?

10. To what extent is the content of the simple and detailed feedback of online CDA for Year Four and Year Five topic of ‘Time’ appropriate?

Based on the research objectives, the second phase of the study intended to answer the following research questions:

11(a) Is there any significant difference in the achievement in the topic of ‘Time’ between Year Four pupils who receive simple feedback from online CDA and the Year Four pupils who receive detailed feedback from online CDA?

11(b) Is there any significant difference in the achievement in the topic of ‘Time’ between Year Five pupils who receive simple feedback from online CDA and the Year Five pupils who receive detailed feedback from online CDA?

12(a) What are the factors that contribute to the effectiveness of the feedback generated in the Year Four online CDA?

12(b) What are the factors that contribute to the effectiveness of the feedback generated in the Year Five online CDA?

1.6 Null Hypotheses

The null hypotheses formulated in this study are as following.

H_o1: There is no significant difference in the achievement in the topic of ‘Time’ between Year Four SK pupils who received simple feedback from online CDA and the Year Four SK pupils who received detailed feedback from online CDA.

H_o2: There is no significant difference in the achievement in the topic of ‘Time’ between Year Four SJKC pupils who received simple feedback from online CDA and the Year Four SJKC pupils who received detailed feedback from online CDA.

H_o3: There is no significant difference in the achievement in the topic of ‘Time’ between Year Four SJKT pupils who received simple feedback from online CDA and the Year Four SJKT pupils who received detailed feedback from online CDA.

H_o4: There is no significant difference in the achievement in the topic of ‘Time’ between Year Five SK pupils who received simple feedback from online CDA and the Year Five SK pupils who received detailed feedback from online CDA.

H_o5: There is no significant difference in the achievement in the topic of ‘Time’ between Year Five SJKC pupils who received simple feedback from online CDA and the Year Five SJKC pupils who received detailed feedback from online CDA.

H_o6: There is no significant difference in the achievement in the topic of ‘Time’ between Year Five SJKT pupils who received simple feedback from online CDA and the Year Five SJKT pupils who received detailed feedback from online CDA.

1.7 Significance of the Study

This study would contribute to the enhancement in the aspect of assessment in mathematics education. The online CDA which would be developed in the study is a ready-made assessment instrument which can be used for diagnostic and remedial purposes in the classroom. Unlike assessments of achievement which tend to focus on only grade and scores, the online CDA emphasizes on pupils' mastery of skills in the topic of 'Time'. The levels of mastery of the skills being measured would be reported instead of total score or a single grade. Thus, the online CDA could serve as a formative assessment which will be given to the pupils after teaching each subtopic to determine the pupils' skill acquisition. In short, the online CDA could support the crucial needs of teachers in implementing classroom assessment (MOE, 2018b).

Apart from the aspect of assessment, this study would also contribute to the aspect of teaching and learning in mathematics education. Immediate corrective feedback was introduced as the additional feature on the online CDA that will be developed in the study. The immediate corrective feedback consisted of the information such as notification about the right or wrong of the answer, the correct answer and the full solution of the question. This feedback was given as remediation to the pupils immediately after they had answered each question in order to help them in reducing the discrepancy between their current understanding and their intended understanding. They would notice about the mistakes they might have conducted through reading the correct solution provided in the feedback and avoid conducting the same mistake in the subsequent item(s). Besides that, pupils would learn how to solve similar questions based on the correct solution provided, which serves as a directive guide.

In addition, the feature of the online CDA such as generating online scoring report would also enhance teaching and learning in mathematics education. The online scoring report would be generated once the pupils finished answering each assessment. Their mastery of the skills for the topic of 'Time' would be reported in a detailed manner in the online scoring report. With this imperative diagnostic information, the instructional decision made might be more accurate. Teachers could plan for remedial activities to help the pupils overcome their weaknesses in the topic of ‘Time’. Meanwhile, the pupils could identify their weaknesses and spot their own skill gaps on the topic of ‘Time’ based on the online scoring report generated. In other words, the online scoring reports would inform the pupils regarding the focus of their revision for the topic of ‘Time’. As such, they could adjust their learning strategies for self-improvement on this topic. In brief, the scoring report generated in the online CDA would benefit both teachers and pupils and would contribute to the enhancement in the aspect of teaching and learning in education.

On the whole, the online CDA could provide a springboard for both intervention and prevention, that helps to produce mathematical proficient pupils, which is in line with the aspiration of the Malaysian Education Blueprint 2013-2025 as well as the needs of the technological age.

1.8 Limitations of the Study

The study was limited in the following ways. Firstly, the findings of the student might be less generalisable due to the population, sampling technique and sample size. Due to the time constraint and the limited budget, the population was realistically chosen, and it was only limited to Year Four and Year Five pupils in Penang. Since instrument development is an iterative process and might be conducted for a few cycles until satisfactory item characteristics are obtained, the sample size of the first phase of study was limited by budget constraint. Only a small number of participants were selected by using convenient sampling. For the second phase of the study, a small number of participants were selected using purposive sampling because it was restricted by availability of school equipped with computer lab in which the computers were well functioning and connected with strong internet access, as well as the capacity of the computer lab. The use of non-probabilistic sampling and the small sample size would eventually reduce the generalisability of the study (Gay, Mills, & Airasian, 2012).

Besides that, the small sample size of this study has also limited the choice of psychometric model used in the study. Instead of Rasch model or IRT model, the psychometric properties of the items can only be evaluated by using Classical Test Theory (CTT) with a more lenient assumption. This eventually reduced the robustness of the findings.

Apart from population, sampling technique and sample size, the limitations of the study were also found on the aspect of research design, data collection and research instrument. Since multimethod research design was used, there were no triangulation of data in Phase 1 of the study. The findings from quantitative data such as model-data fit could not be explained because think aloud interviews were not conducted in the study. Moreover, the use of research instruments which were dichotomously scored could barely reveal pupils' thinking. Consequently, the findings of the study in Phase 1 was left unexplained.

In addition, this study would only focus on the topic of 'Time' in Year Four and Year Five Mathematics. Although cognitive diagnostic assessment has various strengths as outlined in the study carried out by Alves (2012), the procedure of the cognitive diagnostic assessment development is complicated. Thus, the cognitive diagnostic assessment developed would not cover the other topics in Year Four and Year Five Mathematics syllabus. Due to the lack of substantive theory related to the learning of ‘Time’, the item development was guided by the cognitive models which was derived from the curriculum specification. Thus, the assessment result could be hardly explained by the theory.

Furthermore, this study could only infer Year Four and Year Five pupils’ common mistakes in the topic of ‘Time’. Instead of constructed response items, ordered multiple-choice (OMC) items were used in this study. Thus, the respondents would only enable to choose the answer from the distractors which is designed by the researcher based on pupils’ common error found in the previous literature and the groundwork done by the researcher. Consequently, this study might be limited to provide an in-depth understanding of Year Four and Year Five pupils’ common mistakes on the topic of ‘Time’.

Lastly, the OMC items developed in this study might be susceptible to guesswork. Similar to the multiple-choice items, the ordered multiple-choice items also consist of a question stem and a correct option and a few incorrect options. With the limited knowledge, pupils might guess for the answer. Although Hierarchy Consistency Index (HCI) was calculated to determine the mismatch between the expert-based cognitive models and pupil-based cognitive models, the guessing factor cannot be explained fully using HCI. Nevertheless, the use of 3 items to measure each attribute in the online CDA based on the recommendation of Gierl et al. (2010) could reduce the chance of the guessing.

1.9 Definitions of Terms

'Time' – 'Time' is the concept covered under domain of 'Measurement' (NCTM, 2000). It includes the two main skills, namely (i) determining point of time (labelled by clocks, and calendars), and (ii) determining the durations (measuring elapsed time) (Harris, 2008). In this study, 'Time' is a topic which will be taught under the learning area of Measurement and Geometry as outlined in Year Four and Year Five DSKP of KSSR (MOE, 2016e, 2016f)

Achievement on 'Time' – Achievement on 'Time' refers to pupils' performance on the topic of 'Time'. In this study, Achievement on 'Time' is measured by using Achievement Test for the topic of 'Time'.

Attribute – Attribute refers to the specific procedure or skill that is required to solve test problems correctly (Leighton, Gierl, & Hunka, 2004). In this study, the attribute refers to the subskill related to the concept of 'Time' which was used to solve the mathematics task involving 'Time'.

Cognitive model – Cognitive models are illustrated as the hierarchical ordered cognitive attributes which are specified to certain grain size to guide the inference made about pupils' cognitive strengths and weaknesses (Leighton & Gierl, 2007a). In this study, cognitive models developed can be categorized into two types, namely expert-based cognitive model and pupil-based cognitive model. The expert-based cognitive model refers to the cognitive model that is generated by the content expert using task analysis, while the pupil-based cognitive model refer to the cognitive model that is generated based on pupils' responses in CDA (Sia et al., 2019)

Construct map – Construct map is illustrated as hierarchical level representation of unidimensional continuum of understanding about a given construct (Briggs et al., 2006). In this study, the construct map is hierarchical level representation of skill mastery related to ‘Time’ concept.

Feedback – Feedback is the corrective-nature information provided to pupils which attempts to reinforce and extend their understanding (Evans, 2013). In this study, both simple and detailed feedback were generated in Online CDA. The simple feedback consists of the notification of right or wrong of the answer and the correct answer. The detailed feedback consists of the notification of right or wrong of the answer, the correct answer, and the full solution.

Online Cognitive Diagnostic Assessment –. An online assessment system which consists of cognitive diagnostic assessment for Year Four and Year Five topic of ‘Time’. Immediate corrective feedback and pupil’s individual scoring report are generated in this online assessment system.

Ordered Multiple Choice items – Ordered Multiple Choice (OMC) items are the novel item format in which each of the options of the item is linked to the developmental levels of students’ understanding (Briggs et al., 2006). In this study, each option of OMC items was linked to the skill mastery level, rather than the developmental levels of students’ understanding.

Year Five – The fifth year of Malaysian primary education for the eleven-year old pupils.

Year Four – The fourth year of Malaysian primary education for ten-year old pupils.