Chulalongkorn University
Chula Digital Collections

Chulalongkorn University Theses and Dissertations (Chula ETD)

2020

Measurement of Reading Literacy, Growth, and Learning Potential of Grade 9 Students: Application of Computerized Dynamic Assessment Concept

Yanika Lunrasri
Faculty of Education

Follow this and additional works at: https://digital.car.chula.ac.th/chulaetd

Part of the Educational Assessment, Evaluation, and Research Commons

Recommended Citation

Lunrasri, Yanika, "Measurement of Reading Literacy, Growth, and Learning Potential of Grade 9 Students: Application of Computerized Dynamic Assessment Concept" (2020). *Chulalongkorn University Theses and Dissertations (Chula ETD)*. 116.
https://digital.car.chula.ac.th/chulaetd/116

This Thesis is brought to you for free and open access by Chula Digital Collections. It has been accepted for inclusion in Chulalongkorn University Theses and Dissertations (Chula ETD) by an authorized administrator of Chula Digital Collections. For more information, please contact ChulaDC@car.chula.ac.th.

Measurement of Reading Literacy, Growth, and Learning Potential of Grade 9 Students: Application of Computerized Dynamic Assessment Concept

Miss Yanika Lunrasri

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Educational Measurement and Evaluation
Department of Educational Research and Psychology
FACULTY OF EDUCATION
Chulalongkorn University
Academic Year 2020
Copyright of Chulalongkorn University

การวัดการรู้เรื่องการอ่าน พัฒนาการและศักยภาพการเรียนรู้ของนักเรียนชั้นมัธยมศึกษาปีที่ 3: การประยุกต์แนวคิดการประเมินแบบพลวัตด้วยคอมพิวเตอร์

น.ส.ญาณิกา ลุนราศรี

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรดุษฎีบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2563
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title	Measurement of Reading Literacy, Growth, and Learning Potential of Grade 9 Students: Application of Computerized Dynamic Assessment Concept
By	Miss Yanika Lunrasri
Field of Study	Educational Measurement and Evaluation
Thesis Advisor	Associate Professor KAMONWAN TANGDHANAKANOND, Ph.D.
Thesis Co Advisor	Associate Professor Shotiga Pasiphol, Ph.D.

Accepted by the FACULTY OF EDUCATION, Chulalongkorn University in Partial Fulfillment of the Requirement for the Doctor of Philosophy

..... Dean of the FACULTY OF EDUCATION

(Associate Professor SIRIDEJ SUJIVA, Ph.D.)

DISSERTATION COMMITTEE

..... Chairman

(Professor SIRICHAI KANJANAWASEE, Ph.D.)

..... Thesis Advisor

(Associate Professor KAMONWAN TANGDHANAKANOND, Ph.D.)

..... Thesis Co-Advisor

(Associate Professor Shotiga Pasiphol, Ph.D.)

..... Examiner

(Associate Professor SIRIDEJ SUJIVA, Ph.D.)

..... Examiner

(Associate Professor NUTTAPORN LAWTHONG, Ph.D.)

..... External Examiner

(Assistant Professor Sungworn Ngudgratoke, Ph.D.)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ญาณิกา ลุนราศรี : การวัดการรู้เรื่องการอ่าน พัฒนาการและศักยภาพการเรียนรู้ของนักเรียนชั้นมัธยมศึกษาปีที่ 3: การประยุกต์แนวคิดการประเมินแบบพลวัตด้วยคอมพิวเตอร์. (Measurement of Reading Literacy, Growth, and Learning Potential of Grade 9 Students: Application of Computerized Dynamic Assessment Concept) อ.ที่ปรึกษาหลัก : รศ. ดร.กมลวรรณ ตังธนากานนท์, อ.ที่ปรึกษาร่วม : รศ. ดร.โชติกา ภาษีผล

งานวิจัยเรื่องนี้มีวัตถุประสงค์เพื่อ (1) พัฒนาและตรวจสอบคุณภาพของเครื่องมือการประเมินแบบพลวัติด้วยคอมพิวเตอร์ในเรื่องการรู้เรื่องการอ่าน และ (2) ศึกษาผลของการทดสอบเครื่องมือการประเมินแบบพลวัติด้วยคอมพิวเตอร์ในเรื่องการรู้เรื่องการอ่านของกลุ่มที่ได้รับตัวช่วยคำชี้แนะที่แตกต่างกันที่มีต่อความสามารถด้านการรู้เรื่องการอ่าน พัฒนาการและศักยภาพการเรียนรู้ กลุ่มตัวอย่างที่ใช้ในการตรวจสอบเครื่องมือ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 802 คน จาก 6 โรงเรียนในเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 ตัวอย่างในการวิจัย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 541 คน จาก 11 โรงเรียนในเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 และเขต 2 งานวิจัยนี้เป็นงานวิจัยแบบกึ่งทดลอง เครื่องมือที่ใช้ในการวิจัย คือ เครื่องมือการประเมินแบบพลวัติด้วยคอมพิวเตอร์ในเรื่องการรู้เรื่องการอ่าน สถิติที่ใช้วิเคราะห์ข้อมูล ได้แก่ 1) การวิเคราะห์ความแปรปรวนร่วมสำหรับการวิเคราะห์ความสามารถในการรู้เรื่องการอ่าน 2) การวิเคราะห์โมเดลโค้งพัฒนาการที่มีตัวแปรแฝงสำหรับการวิเคราะห์พัฒนาการ และ 3) การวิเคราะห์ความแปรปรวนแบบผสมและการวิเคราะห์ความแปรปรวนพหุคูณสำหรับการวิเคราะห์ศักยภาพการเรียนรู้ โดยผลการวิจัยพบว่า

1. งานวิจัยนี้ได้สร้างเครื่องมือการประเมินแบบพลวัติด้วยคอมพิวเตอร์ในเรื่องการรู้เรื่องการอ่านเป็นเครื่องมือที่มีการให้ข้อมูลย้อนกลับโต้ตอบแบบออนไลน์ เนื้อหาของการรู้เรื่องการอ่านประกอบด้วย 1. การระบุสารสนเทศในเนื้อเรื่อง 2. การมีความเข้าใจในเนื้อเรื่อง 3. การประเมินและสะท้อนความคิดเห็น โดยเครื่องมือเป็นแบบสอบคู่ขนานแบบเลือกตอบ 2 ระดับ จำนวน 20 ข้อ มีทั้งหมด 3 ชุดเพื่อทดสอบจำนวน 3 ครั้ง ผลการตรวจสอบคุณสมบัติทางจิตมิติของเครื่องมือพบว่า ข้อคำถามส่วนใหญ่ในแบบสอบทั้งสามฉบับมีความตรงเชิงเนื้อหาและความเป็นคู่ขนาน ผลการเปรียบเทียบโมเดลเพื่อหาความสอดคล้องของโมเดลพบว่า โมเดลสององค์ประกอบมีความสอดคล้องกับข้อมูลมากที่สุด ข้อคำถามส่วนใหญ่มีค่าอำนาจจำแนกแบบพหุมิติอยู่ในระดับดีและมีค่าความยากแบบพหุมิติในระดับที่เหมาะสม เครื่องมือมีความเที่ยงแบบสอดคล้องภายในสูง เมื่อพิจารณาความเป็นคู่ขนานพบว่าแบบสอบมีความเป็นคู่ขนานทั้งในระดับแบบสอบและระดับข้อสอบ

2. เมื่อพิจารณาความสามารถในด้านการรู้เรื่องการอ่าน รูปแบบการให้ตัวช่วยคำชี้แนะที่แตกต่างกันมีผลต่อคะแนนการรู้เรื่องการอ่านหลังเรียนอย่างมีนัยสำคัญทางสถิติ กลุ่มตัวช่วยคำชี้แนะแบบยืนยันมีคะแนนการรู้เรื่องการอ่านหลังเรียนต่ำกว่ากลุ่มตัวช่วยคำชี้แนะแบบอื่น ๆ อย่างมีนัยสำคัญทางสถิติ เมื่อพิจารณาในด้านพัฒนาการการรู้เรื่องการอ่าน กลุ่มตัวช่วยคำชี้แนะแบบผสมมีพัฒนาการด้านการรู้เรื่องการอ่านมากกว่ากลุ่มอื่น ๆ เมื่อพิจารณาความสัมพันธ์ระหว่างการรู้เรื่องการอ่านในแต่ละด้าน พบว่า คะแนนพัฒนาการในแต่ละด้านไม่ได้มีความสัมพันธ์กับคะแนนพัฒนาการในด้านอื่น ๆ ยกเว้นกลุ่มตัวช่วยคำชี้แนะแบบผสมที่มีคะแนนพัฒนาการในด้านการมีความเข้าใจในเนื้อเรื่องมีความสัมพันธ์กับคะแนนพัฒนาการในด้านการประเมินและสะท้อนความคิดเห็น เมื่อพิจารณาด้านศักยภาพในการเรียนรู้ในด้านการรู้เรื่องการอ่าน พบว่าในด้านคะแนนช่องว่างพัฒนาการ ผลการวิจัยพบว่ากลุ่มที่ได้รับรูปแบบการให้ตัวช่วยคำชี้แนะที่แตกต่างกันมีคะแนนช่องว่างพัฒนาการในการวัดครั้งที่สามแตกต่างกันอย่างมีนัยสำคัญ กลุ่มตัวช่วยคำชี้แนะแบบยืนยันมีคะแนนช่องว่างทางพัฒนาการมากกว่ากลุ่มตัวช่วยคำชี้แนะแบบผสมอย่างมีนัยสำคัญทางสถิติ ในด้านคะแนนส่งผ่าน ผลการวิจัยพบว่ารูปแบบการให้ตัวช่วยคำชี้แนะที่แตกต่างและครั้งในการทดสอบไม่ได้มีปฏิสัมพันธ์กันต่อคะแนนส่งผ่านอย่างมีนัยสำคัญทางสถิติ อย่างไรก็ตามรูปแบบการให้ตัวช่วยคำชี้แนะที่แตกต่างกันมีผลต่อคะแนนส่งผ่านอย่างมีนัยสำคัญทางสถิติ โดยที่กลุ่มตัวช่วยคำชี้แนะแบบยืนยันมีคะแนนส่งผ่านน้อยกว่ากลุ่มอื่น ๆ อย่างมีนัยสำคัญทางสถิติ ในด้านระดับคำชี้แนะ ผลการวิจัยพบว่า กลุ่มที่ได้รับรูปแบบการให้ตัวช่วยคำชี้แนะที่แตกต่างกันมีการใช้คำชี้แนะในระดับที่หนึ่ง คำชี้แนะในระดับที่สอง คำชี้แนะในระดับที่สาม และคำชี้แนะในระดับที่สี่ แตกต่างกันอย่างมีนัยสำคัญทางสถิติ กลุ่มตัวช่วยคำชี้แนะแบบยืนยันมีการใช้ระดับคำชี้แนะในทุกระดับในด้านการรู้เรื่องการอ่านมากกว่ากลุ่มตัวช่วยคำชี้แนะแบบผสม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สาขาวิชา	การวัดและประเมินผลการศึกษา	ลายมือชื่อนิสิต .....
ปีการศึกษา	2563	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....
		ลายมือชื่อ อ.ที่ปรึกษาร่วม .....

5984230027 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION
KEYWORD: READING LITERACY, COMPUTERIZED DYNAMIC ASSESSMENT, GROWTH, LEARNING POTENTIAL, PROMPTING

Yanika Lunrasri : Measurement of Reading Literacy, Growth, and Learning Potential of Grade 9 Students: Application of Computerized Dynamic Assessment Concept . Advisor: Assoc. Prof. KAMONWAN TANGDHANAKANOND, Ph.D. Co-advisor: Assoc. Prof. Shotiga Pasiphol, Ph.D.

The aims of this research were to (1) develop and validate the computerized dynamic assessment for reading literacy and (2) study the effects of different types of prompting of the computerized dynamic assessment on reading literacy, growth of learning, and learning potential. The pilot samples were a total of 802 ninth-grade students in six secondary schools in Bangkok. The data for the main study was gathered from 541 ninth-grade students in eleven secondary schools in Bangkok. A quasi-experimental design was adopted. Research instruments were reading literacy tests for computerized dynamic assessment. The data were analyzed as follows: 1) ANCOVA statistics for reading literacy performance, 2) latent growth modeling for growth of learning, and 3) mixed ANOVA and MANOVA statistics for learning potential. Findings were presented as follows:

1. This research developed computerized dynamic assessment for reading literacy, which was interactive and online feedback program. The tested contents were composed of three dimensions of reading literacy, including 1) locate information, 2) understand, and 3) evaluate and reflect. Three parallel tests were assessed at three different time points, each test comprised twenty two-tier items. For the psychometric properties of the instrument, most items for all reading literacy test forms were appropriate for content validity and parallelism. For the model comparison, the bifactor MIRT model was the best-fitting model. A majority of the items had good multidimensional discrimination values. The multidimensional difficulty estimates were in the acceptable range for most items. Moreover, the instruments yielded highly internal reliability. Regarding the statistical parallelism of the test forms, they showed satisfactory conformity at the test forms and item-by-item levels.

2. For the results of reading literacy performance, different types of promptings of computerized dynamic assessment had a significant effect on students' reading posttest scores. Verification prompting group received significantly lower posttest score than other prompting-based groups. For the growth of learning, mixed prompting group obtained the highest rate of growth in reading literacy than other groups. For the results of the associations of reading literacy subscales, the growth rate in one subscale did not relate to the growth rate in other subscales, except mixed prompting group that the growth rate in understand was associated with the growth rate in evaluate and reflect. In terms of learning potential, the results of the availability score revealed that there were significant differences in availability scores measured at the third testing session among groups with different types of promptings. Verification prompting group had a significantly higher availability score when compared with mixed prompting group. For the mediated score, the result showed that there were no significant two-way interactions between prompting conditions and time on the mediated score. However, there was a significant main effect of prompting conditions. Verification prompting group had a significantly lower mediated score when compared with other groups. For the levels of prompting, the result showed that there were significant differences in first level of prompting, second level of prompting, third level of prompting, and fourth level of prompting among groups with different types of promptings. Verification prompting obtained significantly higher assistance than mixed prompting group in all levels of prompting.

Field of Study:	Educational Measurement and Evaluation	Student's Signature .....
Academic Year:	2020	Advisor's Signature .....
		Co-advisor's Signature .....

ACKNOWLEDGEMENTS

I would like to express my thanks to everyone who supported me throughout the course of my study. First of all, I would like to express my deepest gratitude to my advisor, Assoc. Prof. Dr. Kamonwan Tangdhanakanond, who dedicated her time to assist my work, kindly gave me guidance and valuable knowledge, as well as provided excellent motivation, encouragement, and passion for doing research. In addition, I would like to express my greatest appreciation to my co-advisor, Assoc. Prof. Dr. Shotiga Pasiphol for her valuable advice, useful comments and suggestions.

I would like to enlarge words of thanks to my thesis committee, including Prof. Dr. Sirichai Kanjanawasee, Assoc. Prof. Dr. Siridej Sujiva, Assoc. Prof. Dr. Nuttaporn Lawthong, and Asst. Prof. Dr. Sungworn Ngudgratoke for their valuable suggestions and recommendations for my research study. Their insightful comments were very useful for me to accomplish my success.

I would like to thank all experts who assisted me in the process of validating the research instruments as well as giving suggestions on research development. I also would like to express my warm thanks to all schools and all participants for their willingness to take part in this study.

I would like to express my sincere gratitude and appreciation to my family and friends for their love, moral support, and encouragement during my study. Moreover, I would like to acknowledge my appreciation to my friends for their assistance and encouragement.

Finally, I am deeply grateful to Graduate School, Chulalongkorn University. This work would not have been possible without the financial support of the 100th Anniversary Chulalongkorn University Fund of Doctoral Scholarship and the 90th Anniversary Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund) provided for my study.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Yanika Lunrasri

TABLE OF CONTENTS

Page

ABSTRACT (THAI) iii

ABSTRACT (ENGLISH) iv

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF TABLES xi

LIST OF FIGURES xvi

CHAPTER ONE INTRODUCTION 1

Background of the Study 1

Research Questions 8

Research Objectives 9

Research Hypotheses 9

Scopes of the Study 11

Definition of Terms 12

Significance of the Study 17

CHAPTER TWO LITERATURE REVIEW 18

Part 1: Dynamic Assessment 18

1.1 Concept of dynamic assessment 18

1.2 Approaches of dynamic assessment 23

1.3 Graduated prompting approach 25

1.4 Mediation of dynamic assessment 28

1.5 Transfer 29

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Part 2: Computerized Dynamic Assessment (C-DA) 33
2.1 Procedures of computerized dynamic assessment 33
2.2 Mediation of computerized dynamic assessment 35
2.3 Scoring procedures 45
2.4 The application of the two-tier test 49
2.5 Relevant studies on computerized dynamic assessment (C-DA) 52
Part 3: Reading Literacy 56
3.1 Definition of reading literacy 56
3.2 Dimensions of reading literacy 57
3.3 Reading literacy framework 64
3.4 Reading passages 67
3.5 Levels of reading literacy proficiency 68
3.6 Related studies on reading literacy 72
Part 4: Measurement of Growth 75
4.1 The Traditional method of measuring change 75
4.2 Modern method of measuring change 76
Part 5: Measurement of Learning Potential 96
5.1 Concept of learning potential 96
5.2 Measurement of learning potential 98
Part 6: Test Validation Using Item Response Theory 106
6.1 Concept of item response theory 106
6.2 Assumptions of IRT 106
6.3 Testlet effect 107
6.4 Item parameter analysis 109

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

6.5 Item and test information functions 110
Part 7: Equivalence Analysis 112
Part 8: Research Framework 114
CHAPTER THREE RESEARCH METHODOLOGY 119
1. Participants 119
1.1 Groups of participants 119
1.2 Sample size 121
1.3 Ethical issues 122
2. Research Procedure 123
2.1 Phase 1: Construction of the instrument 123
2.2 Phase 2: Validation of the research instrument 140
2.3 Phase 3: Implementation of the computerized dynamic assessment for reading literacy 145
3. Data Collection 149
3.1 Before the experiment 149
3.2 During the experiment 149
3.3 After the experiment 149
4. Data Analysis 150
4.1 Data analysis for instrument validation 150
4.2 Data analysis for reading literacy, growth, and learning potential 152
CHAPTER FOUR RESULTS 157
Part 1: Results of the Development and Validation of the Computerized Dynamic Assessment for Reading Literacy of Grade 9 Students 158
1.1. Results of the development of the computerized dynamic assessment for reading literacy of Grade 9 students 158

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2 Results of the validation of the computerized dynamic assessment for reading literacy of Grade 9 students 184

Part 2: Results of the Effects of Different Promptings of the Computerized Dynamic Assessment on Reading Literacy Ability, Growth of Learning, and Learning Potential of Grade 9 Students 211

2.1 Results of the effects of different promptings of the computerized dynamic assessment on reading literacy ability of Grade 9 students 211

2.2 Results of the effects of different promptings of the computerized dynamic assessment on growth of learning of Grade 9 students 225

2.3 Results of the effects of different promptings of the computerized dynamic assessment on learning potential of Grade 9 students 245

CHAPTER FIVE DISCUSSION AND CONCLUSION 271

Summary of the Study 271

Discussion 281

Conclusion 302

Implication and Recommendation of the Study 303

REFERENCES 308

APPENDICES 320

Appendix A The List of Experts 321

Appendix B The Results of IOC Index of the Reading Literacy Pretest and Posttest 325

Appendix C The Results of IOC Index of the Reading Literacy Test Sets for Computerized Dynamic Assessment 330

Appendix D The Results of IOC Index of the Prompting Methods 343

Appendix E The Results of IOC Index of the Parallelism of Reading Literacy Test Sets for Computerized Dynamic Assessment 355

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Appendix F Example of Reading Literacy for Computerized Dynamic Assessment 359

VITA 379

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

LIST OF TABLES

Page

Table 1 Summary of types of approaches 24

Table 2 Comparison between DA task and TA task 31

Table 3 Mediation of the inferential reading task 39

Table 4 Summary of instructional prompting from related studies 41

Table 5 Error-explanation prompts 44

Table 6 Summary of previous studies on computerized dynamic assessment 55

Table 7 The PISA 2009-2015 and PISA 2018 reading framework processes 58

Table 8 PIRLS 2016 reading literacy framework 60

Table 9 NAEP reading framework 61

Table 10 Reading literacy dimensions among international programs and Thai research 63

Table 11 Similarity and discrepancy on reading literacy between PISA2018 and the others 63

Table 12 Levels of reading literacy and descriptions of reading proficiency 69

Table 13 Summary of reading literacy research in Thailand 74

Table 14 Summary of the methods for measuring change 95

Table 15 Characteristics of the samples 120

Table 16 The demographics of the participant 121

Table 17 Aspects of the reading literacy test 124

Table 18 Table of specification of reading literacy pretest and posttest 125

Table 19 Table of specification of reading literacy instruments 126

Table 20 The examples of tasks 127

Table 21 Three dimensions of text taxonomies 128

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 22 Selection of passages based on task, text, length, and transferability 130

Table 23 Test blueprint of instruments for computerized dynamic assessment for reading literacy 131

Table 24 Scoring procedure example 133

Table 25 Table of specification of instructional prompts 135

Table 26 Table of specification of error-explanation prompts 136

Table 27 Table of specification of mixed prompts 137

Table 28 The experimental design of this study 147

Table 29 Labels for item discrimination and item difficulty estimates 151

Table 30 The overall model fit of reading literacy Form C-DA1 188

Table 31 The overall model fit of reading literacy Form C-DA2 189

Table 32 The overall model fit of reading literacy Form C-DA3 190

Table 33 Item parameter estimates for reading literacy Form C-DA1 191

Table 34 Item parameter estimates for reading literacy Form C-DA2 193

Table 35 Item parameter estimates for reading literacy Form C-DA3 195

Table 36 Confirmatory factor analysis of reading literacy Form C-DA1 (n = 359) 196

Table 37 The CFA estimates, standard errors, and p-values for the second-order factor analysis model of reading literacy Form C-DA1 (n = 359) 197

Table 38 Confirmatory factor analysis of reading literacy Form C-DA2 (n = 336) 199

Table 39 The CFA estimates, standard errors, and p-values for the second-order factor analysis model of reading literacy Form C-DA2 (n = 336) 199

Table 40 Confirmatory factor analysis of reading literacy Form C-DA3 (n = 344) 202

Table 41 The CFA estimates, standard errors, and p-values for the second-order factor analysis model of reading literacy Form C-DA3 (n = 344) 202

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 42 Reliability of the instruments for computerized dynamic assessment for reading literacy 204

Table 43 Parallel forms of reading literacy of Test Set 1(Form C-DA1 and Form C-DA2) 206

Table 44 Parallel forms of reading literacy of Test set 2 (Form C-DA1 and Form C-DA3) 208

Table 45 Parallel forms of reading literacy of Test set 3 210

Table 46 Descriptive statistics of students' pretest and posttest scores on reading literacy 211

Table 47 Analysis of Covariance for student's reading literacy posttest score with pretest score as covariate 212

Table 48 Results of pairwise comparisons 213

Table 49 Descriptive statistics of students' pretest and posttest scores on reading literacy subscales 214

Table 50 Analysis of Covariance for student's locate information subscale score with pretest score as covariate 215

Table 51 Results of pairwise comparisons of locate information subscale 216

Table 52 Analysis of Covariance for student's understand subscale score with pretest score as covariate 216

Table 53 Results of pairwise comparisons of understand subscale 217

Table 54 Analysis of Covariance for student's evaluate and reflect subscale score with pretest score as covariate 218

Table 55 Results of pairwise comparisons of evaluate and reflect subscale 218

Table 56 Descriptive statistics of group and gender on reading literacy pretest and posttest scores 219

Table 57 A Two-way ANCOVA of prompting groups and gender on student's reading literacy posttest score with pretest score as covariate 220

Table 58 Results of pairwise comparisons of group difference 221

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 59 Descriptive statistics of group and learning achievement on reading literacy posttest score 222

Table 60 Two-way ANCOVA student's reading literacy posttest score with pretest score as covariate 223

Table 61 Results of pairwise comparisons of group difference 224

Table 62 Descriptive statistics for growth of learning 226

Table 63 Fit indices of univariate growth curves for variables 227

Table 64 Univariate model of growth in reading literacy and reading literacy subscales 230

Table 65 Fit indices of multivariate growth curves for variables 234

Table 66 Coefficient estimates and standard error of the multivariate latent growth model for reading literacy variables of all students (N =541) 235

Table 67 Coefficient estimates and standard error of the multivariate latent growth model for reading literacy subscale variables 239

Table 68 Means and standard deviations of reading literacy scores across groups with differently initial reading literacy over three time points 241

Table 69 Means and standard deviations of capacity and ability scores over three time points 246

Table 70 Means and standard deviations of availability scores over three time points 247

Table 71 One-way MANOVA result of availability scores of different prompting groups 248

Table 72 Analysis of variance result of availability scores of different prompting groups 249

Table 73 Results of pairwise comparisons of group difference 249

Table 74 Descriptive statistics of mediated scores of computerized dynamic assessment for reading literacy 250

Table 75 Two-way mixed-effects ANOVA of mediated scores 251

Table 76 Main effect result of pairwise comparison of random-effect (group) 252

Table 77 Main effect result of pairwise comparison of fixed-effect (time) 253

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 78 Means of levels of promptings across three time points 255

Table 79 One-way MANOVA result of levels of prompting of reading literacy 257

Table 80 Analysis of variance result of levels of prompting of reading literacy 257

Table 81 Multiple comparisons in levels of prompting of reading literacy 258

Table 82 Means of levels of promptings in locate information subscale across three time points 259

Table 83 One-way MANOVA result of levels of promptings in locate information subscale 260

Table 84 Analysis of Variance result of levels of promptings in locate information subscale ... 260

Table 85 Multiple Comparisons of levels of promptings in locate information subscale 261

Table 86 Means of levels of promptings in understand subscale across three time points .. 263

Table 87 One-way MANOVA result of levels of promptings in understand subscale 264

Table 88 Analysis of Variance result of levels of promptings in understand subscale..... 264

Table 89 Multiple Comparisons of levels of promptings in understand subscale..... 265

Table 90 Means of levels of promptings in evaluate and reflect subscale across three time points 267

Table 91 One-way MANOVA result of levels of promptings in evaluate and reflect subscale..... 268

Table 92 Analysis of Variance result of levels of promptings in evaluate and reflect subscale..... 268

Table 93 Multiple Comparisons of levels of promptings in evaluate and reflect subscale ... 269

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

LIST OF FIGURES

Page

Figure 1 The AnimaLogica training item 27

Figure 2 Graduated prompting flowchart 27

Figure 3 Listening transferred item 32

Figure 4 Construction of computerized dynamic assessment 34

Figure 5 Prompted items of a French reading comprehension test 36

Figure 6 Prompted item 38

Figure 7 Abstract, semi-specific, and specific prompts 40

Figure 8 Prompting interface 43

Figure 9 Scores of the Chinese listening test 46

Figure 10 Learning potential score formula 47

Figure 11 Learning profile of the C-DA Chinese listening test 48

Figure 12 Group profile of the C-DA Chinese listening test 49

Figure 13 Two-tier testing process adapted in this study 50

Figure 14 Second-tier options 51

Figure 15 Dimensions of reading literacy 65

Figure 16 Relationship between items and students on a proficiency scale 69

Figure 17 Path diagram for no change latent growth curve 80

Figure 18 Path diagram of linear latent growth curve 81

Figure 19 Path diagram of quadratic latent growth curve 83

Figure 20 Path diagram of unspecified latent growth curve 85

Figure 21 Path diagram of associative latent growth model 89

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Figure 22 Path diagram of the factor-of-curves latent growth model..... 90

Figure 23 Path diagram of the curve-of-factors latent growth model..... 91

Figure 24 Latent growth curve model with gender differences 92

Figure 25 Linear second-order latent growth curve model for four measurement occasions..... 94

Figure 26 Three components of potential 97

Figure 27 Michaelis-Menten function 102

Figure 28 Number of prompts per training session 105

Figure 29 Conceptual framework of the present study 118

Figure 30 Research procedure of the present study..... 123

Figure 31 Classes in Google classroom for different prompting groups..... 145

Figure 32 Reading literacy pretest in Google Form 146

Figure 33 Reading literacy posttest in Google Form 148

Figure 34 Flowchart of data analysis of the study 156

Figure 35 Structure of computerized dynamic assessment 159

Figure 36 User interface 161

Figure 37 Example of reading passage..... 162

Figure 38 Example of reading literacy item (first-tier item)..... 162

Figure 39 The prompt given when student answered first-tier item correctly..... 163

Figure 40 Example of reading literacy item (second-tier item)..... 163

Figure 41 Instructional prompt of computerized dynamic assessment..... 165

Figure 42 Error-explanation prompt of computerized dynamic assessment..... 167

Figure 43 Mixed prompt of computerized dynamic assessment 169

Figure 44 Scoring report of computerized dynamic assessment..... 171

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Figure 45 Example of reading literacy item in locate information subscale (Item 1)..... 171
Figure 46 Example of reading literacy item in understand subscale (Item 2) 172
Figure 47 Example of reading literacy item in evaluate and reflect subscale (Item 4)..... 172
Figure 48 Instructional prompting method in locate information (Item 1) 173
Figure 49 Instructional prompting method in understand subscale (Item 2) 173
Figure 50 Instructional prompting method in evaluate and reflect subscale (Item 4)..... 174
Figure 51 Error-explanation prompting method in locate information subscale (Item 1).... 175
Figure 52 Error-explanation prompting method in understand subscale (Item 2)..... 176
Figure 53 Error-explanation prompting method in evaluate and reflect subscale (Item 4).... 177
Figure 54 Mixed prompting method in locate information subscale (Item 1)..... 179
Figure 55 Mixed prompting method in understand subscale (Item 2)..... 181
Figure 56 Mixed prompting method in evaluate and reflect subscale (Item 4)..... 183
Figure 57 Second-order factor analysis model of reading literacy Form C-DA1 198
Figure 58 Second-order factor analysis model of reading literacy Form C-DA2 201
Figure 59 Second-order factor analysis model of reading literacy Form C-DA3 203
Figure 60 Test information and test characteristic curve of Test set 1 206
Figure 61 Test information and test characteristic curve of Test set 2..... 208
Figure 62 Test information between reading literacy of Test set 3 210
Figure 63 Students' reading literacy pretest and posttest scores..... 211
Figure 64 Line plot of the adjusted posttest score..... 212
Figure 65 Students' posttest subscale scores 214
Figure 66 Line plot of the adjusted posttest score for locate information subscale..... 215
Figure 67 Line plot of the adjusted posttest score for understand subscale 217
Figure 68 Line plot of the adjusted posttest score for evaluate and reflect subscale..... 218

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Figure 69 Students' posttest scores with respect to different prompting groups and gender 220

Figure 70 Line plot of prompting groups and gender on student's reading literacy posttest score 221

Figure 71 Students' posttest scores with respect to different prompting groups and levels of learning achievement 223

Figure 72 Line plot of prompting groups and levels of learning achievement on student's reading literacy posttest score 224

Figure 73 The trajectory plot of reading literacy for all students 226

Figure 74 The trajectory plots of reading literacy subscales for different prompting groups 231

Figure 75 The trajectory plots of overall reading literacy for different prompting groups 232

Figure 76 The univariate latent growth curves of overall reading literacy for different prompting groups 233

Figure 77 The multivariate latent growth curves of overall reading literacy 236

Figure 78 Examples of Individual growth trajectories over three time points 240

Figure 79 Growth trajectories of reading literacy of low initial reading literacy score 242

Figure 80 Growth trajectories of reading literacy of medium-low initial reading literacy score 243

Figure 81 Growth trajectories of reading literacy of medium-hiigh initial reading literacy score 244

Figure 82 Growth trajectories of reading literacy of high initial reading literacy score ... 244

Figure 83 The Michaelis-Mentens curve of reading literacy across three time points 245

Figure 84 Trajectory plots of reading literacy ability and estimated capacity over three time points 247

Figure 85 Line graphs of availability scores over three time points 249

Figure 86 Box plots of mediated scores of different prompting groups over three time points 250

Figure 87 Line graphs of mediated scores of different prompting groups over three time points 251

Figure 88 Box plots of group effect 252

Figure 89 Box plots of time effect 253

Figure 90 Line graphs of levels of prompting of different groups over three time points 256

Figure 91 Line graphs of levels of promptings in locate information subscale 262

Figure 92 Line graphs of levels of prompting in understand subscale 266

Figure 93 Line graphs of levels of prompting in evaluate and reflect subscale 270

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER ONE
INTRODUCTION

Background of the Study

Literacy is one of the universal education goals believed to be the foundation to create sustainable development. One of the goals of the United Nations' Agenda of Sustainable Development Goals (SDGs) is the quality education to ensure that youth and adults equally achieve literacy and numeracy by 2030 (United Nations, 2018). Moreover, the 20-year National Education Plan of Thailand (2017-2036) has considered literacy as one of the strategies to improve Thai educational system; it states, "people of all ages have essential knowledge and skills as indicated in the educational and professional standards, and can use their potential to improve their lives" (Office of the Education Council, 2017). The mentioned indicators include the improvement of reading proficiency level and the rise of national and international assessment results, the Programme for International Student Assessment (PISA) included (Office of the Education Council, 2017) .

The Programme for International Student Assessment (PISA) is an international large-scale assessment study administered by the Organization for Economic Co-operation and Development (OECD) (OECD, 2016). Thailand has participated in the study since 2000 as the partner country. The survey has administered three domains, including math, science, and reading literacy, for 15-year-old students to measure their knowledge and skills needed to be a successful member in the society (Dechsri, 2016). PISA is the measurement of future human capital, an indicator of the success and competitiveness in economic development (The Institute for the Promotion of Teaching Science and Technology, 2015). Moreover, PISA aims to improve the quality of education by providing a deeper perspective of the strengths and weaknesses in the education system whether the school system readily prepares students for the global economy in the twenty-first century (The Institute for the Promotion of Teaching Science and Technology, 2015). Reading literacy is the foundation of other subject areas in the educational system; it is essential skills needed for individual growth, educational and career success, and global citizenship (OECD, 2016). The attainment in reading literacy successfully leads to the prerequisite for participation in real-life situations in

the rapidly changing world. Moreover, if the population lacks basic skills, the impact will happen on a country's long-term economic growth (OECD, 2016).

The most recent evidence from PISA 2015 and PISA 2018 showed that there was no improvement of Thai students' reading literacy results. The average reading literacy scores of Thai students decreased steadily since PISA 2000 (The Institute for the Promotion of Teaching Science and Technology, 2018). In recent PISA 2018, Thai students' average scores in reading literacy were 393 points below the OECD average (487 points) (OECD, 2019a; The Institute for the Promotion of Teaching Science and Technology, 2020). Half of students (50%) had reading literacy scaled scores below Level 2, which is the recommended minimum level benchmark (The Institute for the Promotion of Teaching Science and Technology, 2018). Those students who scored below Level 2 on the PISA reading scale are defined as low performers in PISA reading literacy (OECD, 2016). Level 2 is considered as the baseline level of proficiency for those who are required to fully participate in the society. Students below this baseline are those who cannot engage in solving complex reasoning and problems that are faced by adults in modern societies (OECD, 2016). Thus, it has been shown that the number of low-performing students in reading literacy are increasing in Thai education. This is coherent with UNESCO's 2017/8 Global Education Monitoring (GEM) Report who found that only 50% of students have achieved a minimum proficiency level in reading at the end of lower secondary education (UNESCO Bangkok, 2016).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Reading literacy is the skill measured in how students understand the text, interpret the meaning of the text, evaluate the text and apply their reading ability into their real-life situations (Mullis et al., 2015; OECD, 2016). The international assessment programs assessing student's reading literacy are the Programme for International Student Assessment (PISA) and Progress in International Reading Literacy Study (PIRLS). PISA has categorized reading literacy for 15-year-old students as the following: 1) Locate information, 2) Understand, and 3) Evaluate and reflect (OECD, 2016). PIRLS has divided reading literacy for students in their fourth year of schooling as 1) Focus on and retrieve explicitly stated information, 2) Make straightforward inferences, 3) Interpret and integrate ideas and information, and 4) Evaluate and critique content and textual elements (Mullis et al., 2015). Moreover, National Assessment of Educational Progress (NAEP) which

measures American student's academic achievement has grouped reading literacy into three dimensions, including 1) Locate and recall, 2) Integrate and interpret, and 3) Critique and evaluate (National Assessment Governing Board, 2017). When considering research on reading literacy in Thailand, Diowvilai et al. (2012) have categorized reading literacy into five processes, including 1) Decoding, 2) Reading fluency, 3) Reading comprehension, 4) Analytical reading and interpretation, and 5) Critical reading. In addition, Praputtakun et al. (2013) have divided reading literacy into two aspects: 1) R etrieving information and 2) Interpretation. As a result, reading literacy has been similarly divided into three main processes, including 1) the ability to locate and receive information, 2) the ability to understand, integrate and generate inferences, and 3) the ability to evaluate and reflect information (Diowvilai et al., 2012; Mullis et al., 2015; National Assessment Governing Board, 2017; OECD, 2016). This study adopted the reading framework from PISA 2018 as the research reading literacy framework due to its up-to-date framework emphasizing online reading literacy as well as the inclusion of competencies to measure the ability to read multiple texts. Moreover, PISA reading literacy has been promoted as one of the indicators for the quality of education in Thailand.

The Office of the Basic Education Commission has encouraged teachers and educators to promote reading literacy through formative assessment for the benefits of improving students' reading ability (Office of the Basic Education Commission, n.d.). Formative assessment has been suggested to evaluate the process of learning. Good formative assessment helps teachers to diagnose their students ' learning needs and engage students in their own learning process (OECD/UNESCO, 2016). Moreover, formative assessment helps lower achieving students, who have little opportunity to develop their capacity and score poorly on the standardized assessment, to get the higher achievement. The formative assessment can provide appropriate learning strategies and remedial learning opportunities for them (OECD/UNESCO, 2016).

Dynamic assessment (DA) has been defined as a procedure integrating teaching and assessment simultaneously to assess and to promote learner's zone of proximal development by offering appropriate forms of mediation to learners during the assessment process (Davin, 2013; Elliott, 2003; Lantolf & Poehner, 2004). The idea of dynamic assessment has originally come from the notion of Vygotsky' Zone of Proximal

Development. In his belief, learner independent performance or 'individual's zone of actual development' (ZAD) is interpreted as an ability that has fully formed. The ability that has not yet completed is called 'individual's zone of proximal development' (ZPD), that is, what the individual is able to do with help presently, he or she is able to do it later alone (Grigorenko & Sternberg, 1998; Guthke & Wingenfeld, 1992).

Dynamic assessment differs from the standardized assessment, the traditional paper-and-pencil test, especially in terms of mediation. Mediation is the combination of cooperative activity and interpersonal interaction to assist learners to improve their hidden performances (Poehner et al., 2015). The examiner plays a crucial role as a mediator intervening learners by using a variety of mediations to evaluate learners' learning process and the quality of mediation (Grigorenko & Sternberg, 1998). Resultant diagnoses include not only whether learners answered correctly but also the amount of support or mediation they required during the test which can help learners to attain their learning potential (Poehner et al., 2015). Grigorenko and Sternberg (1998)said that learning potential is something developing and modifiable rather than something developed. In other words, learning potential is the measurement of ability to learn by teaching them something and observing their learning to predict their future performance. Moreover, learning potential may be understood as an openness to mediation, providing insights into the degree of instructional effort that will likely be required to help individual learners attain their learning fulfillment (Poehner et al., 2015). According to McNeish and Dumas (2017), they said that learning potential takes account of (1) ability, the amount of skill that a student has at the current time; (2) availability, the differences between capacity and ability; and (3) capacity, the maximum amount of ability that a student can be expected to attain. The availability importantly captures the zone of proximal development showing learner responsiveness to mediation during the test, which is an essential part of diagnosis and more relevant to future teaching and learning. Poehner et al. (2015) said that learners who performed at the same level independently may differ in their learning potential or responsiveness to mediation, which can be supposed that learners require very different levels of instructional support in their learning in the future. Therefore, the goal of dynamic assessment is to evaluate, to intervene, and to change (Grigorenko & Sternberg, 1998).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

There are numerous approaches for dynamic assessment, varying in terms of the degree of structure and the timing of the intervention, as well as the content of the intervention procedure. Dynamic approaches can be categorized into two forms, interventionist DA and interactionist DA (Lantolf & Poehner, 2004). The interventionist DA addresses in more standardized and systematic ways. It measures learning potential by quantifying numerical scores and calculating the amount of help learner required to attain the predetermined endpoint. This form consists of learning potential testing from Budoff, Leipzig learning test approach from Guthke, testing-the-limits approach from Carson and Wiedl, and graduated-prompts approach from Campione and Brown. On the other hand, the interactionist DA focuses on more holistic and flexible measures to match with an individual without concerning the predetermined endpoint. This form of DA is mediated learning experience from Feuerstein and colleagues (Poehner, 2008). This study utilized the interventionist approach because of its standardized procedures to quantify learning potential in systematic ways and could generalize the results to the large number of students.

With the use of computer and communication technology, the computer-based dynamic assessment or computerized dynamic assessment (C-DA) can be administered to large numbers of students and generates results for the benefits of teachers and learners (Poehner & Lantolf, 2013). Learners are provided mediation, such as prompts, leading questions, or hints, during the assessment procedure. The survey of PISA study shows that Thailand still lacked teaching and learning resources especially internet connection and computer software for instruction, which is an important factor to promote student learning (Dechsri, 2016). Accordingly, computerized dynamic assessment can automate and individualize feedback to students, provide immediate and customized coaching, and generate progress and proficiency reports to teachers and students individually or in groups (Poehner & Lantolf, 2013; Poehner et al., 2015).

Most C-DA studies have utilized the graduated-prompts approach as a training procedure (Poehner & Lantolf, 2013; Resing & Elliott, 2011; Stevenson et al., 2016; Teo, 2012; Wang & Chen, 2016; Wang, 2011; Wu et al., 2017; Zhang & Lu, 2019; Zhang et al., 2017). In this procedure, mediation is provided during the assessment procedure and relies on the use of ‘prompts’ that are graduated in the sense that they are arranged from most

implicit to most explicit (Poehner et al., 2015). There are several patterns of prompts in accordance with content as follows: 1) instructional prompt, 2) error-explanation prompt, 3) mixed prompt, and 4) verification prompt. Regarding the instructional prompt, the prompt relies on instruction or reading strategies to guide students to answer the questions correctly (Poehner & Lantolf, 2013; Teo, 2012; Wang & Chen, 2016; Zhang et al., 2017). For the multiple-choice questions with five answers, there are a total of four mediating prompts in each item (Poehner & Lantolf, 2013). Prompts are provided from the most general to more explicit until student could answer the question correctly. Prompts guided in gradually instructional method can assist students to discover or apply some principles to independently solve problems (Wang, 2010). Students can transfer the prompt to other questions by detecting similarities of the reading strategies between questions (Golke et al., 2015). For the error-explanation prompt, the prompt is provided in accordance with the incorrect option student selected, ranging from implicit to explicit. Each distractor provides different prompts to emphasize the error pattern that corresponded to selected response (Golke et al., 2015; Ting & Kuo, 2016). This type of prompt is preferred for tasks requiring higher-order cognitive processes because it aims to repair false links or gaps that student maintains (Golke et al., 2015; Petrović et al., 2017). In regard to mixed prompt, prompt is provided in accordance with the instruction to guide a student answers the question correctly and the reason why the option student chose is incorrect (Ting & Kuo, 2016). For verification prompt, the prompt is given only to tell a student that the response is either correct or incorrect. If a student answers incorrectly, the prompt is shown that the answer choice is incorrect (Golke et al., 2015). Different types of prompts might lead to different types of responses and possibly benefit students in different ways (Lin, 2016; Wu et al., 2017). As a result, the appropriate prompts are necessary to be implemented in classroom context.

Utilizing graduated prompts enables educators to determine the amount and type of instruction a student requires to perform to reach their potential levels (Stevenson et al., 2016). It is an effective approach in providing the assessment information for the teacher to provide remedial teaching and for students to give extra learning (King et al., 2015; Stevenson et al., 2016; Wu et al., 2017). This approach matches with a wide range of tested contents and meets the need in using for classroom and individual settings. Wang

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

and Chen (2016) said that computer-based assessment is applicable for most students with a wide range of reading abilities. The questions had been raised whether computerized dynamic assessment for reading literacy matched with students in varying groups. As a result, this study investigated interaction effects between different prompting methods and other external factors, including gender and levels of learning achievement, on reading literacy.

Dynamic assessment is also a technique featuring multiple testing occasions for estimating student potential (Dumas & McNeish, 2017). Feuerstein (1979, as cited in Dumas & McNeish, 2017) found out that cognitive abilities should be tested multiple times for increasing opportunities of learning integrated with measurement. According to Dumas and McNeish (2017), single-timepoint measurement underestimates the potential of students because they cannot provide valid information about the capacity of an individual or group in order to develop particular knowledge. Students who score poorly on the test may not be given the resources and attention or effort from educators to help them meet their potential; they may be perceived as low-capacity students. Stevenson et al. (2016) said that multiple learning opportunities through feedback or training are advantages because the resultant can take account of initial performance, learning rate, and instructional needs.

Consequently, dynamic assessment is a tool helping students to determine their weaknesses and improve their performances to attain full potentiality. Several studies on dynamic assessment have paid attention to the measurement of current performance of students. There have been few studies addressing the measurement of growth and the measurement of learning potential; repeated measures of two or more time points are hardly used in dynamic assessment research (Zhang et al., 2017). Moreover, research on computerized dynamic assessment and how different types of prompts influence students' performance is at a preliminary stage. There has been little research comparing different types of prompting methods (Wu et al., 2017). Although the graduated prompting method has been used to measure a wide range of subjects, little research has been conducted to measure reading literacy of students. Therefore, the goals of this research study were to measure reading literacy, growth, and learning potential of Grade 9 students by developing computerized dynamic assessment for reading literacy and comparing

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

different types of prompting methods to find out what prompting method best suits Grade 9 students. Ninth graders were selected to participate in this study because reading literacy is the solid foundation of basic knowledge and skills for pursuing their further education or career in the future. As a result, this study would bring usefulness in applying the theoretical concept of dynamic assessment into practice by developing the computerized dynamic assessment for measuring reading literacy, growth, and learning potential.

Research Questions

1. How computerized dynamic assessment for reading literacy of Grade 9 students is constructed and validated?
   1.1 How computerized dynamic assessment for reading literacy of Grade 9 students is constructed?
   1.2. How are the qualities of the computerized dynamic assessment for reading literacy of Grade 9 students in terms of psychometric properties and test parallelism?
2. Do different prompting methods of the computerized dynamic assessment for reading literacy affect reading literacy, growth of learning, and learning potential of Grade 9 students?
   2.1 Do different prompting methods of computerized dynamic assessment affect reading literacy score of Grade 9 students?
   2.2 Do different prompting methods of computerized dynamic assessment affect growth trajectories of reading literacy and reading literacy subscales of Grade 9 students?
   2.3 Do different prompting methods of computerized dynamic assessment affect availability score, mediated score, and levels of promptings of Grade 9 students?

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Research Objectives

1. To develop and validate computerized dynamic assessment for reading literacy of Grade 9 students
   1.1 To develop computerized dynamic assessment for reading literacy of Grade 9 students
   1.2 To validate the quality of computerized dynamic assessment for reading literacy of Grade 9 students in terms of psychometric properties and test parallelism
2. To measure reading literacy, growth of learning, and learning potential of Grade 9 students by comparing different prompting methods of the computerized dynamic assessment for reading literacy
   2.1 To measure reading literacy score of Grade 9 students by comparing different prompting methods of computerized dynamic assessment for reading literacy
   2.2 To investigate growth trajectories of reading literacy and reading literacy subscales of Grade 9 students by comparing different prompting methods of computerize dynamic assessment for reading literacy
   2.3 To examine availability score, mediated score, and levels of promptings of Grade 9 students by comparing different prompting methods of computerized dynamic assessment for reading literacy

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Research Hypotheses

Hypotheses regarding the measurement of reading literacy, growth, and learning potential were as follows:

Research studies on dynamic assessment have compared different types of prompting methods in assessing student performance (Ebadi et al., 2018; Wang, 2010; Wu et al., 2017). Some studies have utilized the instructional prompting compared with the control condition. Ebadi et al. (2018) said that students in computerized dynamic assessment had significantly higher vocabulary gains than those in the control group. Wang (2010) also said that students in web-based dynamic system enhanced learning effectiveness in Biology than those in the normal web-based test. When considering the

comparison between instructional prompting and error-explanation prompting, the study of Wu et al. (2017) compared three remedial instruction programs, including the direct instruction group, the instructional prompting group, and the error-explanation prompting group on Mathematics. It was found that students with the error-explanation prompting group outperformed other groups in learning mathematics. There has been no research focused on the combination between instructional prompting and the error-explanation called mixed prompting. With the advantages of both prompting methods, the combination of both methods may lead to better performance of students. Thus, even though there are no studies comparing three prompting methods directly, it can be hypothesized that prompting-based groups might have higher reading literacy scores when compared to the control condition, and mixed prompting might be the most effective method in measuring reading literacy score of Grade 9 students.

Regarding growth of learning, it was found that no studies related to dynamic assessment have compared different types of prompting methods to measure growth of learning. Most studies have utilized the instructional prompting method compared with control condition (Wang & Chen, 2016; Yang & Qian, 2020). Wang and Chen (2016) investigated the effects of the web-based dynamic assessment using the instructional prompting method on reading ability during four-week periods. The results found that the assessment had significant effects on the change or growth in students' reading ability. In addition, Yang and Qian (2020) showed that students in the experimental group performed significantly better than the control group in English language reading after four weeks of learning. As a result, computerized dynamic assessment probably lead to change of growth. In addition, prompt-based groups might have higher growth trajectories of reading literacy and reading literacy subscales than those in the control group.

In terms of learning potential, it was found that no dynamic assessment studies have compared different types of prompting methods to measure learning potential. Most of them have utilized the instructional prompting method compared with control condition. Concerning mediated score, Poehner et al. (2015) said that learners with high ability scores had higher mediated scores. Qiu and Van Compernolle (2021) said that students with high actual scores received higher

mediated scores. For levels of prompting, Wang (2011) showed that students in computerized dynamic assessment with instructional prompt needed fewer prompts to answer items correctly. Moreover, Resing and Elliott (2011) said that the number of prompts of students needed to solve the item decreased from the first training to the second training session. Consequently, in order to assess how students profited from the prompts, it can be hypothesized that students with prompting-based groups might have higher mediated scores than the control group. Moreover, prompting-based groups might have lower number of prompts required than the control condition from the first testing session to the third testing session.

As has been shown previously, no studies have compared different types of prompting methods on reading literacy. Thus, it could be hypothesized that reading literacy, growth, and learning potential might probably be varied based on different prompting methods. The hypotheses of this study were:

H1: Prompting-based group might have higher reading literacy scores when compared to the control condition, and mixed prompting might be the most effective method in measuring the reading literacy score of Grade 9 students.

H2: Prompting-based groups might have higher growth trajectories of reading literacy and reading literacy subscales than those in the control group.

H3: Prompting-based groups might have higher mediated scores and lower number of prompts required over time than those in the control group.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Scopes of the Study

1. Population

Population in this study was Grade 9 students studying at schools under the Secondary Educational Service Area Office 1 and the Secondary Educational Service Area Office 2.

2. Dynamic assessment approach

This study applied a graduated prompts approach from Campione and Brown, one of the interventionist dynamic assessment.

3. Tested content

The content was adapted from the framework of PISA reading literacy 2018, which was the newest framework and the major assessed domain in 2018. The framework of PISA reading literacy comprised three dimensions including 1) locate information, 2) understand, and 3) evaluate and reflect.

4. Research Instrument

The researcher developed two research instruments as follows: 1) a computerized dynamic assessment for reading literacy and 2) reading literacy pretest and posttest. There were three parallel tests of computerized dynamic assessment for reading literacy. Each test comprised 20 two-tier multiple-choice items with one correct answer.

5. Prompting methods

There were four prompting methods operated in computerized dynamic assessment for reading literacy as follows: 1) instructional prompting method, 2) error-explanation prompting method, 3) mixed prompting method, and 4) verification prompting method operated as the control condition.

6. Time

Computerized dynamic assessment for reading literacy was measured three time points. Each time point was four-week interval.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Definition of Terms

1. Computerized dynamic assessment for reading literacy

Computerized dynamic assessment for reading literacy refers to the system of computerized dynamic assessment developed to measure reading literacy skills. The system was based on research studies and theories related to dynamic assessment. The test format was 20 two-tier questions. For the first-tier item, there were five multiple-choice questions with one correct answer. The second-tier items were three multiple-choice questions that applied to avoid guessing. The instrument differed from other computer-based tests in terms of the application of graduated prompting approach and two-tier test. The graduated prompting approach was used during the assessment procedure as the mediated intervention provided student to learn and find

out the correct answer. Student would be allowed to do the second-tier item after he/she answered the first-tier item correctly at the first attempt. If student could not answer item correctly, prompts were provided in accordance wtih group conditions.

2. Psychometric properties of the computerized dynamic assessment for reading literacy

Psychometric properties of the instrument refer to the process of instrument validation. This study utilized the content validity to measure the reading literacy test items and the prompts written for different prompting methods. Moreover, this study demonstrated the application of item response theory for model fit and item parameter estimates, both item discrimination and item difficulty. Moreover, this study analyzed confirmatory factor analysis and reliability of the reading literacy instruments.

3. Test parallelism of computerized dynamic assessment for reading literacy

Test parallelism of the instrument refers to the content and statistical parallelism of the computerized dynamic assessment at the levels of test forms and individual items. The statistical parallelism of test forms was analyzed in terms of test information function (TIF) and test characteristic curve (TCC). Moreover, the item-by-item parallelism was measured by item information function (IF) and item characteristic curve (ICC).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4. Graduated prompting approach

Graduated prompting approach refers to the dynamic assessment approach that provided standardized prompts or mediation during the test administration. The outcome of this approach was the measurement of the student's efficiency of learning; in other words, the number of prompts or mediation needed. The use of 'prompts were divided into four types of prompting methods as follows:

4.1 Instructional prompting method

Instructional prompting method was the prompt that relied on the instruction or reading strategies to guide a student better understood the text or answered the questions correctly. The prompts were arranged from most implicit to most explicit instructions.

4.2 Error-explanation prompting method

Error-explanation prompting was the prompt that relied on the reason why the selected choice was wrong. If a student selected the incorrect answer choice, they would receive the prompt showing why the choice was incorrect, ranging from implicit to explicit.

4.3 Mixed prompting method

Mixed prompting was the prompt that based on the instruction to guide a student answered the question correctly and the reason why the selected choice was incorrect. If a student answered incorrectly, they would receive the prompt showing why the choice was incorrect and the instructions provided to better understand the questions, ranging from implicit to explicit.

4.4 Verification prompting method

Verification prompt was the prompt that given only to tell a student that the response was either correct or incorrect. If a student answered incorrectly, the prompt would be provided that the answer choice was incorrect.

5. Reading literacy

Reading literacy refers to the constructive process involving interaction between the reader and the texts with the intention of reading for growth and for participation in society as a local and global citizen. Reading literacy comprises the ability of a student to locate information, to understand, to make inference, to evaluate, and to reflect on information. There were three processes of reading literacy including: locate information, understand, and evaluate and reflect. The present framework was applied from PISA the 2018 reading literacy framework. The three dimensions or processes would be the following:

5.1 Locate information

Locate information was the ability of a student to quickly read the texts for specific purposes, without considering the rest of them. There were two tasks including 1) access and retrieve information within a text and 2) search and select the relevant task. The former task required a student to scan a single text to retrieve target information in the form of words, phrases or numerical values. The

latter task required a student to search and select information from several texts to retrieve information given the demands of the task.

5.2 Understand

Understand was the ability of a student to understand the meaning of the text and draw a conclusion of the text; in other words, the ability to read between the lines. There were two tasks in this dimension including 1) represent literal meaning and 2) integrate and generate inferences. The former task required a student to comprehend sentences or short passages by matching a direct or close paraphrasing of information. The latter task required a student to integrate and generate inferences from simple connecting inferences to more complex coherence relationships by creating the main idea, producing a summary, or giving a title for a passage.

5.3 Evaluate and reflect

Evaluate and reflect was the ability of a student to reflect on the content and form of the text and to evaluate the quality and validity of the information. There were three tasks in this dimension including 1) assess quality and credibility, 2) reflect on content and form, and 3) detect and handle conflict. The first task required a student to evaluate the quality and credibility of the text in terms of validity, accuracy, and bias. In addition, the task sometimes required a student to evaluate the source of information such as author, content, and form. The second task required a student to reflect on the quality and style of the writing from their own experience by determining the author's purpose of writing and the author's point of view. The third task asked a student to detect the soundness of the claims and handle conflict of the contradicting information within a text or multiple texts.

6. Growth

Growth refers to the measurement of change in reading literacy from multiple measurement occasions. The study aimed to investigate whether computerized dynamic assessment led to change in reading literacy over three time points by using latent growth curve models. The growth in reading literacy was estimated as follows: 1) the univariate latent growth curve models estimated the overall reading literacy and three measured constructs separately, which were locate information,

understand, and reflect and evaluate, to determine the most fitted shape of curve of each construct and understood which prompting groups changed in each construct over time. Moreover, the multivariate growth curve model estimated the association among the three measured constructs over time in order to understand how change in one variable related to change in other variables. The model would estimate all three constructs from each univariate growth curve model simultaneously. The covariances of intercepts and covariances of slopes of the interrelated constructes were compared among different prompting groups.

7. Learning potential

Learning potential refers to the reading literacy ability that a student could not do presently but needed help or assistance to develop his or her own potential in reading literacy in order to be able to accomplish it independently in the future. It was consistent with the idea of the zone of proximal development (ZPD) of Vygotsky. Learning potential in reading literacy could be presented by estimating the mediated score in reading literacy. Otherwise, in order to support the efficiency of prompting, this study estimated three terms of learning potential as follows: availability score, mediated score, and levels of prompting.

7.1 Availability score

Availability score in reading literacy was the measurement of distance between the estimated reading literacy capacity and the current performance in reading literacy of a student over time. Availability score in reading literacy was individually calculated by subtracting the estimated ability from the subject-specific capacity estimate. It indicated how far a student would reach the attainment. The availability score in reading literacy of different prompting methods in each time point would be compared.

7.2 Mediated score

Mediated score in reading literacy was the amount of external mediation or support a student needed to attain their goals and it was also an evidence of learning occured during the assessment. The number of prompts counted in the form of mediated scores in reading literacy, ranging from 4-0 points. In each item, the total 4 points decreased by 1 if a student received each mediating

prompt. The total number of mediated scores in each treatment would be counted and compared. It was assumed that a student who benefitted from the prompts would require fewer prompts during the course of training and obtain higher mediated scores.

7.3 Levels of prompting

Levels of prompting were the count of prompts in each level that a student required to answer items correctly. The prompts were arranged from implicit to explicit. The first level of prompting was the most implicit prompt a student needed for the first time when he/she answered incorrectly. The second level of prompting was the implicit prompt a student required for the second time when he/she answered incorrectly. The third level of prompting was the explicit prompt provided for a student when he/she responded incorrectly. The fourth level of prompting was the correct answer choice provided for a student when he/she answered incorrectly. The levels of promptings in each different prompting treatment would be analyzed and compared.

Significance of the Study

1 . The results of this study would be beneficial for teachers to employ computerized dynamic assessment as an alternatively innovative instrument to measure student's reading literacy.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2 . The results of this study would provide guidance for teachers to use appropriate prompting methods to measure both current ability, growth and learning potential in the classroom.

3. The results of this study would provide useful information on a student's growth and learning potential for teachers to diagnose students and design their future classroom instructions that might help students achieve his/her fulfillment in reading literacy.

CHAPTER TWO
LITERATURE REVIEW

This chapter involved reviewing theoretical and practical concepts of research related to dynamic assessment, computerized dynamic assessment, measurement of reading literacy, measurement of growth, measurement of learning potential, and test validation using item response theory, equivalence analysis, and research framework. This study consisted of eight parts as follows:

Part 1: Dynamic assessment

Part 2: Computerized dynamic assessment (C-DA)

Part 3: Measurement of reading literacy

Part 4: Measurement of growth

Part 5: Measurement of learning potential

Part 6: Test validation using item response theory

Part 7: Equivalence analysis

Part 8: Research framework

Part 1: Dynamic Assessment

Dynamic assessment (DA), an alternative assessment approach, is widely popular in the areas of clinical and pedagogical research. It is grounded by Vygotsky's notion on the zone of proximal development which aims to promote learner's potential development. It is believed that using appropriate forms of mediation can help a learner to attain his/her learning potential. In this part, the concept, theory and approaches related to dynamic assessment are reviewed.

1.1 Concept of dynamic assessment

1.1.1 Definition of dynamic assessment

The term ***'dynamic assessment'*** is an umbrella term for the assessment with mediation. It has been defined as a procedure integrating teaching and assessment simultaneously to assess and promote learner's zone of proximal development or learning potential by offering assistance to learners during the

assessment process (Davin, 2013; Elliott, 2003; Lantolf & Poehner, 2004). Moreover, it has been referred to an assessment procedure combined with active teaching to modify cognitive functioning (Feuerstein et al., 2003; Tzuriel, 2001). Therefore, it can be said that the definition of dynamic assessment is the assessment to evaluate, to intervene, and to change a learner's ability (Grigorenko & Sternberg, 1998). Several names have been used referring to 'dynamic assessment', including interactive assessment, process assessment, measuring the zone of proximal development, assisted assessment, and tests of learning potential (Grigorenko & Sternberg, 1998).

1.1.2 Theory grounded for dynamic assessment

The central idea of dynamic assessment attaches with Vygotsky's notion about a zone of proximal development and mediation (Guthke & Wingenfeld, 1992). The theoretical work of Lev Vygotsky about a zone of proximal development (ZPD) has been applied to the research on dynamic assessment. ZPD is the discrepancy between the actual development level of learner determined by the number of solved tasks independently and the higher level of potential development that measured by solving a problem through assistance or guidance from adults or capable peers (Grigorenko & Sternberg, 1998; Guthke & Wingenfeld, 1992). This notion has initially been developed to stress the problem of measuring mental age which may not disclose the true performance. Instead of focusing on the learner's developed ability, Vygotsky is interested in learner's developing ability. He believed that what the learner currently does by being assisted of others will lead to what he or she can do independently in the future (Davin, 2013; Ku et al., 2014; Lantolf & Poehner, 2004; Poehner et al., 2015; Shrestha, 2017; Wu et al., 2017). The ability that has fully formed or the test task that student is able to solve by themselves is called *independent performance* or an *individual's* ***zone of actual development (ZAD),*** whereas the ability has just emerged from interaction or help with mediator or more capable peer is referred to *learner responsiveness to support* or *individual's* ***zone of proximal development (ZPD).*** In Vygotsky's view, most assessments reveal only ZAD because they do not allow mediation into the process (Lantolf & Poehner, 2004; Poehner et al., 2015). Therefore, the heart of DA is not only

to find out actual abilities but also to stress developing potential (Guthke & Wingenfeld, 1992; Lantolf & Poehner, 2004).

The developing potential is not static but varies from person to person. Two persons who have the same level of actual development or ZAD may be different in their level of potential development or ZPD even though the same assistance has been provided; one child may be more capable when he is assisted with mediator than others (Elliott, 2003; Lantolf & Poehner, 2004). A learner with broad ZPD benefits more from the intervention and needs less help than a learner with narrow ZPD (Tzuriel, 2001). Therefore, the nature of ZPD comprises development, social interaction and forward-looking concepts (Cotrus & Stanciu, 2014; Grigorenko & Sternberg, 1998; Lantolf & Poehner, 2004).

The core concept of a zone of proximal development (ZPD) is ***'mediation'*** which emerges from learner interaction with stimuli. Cognitive ability is not constructed directly by itself, but it is mediated from the surroundings by using various resources in the form of examples, leading questions, and correctly answering demonstrations (Lantolf & Poehner, 2004). Relying on the internalization process, the actual development can be transformed and becomes internalized after guidance and responsiveness to assistance (Grigorenko & Sternberg, 1998).

1.1.3 Goal of dynamic assessment

The goal of dynamic assessment differs from other measures. It is aimed to measure learning potential, not current performance. The initial idea of learning potential came from Vygotsky's notion of the zone of proximal development. He mentioned the greater development of learner performance through the assistance of older or more capable peers (Grigorenko & Sternberg, 1998; Lantolf & Poehner, 2004; Tzuriel, 2001). Regarding Haywood and Lidz (2007, p.321), learning potential is: "the capacity to learn more effectively than was either demonstrated by present performance or predicted by standardized tests of intelligence." In other words, the concept emphasizes the greater capacity of learning rather than the learner's current abilities (Haywood & Lidz, 2007).

1.1.4 Discrepancy between dynamic assessment and static assessment

Standardized assessment is often called static assessment, which refers to assessment conducted without intervention to change or improve a learner's performance (Tzuriel, 2001). In order to better understand the term 'dynamic assessment', several scholars have compared it with a static assessment in several dimensions.

There are several differences between dynamic assessment and static assessment including goals, process, and result interpretation. In terms of goals of assessment, Feuerstein et al. (2003) stress that while static tests focus on stage and progress of development, dynamic assessment emphasizes the character and process of mental development. Lantolf and Poehner (2004) and Feuerstein et al. (2003) state that dynamic assessment's perspective helps a learner to attain the future by taking into account of learning potential or modifiability based on changes during the assessment, whereas static assessment mentions on the current state of a learner which developed abilities take place. This is consistent with Grigorenko and Sternberg (1998) and Tzuriel (2001) that mention the focus of dynamic assessment is on what a learner achieves with adult's assistance, the process of learning, deficient cognitive function, and strategies to facilitate learning, whereas the static assessment is focused on a product of existed skills.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

When taking into account the assessment process, Lantolf and Poehner (2004) state that dynamic assessment is involved with mediation, not the instrument itself. Grigorenko and Sternberg (1998) and Elliott (2003) point out that feedback is considered as a crux of learning for dynamic assessment to help a learner formulate their knowledge, whereas there is no feedback for static assessment due to the concern of the quality of testing. Moreover, for the relationship between examiner and examinee, Grigorenko and Sternberg (1998) and Elliott (2003) mention that it has a closer relationship for dynamic assessment since involvement and interaction play the main role mediating a learner to transform his/her knowledge. Moreover, Tzuriel (2001) says that the static test does not relate the non-intellective factors such as intrinsic motivation, anxiety, frustration, self-

confidence, locus of control, while DA is more holistic that involves cognitive and affective factors.

Regarding result interpretation, Haywood and Lidz (2007) and Feuerstein et al. (2003) state that static assessment and dynamic assessment differs in terms of comparison of the results. Score or interpretation of static assessment is generally applied to all learners and influences learners to compare their results with normative groups of peers. On the other hand, for dynamic assessment, the comparison is within the person and leads the learner to compare the results with himself/herself at different times and under different conditions.

1.1.5 Target groups of dynamic assessment

Dynamic assessment can probably be applied to multiple contexts. Normally, it can be used as a basis for teaching and improving cognitive skills. Teaching is an intervention to strengthen learners' ability and help them to learn independently. In addition, it can be used with certain groups of learners, especially disadvantaged learners, to promote learning potential. Disadvantaged learners are defined as those who performed poorly on traditional testing, who lacked an opportunity to learn, who suffered from language problems, who had cultural differences, and who suffered from mental retardation, learning disability, emotional disturbance, personality disorder, or motivational deficit. Dynamic assessment is fair and equitable for assessing learning (Grigorenko & Sternberg, 1998; Haywood & Lidz, 2007). It is believed that these learners were deprived of learning experiences more than others which caused them to fail from cognitive functioning. Kozulin (2014) said that creating an assessment situation for disadvantaged learners to learn prerequisite skills and demonstrating the ability to apply these skills in problem solving is one of the goals of dynamic assessment.

1.2 Approaches of dynamic assessment

As mentioned before, dynamic assessment is an umbrella term for assessment with mediation. Dynamic assessment has been categorized in various ways (Feuerstein et al., 1987, as cited in Lantolf & Poehner, 2004; Sternberg & Grigorenko, 2002; Tzuriel, 2001) as follows:

Sternberg and Grigorenko (2002) have categorized DA approaches into two formats: 'the sandwich format' and 'the cake format'. (1) For the sandwich format, the intervention or instruction is sandwiched between the pretest and posttest.

The content of the pretest should be beyond the learner's zone of actual development. After taking the pretest, intervention is provided. The intervention can be either an individual or group setting. Type of instruction and amount of instruction can be varied in an individual setting, but they will be the same for all learners in group settings. Following the intervention, posttest will be provided to measure performance. Posttest should be the parallel form of the pretest; they may be the same test in some situations. (2) For the cake format, the intervention is layered during test administration. Examinees will be provided item by item. If the examinee responds to the item correctly, they will turn to do the next one. If their answers are incorrect, the series of prompts will be provided immediately, ranging from implicit to explicit. The layered format shows how many and what kind of prompts the examinee needs to solve the problem (Elliott, 2003; Sternberg & Grigorenko, 2002; Wang & Chen, 2016).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Feuerstein et al. (1987, as cited in Tzuriel, 2001) have proposed two approaches of DA: 'functional modification' and 'structural modification'. (1) The functional modification is the intervention intended to change the functioning of a person related to a specific task such as the coaching of learning potential testing from Budoff and graduated prompting from Campione and Brown. It is a more standardized procedure than another approach. (2) The concept of structural modification, on the other hand, aims to produce and seek changes in cognitive ability inside. Structural changes are the outcome of this approach such as the mediated learning experience from Feuerstein and colleagues.

Lantolf and Poehner (2004) have proposed two forms of dynamic assessment which are 'interventionist' and 'interactionist'. (1) Interventionist DA, which is a more standardized form that focuses on quantifying the number of assistance provided to help a learner achieves their learning potential. Mediators do not allow providing free responses to an examinee but they have to follow pre-scripted procedure that prompts, hints and leading questions are sequenced from implicit to explicit. (2) On the other hand, interactionist DA, that is closely related to Vygotsky's notion of the zone of proximal development. It emphasizes the process of learning and how the mediator interacts with the learner to help his/her development. The decision to use either interventionist or interactionist depends on the goal and circumstances of the assessment (Davin, 2013; Lantolf & Poehner, 2014).

In summary, there are similarities between Feuerstein's and Lantolf and Poehner's types of approaches. Both studies divide approaches based on their functioning. However, the study of Sternberg and Grigorenko emphasized the importance of formats of learning as shown in Table 1.

Table 1 Summary of types of approaches

Study	Type of approach
Feuerstein et al. (1987, as cited in Tzuriel, 2001)	Functional modification e.g., learning potential testing (Budoff) and graduated prompting (Campione and Brown)
	Structural modification e.g., mediated learning experience (Feuerstein and colleagues)
Lantolf and Poehner (2004)	Interventionist DA e.g., learning potential testing (Budoff), Leipzig learning test (Guthke and colleagues), and graduated prompting (Campione and Brown)
	Interactionist DA e.g., mediated learning experience (Feuerstein and colleagues)
Sternberg and Grigorenko (2002)	The sandwich format e.g., learning potential testing (Budoff), long-term Leipzig learning test (Guthke and colleagues), and mediated learning experience (Feuerstein and colleagues)
	The cake format e.g., graduated prompting (Campione and Brown), short-term Leipzig learning test (Guthke and colleagues)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.3 Graduated prompting approach

1.3.1 Concept of graduated prompting approach

This approach is proposed by Campione and Brown to assist learner's ability in solving problems gradually. The emphasis is on learning gradually (Campione & Brown, 1985). The student's learning should be guided by an adult, and mediation or guided instruction should highlight the acquisition or rules or principles (Campione & Brown, 1985). The standardized hints and leading questions are provided during the administration of the test, along with mediation arranged from most implicit to most explicit (Poehner, 2008). The uniqueness of this approach is transfer (maintenance of learning) which aims to explore not only the ability of learners to master relevant principles but also the ability to apply what they have learned to the new situation or in a variety of contexts (Poehner, 2008). The outcome of this approach is to measure the learner's efficiency of learning, which is defined as the number of prompts or mediations needed to be required as well as how far a learner transfer their knowledge to reach the intended level (Grigorenko & Sternberg, 1998; Lantolf & Poehner, 2004; Poehner, 2008; Tzuriel, 2001; Wu et al., 2017).

1.3.2 Tested content

Assessment should be conducted in a specific domain rather than general intelligence, especially the domain that rules and principles can learn and can be adapted to novel problems or tasks (Campione & Brown, 1985). Types of tasks for this approach are composed of inductive-reasoning problems (e.g., progressive-matrices problems, geometric matrices, series-completion problems), mathematics problems, reading, and listening comprehension tasks.

1.3.3 Testing procedure

Normally, the sequence of testing comprises 1) pretest stage, 2) learning or testing stages, which includes mediation and transfer task, and 3) posttest stage (Grigorenko & Sternberg, 1998). In this procedure, the initial evaluation of a student competence is tested as the pretest. Then, a student will be placed in a mini-learning environment where they collaborate with a mediator until they are able to solve problems autonomously. When a student has struggle or failure in answering the question, he or she will be offered mediation as predetermined hints, prompts, and leading questions, ranging from implicit to

explicit, to help them apply principles necessary for problem solving. The mediation procedure aims to measure the minimum amount of help a student needs to solve each problem as the learning efficiency of each student (Campione & Brown, 1985). When a student is finally able to master their task independently, he or she will be tested on the ability to transfer their knowledge to novel problems (Campione & Brown, 1985; Poehner, 2008). After this session, the posttest is given (Campione & Brown, 1985).

1.3.4 Prompting procedure

For the promptings used in this approach, the prompting often ranges from implicit to explicit. The example of using prompting is from Stevenson et al. (2016), learners were assigned into three groups: graduated prompts, practice control, and attention control. For pretest and posttest, learners were assessed 22 analogy problems. During the intervention session, the first group received the graduated prompts training. The second group received no mediation and the final group received a maze coloring task. In the group of graduated prompting procedure, the task was analogical reasoning presenting in 2×2 matrix format as shown in Figure 1 learners performed the task by inferring the relation of horizontal and vertical pictures to solve the analogy (A:B::C:?) and responded by choosing the animal cards. If they answered correctly, the mediator asked the child for his/her reasoning before moving to the next question. If the child answered incorrectly, the prompt was provided as follows: 1) the first prompt helped learners to recognize the problem and redefined how the learners previously solved such a problem. 2) In the second prompt, learners received the aid card presenting the general steps to solve the analogies: compare, think, solve, and check. 3) The third prompt was related to guiding and inference in which the examiner worked through the aid card showing solutions by using words and gestures. 4) The fourth prompt was related to transformations both horizontally and vertically as well as inference and mapping. 5) For the final prompt, the examiner provided scaffold questions such as “Which color should the elephant be?” or “Which direction should the dog face?” After that, direct feedback was provided to guide learners step-by-step to the correct solution. The flowchart of the graduated prompting procedure was shown in Figure 2 (Stevenson et al., 2016, p.30).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

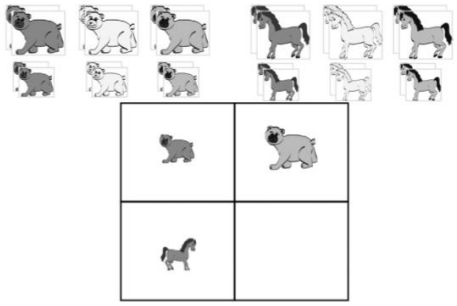


Figure 1 The AnimaLogica training item

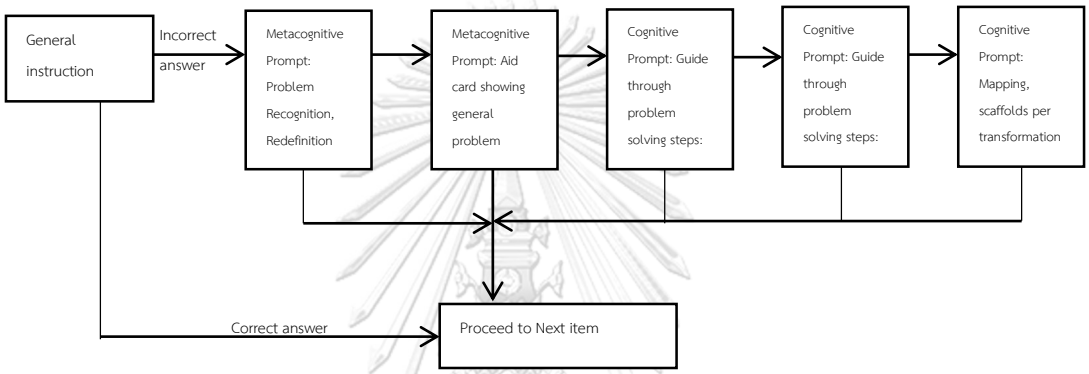


Figure 2 Graduated prompting flowchart

1.3.5 Scoring

The indicators of learning of this approach are on how much assistance needed and how far an individual can apply to novel tasks or situations. The outcome of this approach emphasizes the amount of help or assistance needed to reach the predetermined endpoint or to acquire a rule or procedure (Campione & Brown, 1985).

Therefore, the graduated prompting approach was used in this study because of its suitability to measure cognitive ability in specific areas as well as its adaptability to measure the normal groups of students.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

1.4 Mediation of dynamic assessment

Mediation is considered as the center of dynamic assessment. It provides an opportunity for development. Shrestha (2017) said that mediation is the process that humans use artifacts, concept, activity and other people to mediate or regulate any human abilities' mental functions.

There are several types of mediation for interventionist DA such as leading questions, models, prompts, and feedback. The type of mediation has relied on the format of dynamic assessment. In a sandwich format, the mediation is explicit for the individuals, but it will be more implicit in a group. There are three conditions related to mediation to the ZPD. First, the assistance should be graduated from implicit to explicit. Second, help should be offered only when needed. Explicit assistance should be provided only when implicit assistance is done unsuccessfully. Third, help should be provided in dialogue form conveying the intended meaning between mediator and learner (Davin, 2013).

The widely used mediation is prompting. Prompts are questions or hints to guide and support students' solving problems indirectly as well as promote metacognitive strategies (Davis, 2009). Prompts can be simple questions (e.g., "What is the first step of doing this activity"), incomplete sentences (e.g., "To solve the problem step by step, I have to...."), explicit instruction (e.g., "First,.....) or pictures or graphics provided (Bannert, 2009, as cited inIfenthaler, 2012).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

According to Davin (2013) who conducted graduated prompt in the classroom context, she utilized mediation in language classroom to assess lexical and grammatical errors of 17 Grade 4 and Grade 5 primary school students. The mediation was prompted on the level of explicitness as follows: 1) prompt 1 – pause with a questioning look, 2) repetition of the entire phrase, 3) repetition of the specific site of error, 4) forced choice option, and 5) correct response and explanation provided. It was found that this approach could mediate lexical and grammatical errors that were routine and predictable. She found two advantages in systematic mediation. First, it helped mediators to respond and correct learners' errors systematically by using the same pre-scripted prompts. Second, the mediator could respond quickly and kept track of learners' progress more easily by pointing to the

number and types of prompts learners require. She suggested that it was useful for both mediator and learner for remedial classrooms in the future. This method was useful to develop individual ZPDs. Even though the interaction took place between teacher and a learner, other learners in class benefited indirectly from the mediation.

1.5 Transfer

Transfer is the ability to transform learning into a new context (Shrestha, 2017). Regarding Poehner and Lantolf (2013), transfer is important to diagnose the learner ability from the increasing level of task difficulties as well as to predict learning potential in the future. The concept of transfer is used in Brown's and Feuerstein's approaches. The graduated prompt approach from Brown and colleagues applies transfer to use in standardized format of pre-scripted prompts ranging implicit to explicit. Transfer items are provided to measure the ability of learners to apply what they have learned to new situations or in a variety of contexts or the ability to encounter difficult tasks. The outcome of this approach is to measure the learner's efficiency of learning, which is defined as the number of prompts or mediations needed to be required as well as the breadth of transfer or how far learners transfer their knowledge to reach the intended level. In Feuerstein's approach, the concept of mediated learning experience (MLE) also involves transfer. It is the concept of transformation that the cognitive change is applicable in other contexts and tasks. Mediators play an essential role in order to reinforce a learner's progression from the task they previously perform with help to the task they will be able to perform independently in the future (Poehner & Lantolf, 2013; Shrestha, 2017).

There are two main types of transfer: 'near transfer' and 'far transfer'. ***Near transfer*** is the transfer of knowledge and skill to the context or task that is similar to the previous state. On the other hand, ***far transfer*** refers to the transfer of knowledge and skill to more complex and challenging contexts or tasks (Shrestha, 2017). However, Lantolf and Poehner (2004) argued that the breadth of transfer can be extended to 'very far transfer'. Learners will be provided specific questions to

tackle the learner's ability to transfer. Firstly, the learner is given ***'novel exemplars'*** of the relevant principles. After that, ***'near transfer'*** problems are given which integrate the same principles and the new ones. Then, the ***'far transfer'*** problems are presented combined with the new problem but related to the same principle. Finally, the ***'very far transfer'*** problems are presented with more complex problems to be solved (Lantolf & Poehner, 2004; Tzuriel, 2001).

The example of transfer task using the breadth of transfer developed by Ferrara, as cited in Grigorenko and Sternberg (1998), are shown as follows: (1) the pretest stage, learner was tested the ability to solve simple two-digit problem (e.g., 3+2 = ?). (2) During the learning session, the examiner worked with the examinee collaboratively, and the math problem was presented as word problems for better understanding. In this stage, the examiner provided a sequence of hints and suggestions. The amount of help to master the procedure was the outcome of learning. (3) Following the learning session, the transfer problems were presented, sequencing from **near transfer** (e.g., 3+1 = ?), **far transfer** (e.g., 4+3+2=?), and **very far transfer** (4+? = 6). The outcome of this session was the amount of help provided to solve these transfer items. (4) The posttest stage, the learner was assigned to master the task.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Several studies have included transfer in their procedures especially in the context of higher education (Gellert & Elbro, 2017; Poehner & Lantolf, 2013; Shrestha, 2017). The study of Shrestha (2017) revealed the potential of dynamic assessment may contribute to transfer to the new assessment context. She studied the writing assignment in distance learning. There were two DA tasks and a transfer assignment. She said that the differences between DA and transfer assignments were (1) the transfer task was longer than the DA task; (2) there were no data provided for students; they had to find their own data to support the analysis; and (3) the TA task required students to apply several business concepts and framework taught in the

modules rather than the only one framework. The results showed that each participant had a different level of transfer. The detail of transfer assignment was as presented in Table 2:

Table 2 Comparison between DA task and TA task (Shrestha, 2017, p.6)

Task	Question
DA 2 task	Read the three case study texts about Vodafone's broadband market mentioned above and write a SWOT analysis of this product based on the articles. Your SWOT analysis should be of about 500 words.
TA task	Using tools and concepts from Units 1-4 of Block 3 (including, if you wish, the analysis referred to in Part 1 of this assignment), write a report which identifies the key challenges facing an organisation with which you are familiar (you may use the same organization you described in Part 1, or in Assignment 01 Part 2 or Assignment 02). (2000 words)

The study of Gellert and Elbro (2017) also applied the transfer in their study for predicting students' ability of decoding at the end of kindergarten to the end of the first grade. Children in this study were taught 3 novel letters and how to blend the sound into the new words. Then, they were assessed on different words combining these three letters. After that, the parallel test was assessed after 5 months of the first administration which employed the same learning concept but different letters and sounds. The ability of transfer was measured by scoring of reading outcome which resulted from the unaided performance in independent reading. According to Poehner and Lantolf (2013), they included transfer items in their computerized dynamic assessment (C-DA) of French and Chinese listening and reading tests. The transfer items were the same as other test items in terms of response options, prompts and particular constructs. However, it differed in terms of level of difficulty. It was not intended to recall a particular lexical item but to interpret the texts with the higher level of lexical, grammatical, cultural, and phonological

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

dimensions of the texts. The example of a transfer item was a listening test about family dinner routine as in Figure 3 (Poehner & Lantolf, 2013, p.333). The learner had to interpret the meaning of the dialogue he or she heard and make a conclusion. The scores of transfer items were calculated as 'the transfer scores' which was later compared with learning potential scores (LPS), the scores predicting learning outcomes in classrooms.

According to the passage: (a) The family is vegetarian

(b) The family almost always has a dessert.

(c) The family eats mostly meat.

(d) The family rarely has dessert.

(e) The family always eats cheese with their meal.

Figure 3 Listening transferred item (Poehner and Lantolf, 2013, p.333

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Part 2: Computerized Dynamic Assessment (C-DA)

In order to lower the drawbacks of dynamic assessment in terms of the need of experienced mediator and the large quantity of time for test administration, the concept of using dynamic assessment procedure on computer called computerized dynamic assessment (C-DA) has been developed (Ku et al., 2014; Lin, 2016; Poehner & Lantolf, 2013; Poehner et al., 2015; Shih et al., 2013; Ting & Kuo, 2016; Wu et al., 2017; Zhang et al., 2017). This instrument is used as an assessment to better capture learner's independent performance and mediated performance as well as to predict learning potential hidden in each learner (Poehner & Lantolf, 2013; Poehner et al., 2015). The C-DA has been popular because of its advantages in administering a large number of students as well as generating the diagnostic results for the classroom settings (Poehner & Lantolf, 2013). This instrument has been widely applied in many fields of education such as accounting, mathematics and second language learning.

2.1 Procedures of computerized dynamic assessment

The computerized dynamic assessment is developed based on the cake format, in which mediation will be provided during the test administration. Students will receive the prompt while having some difficulty during the assessment. Several studies have developed the computerized dynamic assessment (Poehner & Lantolf, 2013; Teo, 2012; Wang & Chen, 2016). The work of Poehner and Lantolf (2013) and Poehner, Zhang, and Lu (2015) developed the computerized dynamic assessment of L2 reading and listening comprehension tests. Before administering on a computer, the developed test was piloted in paper-based form and administered without mediation to investigate whether the test items were in line with the intended constructs. Then, the test was re-administered in the paper-based version by using mediation during the administration provided prompts and feedback to help a student attained the intended level. This session was used to develop the appropriate mediating prompts for the computerized version afterward. Then, the computerized dynamic assessment based on intervention approach was administered. Mediation provided as cake format during the computerized dynamic assessment procedure. Several studies have represented the C-DA procedures similarly. If a student's first response is incorrect, a mediating prompt will be offered and he or she will be able to re-attempt the question. The first mediating prompt is the most

implicit. If the second response is incorrect, another and the more explicit mediating prompt will be offered. This process continues until student answer correctly or all prompts have been disclosed. The procedures C-DA are as shown in Figure 4.

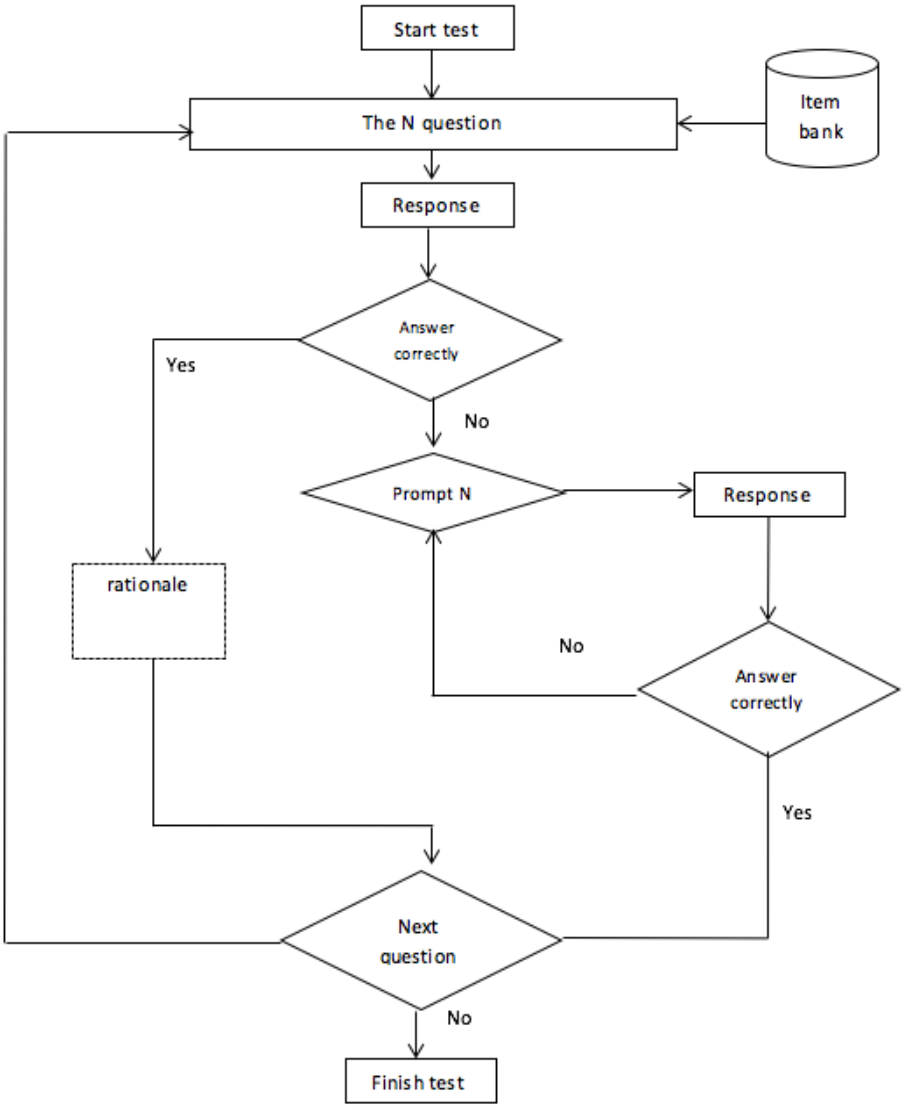


Figure 4 Construction of computerized dynamic assessment (adapted from Wu, Kuo, & Wang, 2017)

2.2 Mediation of computerized dynamic assessment

The mediation of computerized dynamic assessment is different based on the tested contents, approaches, and intended measure. Most of the mediations in the computerized dynamic assessment are prompting.

2.2.1 Levels of prompting

The prompts are statements, hints, leading questions that are provided gradually from most implicit to most explicit in order to help a student achieve their attainment which is coherent with the notion of zone of proximal development (ZPD). The prompt will be provided when a student answers incorrectly. The first prompt will be the most implicit or the most general. If a student answers incorrectly again, the explicit prompt will be provided. A student will be allowed to answer the question again. If he or she answers incorrectly, the third prompt with more explicit information will be shown. This process continues until either a student answers correctly or a student gets all four mediating prompts. For the multiple-choice questions with five answers, there will be a total of four mediating prompts in each item.

According to Poehner and Lantolf (2013), they initially developed the C-DA of L2 reading and listening comprehension tests for Chinese, French, and Russian languages. The format used was multiple-choice questions with four distractors and one correct choice. The test questions and prompts were in English. The instructional prompts layered from most implicit to most explicit were provided to a learner if he or she did not answer correctly. In the mediation process, the first prompt presented if a learner answered incorrectly, *"Sorry, that's not the right answer."* Then, the C-DA provided text highlighting in color or with auditory repetition of the main text. If the second attempt failed, the text appeared, *"Sorry, that's not right either."* It then showed the narrower scope of the text. Each item followed this pattern. The explanation was shown if the correct response was selected in order to avoid uncertainty and guessing (Poehner & Lantolf, 2013). The example of mediation is presented in Figure 5.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

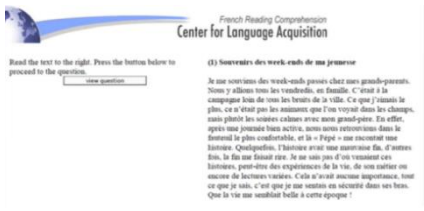
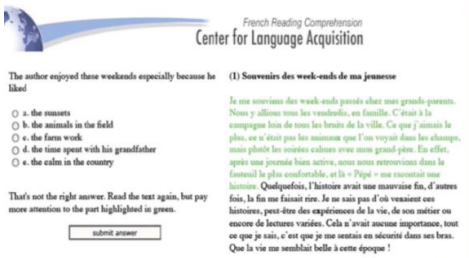
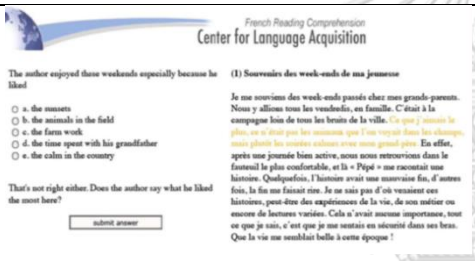
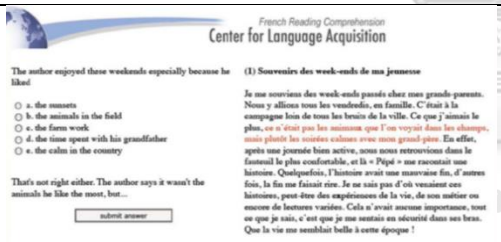
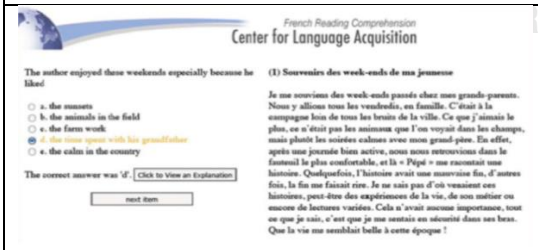
The screen	Mediation (prompts)
	
	If the first response is incorrect, learner received the mediating prompt which said, ***"that's not the right answer. Read the text again, but pay more attention to the part highlighted in green."*** This prompt is most implicit.
	If the second response is also incorrect, learner received another mediating prompt which said, ***"That's not right either. Does the author say what he liked the most here?"*** This prompt is more explicit than the first one.
	If the third response is incorrect, learner receives a more explicit prompt saying, ***"That's not right either. The author says it wasn't the animals he like the most, but…"***
	If the learner answers incorrectly, the fourth and most explicit prompt is revealed, which shows the correct answer.

Figure 5 Prompted items of a French reading comprehension test (adopted from Poehner & Lantolf, 2013, p.341)

2.2.2 Types of prompting

There are two types of prompting from related studies: the instructional prompting and error-explanation prompting

2.2.2.1 Instructional prompting

Instructional prompting can be used for both 'cognitive construct' and 'practical construct' (Poehner & Lantolf, 2013; Wang & Chen, 2016; Zhang et al., 2017). Each construct has differed in terms of the format used in computerized version, multiple-choice questions for cognitive constructs and open-ended questions for practical ones (Poehner & Lantolf, 2013; Wang & Chen, 2016; Zhange et al., 2017). The pattern of mediation will be as follows: 1st incorrect attempt ⟶ prompt 1 (most implicit), 2nd incorrect attempt ⟶ prompt 2 (more explicit), 3rd incorrect attempt ⟶ prompt 3 (very explicit), and 4th incorrect attempt ⟶ prompt 4 (most explicit, show the correct answer).

The study of Wang and Chen (2016) is cognitive-based assessment. The DARCST utilized a graduated prompts approach in the assessment procedure providing multiple-choice questions with four options. The assessment was administered online. Students could review the texts for re-reading and oral reading. It consisted of 14 text units with 73 follow-up items and 52 instructional prompts. The instructional prompts were designed to help students better understand the texts. There were three sequential instructional prompts as follows (Wang & Chen, 2016, p.378): "(1) ***IP1*** rephrase the question in an oral form; (2) ***IP2*** showed the reading strategy; and (3) ***IP3*** explained why the remaining choice was the correct answer derived from the appropriate reading strategy. Students clicked on the option. If the selected option was wrong, the IP1 appeared on the right side of the screen. Then, if the second response was wrong, the next IP2 appeared. If the third attempt was incorrect, the final IP3 was shown to represent the correct answer. The statements of IPs differed depending on the reading passages. The pattern of mediation was: 1st incorrect attempt ⟶ prompt 1 (rephrasing the question), 2nd incorrect attempt ⟶ prompt 2 (showing reading strategies), 3rd incorrect attempt ⟶ prompt 3 (providing a correct answer). The example of mediation is shown in Figure 6.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Do you agree with scientists' prediction that the neck of the Queen's Head rock will break and the rock will lose its head in the future?
(a) Agree, because the neck of the Queen's Head rock is being eroded and becoming too thin, and eventually cannot support the head. (b) Agree, because the gravity will pull the head of Queen's Head rock down. (c) Not agree, because the Queen's Head rock is very hard and cannot be broken down. (d) Not agree, because of erosion effects, the head will become smaller and stand on the top.
IP1: Will the neck of Queen's Head rock break down in the future? IP2: Please make a reasonable inference based on the statements in the fourth paragraph regarding the future crisis of the Queen's Head. IP3: The neck of Queen's Head rock is formed of sandstone, which contains only small grains. Without sediments of carbonate grains of marine organisms, the neck is very fragile and can be easily eroded by strong winds and ocean waves.

Figure 6 Prompted item (Wang & Chen, 2016, p.379)

Another example is from Teo (2012) who investigated the effect of computerized dynamic assessment on EFL students' inferential reading ability. The task was inferential reading. After students finished reading the passage, the questions with five-multiple choices were provided. If students answered incorrectly, four-level mediation, ranging from implicit to explicit was provided. The pattern of mediation was: 1st incorrect attempt ⟶ prompt 1 (most implicit), 2nd incorrect attempt ⟶ prompt 2 (more explicit), 3rd incorrect attempt ⟶ prompt 3 (very explicit), and 4th incorrect attempt ⟶ prompt 4 (most explicit, show the correct answer). The example of mediation is in Table 3 (Teo, 2012, p.14).

Table 3 Mediation of the inferential reading task (Teo, 2012)

Mediation Level	Details	Example
Level1: the most implicit mediation	• Explaining the meaning of inferential reading • Asking a general question to help the learners to find the main idea in the passage. • Providing the definitions of keywords in the passage	"To make inferences successfully, you need to read between the lines. This means that an inferential question cannot be answered by looking at the text itself. Instead, you will need to use the information stated in the passage to infer what is not stated. Your job is to make your best guess based on what you read. First, think about what the main idea of the passage is. The list below gives you the meanings of the keywords and places where you can possibly find the main ideas. Use the information to help you."
Level2: more explicit mediation	• Narrowing hints to guide learners focusing on certain paragraphs or sentences.	"Read **the third sentence through the last sentence** of the passage very carefully. The author is trying to point out something serious about the society at that time. **What is it?** "
Level3: very explicit mediation	• Focusing on ONE sentence, phrase, or word • Giving specific explanation	"Pay attention to the **second sentence**. The author said two things "bore little relation," So, ask yourself: according to the author, what are the two things that were not related to each other? Since these two things were not related, it caused some serious societal problems. Think about it, and then answer the inferential question.
Level4: providing the correct answer	• Providing the correct answer • Explaining the answer	"The word "critical" in the 5th sentence gives out a sense of urgency. The author used this word to express how serious the air pollution problem is. Although she did not say it directly in the passage, she was implying that the problem needed to be solved urgently because of its critical condition.

CHULALONGKORN UNIVERSITY

In terms of a practical construct, the work of Zhang et al. (2017) utilized the graduated prompting approach for computerized dynamic assessment. They assessed 2D graphic drawing in Graphing course. Students were assessed by drawing freely from their thoughts, no multiple-choice questions. If students encountered difficulty in performing the task, three levels of instructional prompting were provided, including abstract, semi-specific and specific. The abstract

prompt provided only text describing the graphic task. The semi-specific prompt added texts and numbers and the third prompt, which was most specific, described more details about the operational steps. The system provided the first prompt to a student if he or she was unable to perform the task. There was a five-minute lock-up for clicking the next prompt to prevent a student from taking a more explicit prompt without thinking. Then, if students still did not perform successfully, the next prompt was provided. A student could make a decision to use more prompts or move to the next item depending on their awareness of the picture or their understanding of the task. The pattern of mediation was: 1st facing difficulty ⟶ prompt 1 (abstract), 2nd facing difficulty ⟶ prompt 2 (semi-specific), and 3rd facing difficulty ⟶ prompt 3 (specific). The example is shown in Figure 7.

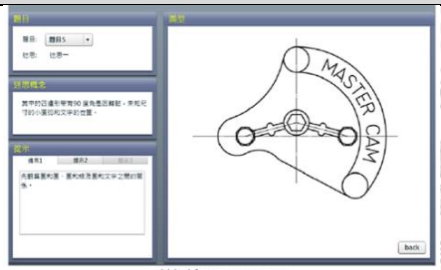
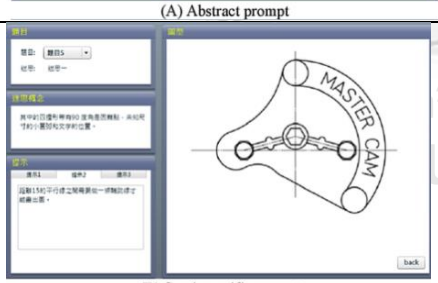
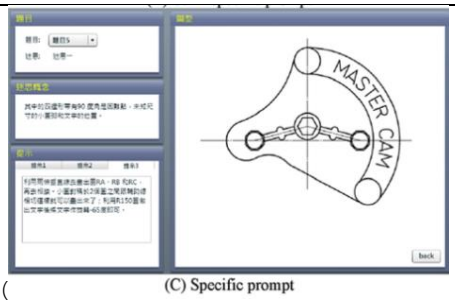
Stage	Screen	Feature	Example
Stage 1: Abstract prompt	 (A) Abstract prompt	Uses only text to describe	Observe the relations between circles, between circles and lines, and between letters.
Stage 2: Semi-specific prompt	 (B) Semi-specific prompt	Uses texts and numbers to describe it	a subsidiary line between the parallel lines with a distance of 15 is required for the circle, and the letter requires rotation.
Stage 3: Specific prompt	 ((C) Specific prompt	Uses texts, numbers, and operation steps to describe it.	Have a small circle symmetric to two circles and tangent to the subsidiary line for graph; draw a letter with an R150 circle and further rotate it 65

Figure 7 Abstract, semi-specific, and specific prompts (Zhang et al., 2017, p.187)

Table 4 Summary of instructional prompting from related studies

Study	Construct	Mediation Patterns
Poehner and Lantolf (2013)	Cognitive construct	1st incorrect attempt ⟶ prompt 1 (most implicit) 2nd incorrect attempt ⟶ prompt 2 (more explicit) 3rd incorrect attempt ⟶ prompt 3 (very explicit) 4th incorrect attempt ⟶ prompt 4 (most explicit, show the correct answer
Wang and Chen (2016)	Cognitive construct	1st incorrect attempt ⟶ prompt 1 (rephrasing the question) 2nd incorrect attempt ⟶ prompt 2 (showing reading strategies) 3rd incorrect attempt ⟶ prompt 3 (providing correct answer)
Teo (2012)	Cognitive construct	1st incorrect attempt ⟶ prompt 1 (most implicit) 2nd incorrect attempt ⟶ prompt 2 (more explicit) 3rd incorrect attempt ⟶ prompt 3 (very explicit) 4thincorrect attempt ⟶ prompt 4 (most explicit, show the correct answer)
Zhang et al. (2017)	Practical construct	1st facing difficulty ⟶ prompt 1 (abstract) 2nd facing difficulty ⟶ prompt 2 (semi-specific) 3rd facing difficulty ⟶ prompt 3 (specific)

In summary, Table 4 addresses the instructional prompting used on studies related to computerized dynamic assessment to help students better understand the texts. Studies emphasized the cognitive constructs used multiple-choice questions. If a student answered incorrectly, the provided prompts ranged from most implicit to most explicit. On the other hand, the study focused on the practical constructs used for practical task with no multiple-choice questions. Mediation was provided when students had difficulty in performing the task. The mediation was the prompts ranging from abstract to specific.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.2.2.2 Error-explanation prompting

The system provides different prompts in accordance with selected options. When a student chooses the incorrect answer, the prompt will be provided depending on the number of times the student answers incorrectly and the error from his or her selected response. Each distractor provides different prompts to emphasize the error he or she chooses ranging from implicit to explicit (Ting & Kuo, 2016). Error-explanation prompts can be formatted as follows: if the correct answer is choice C, a student's 1st attempt answers incorrect choice A⟶ most implicit prompt 1 from choice A, a student's 2nd attempt answers incorrect choice B⟶ explicit prompt 2 from choice B, a student's 3rd attempt answers incorrect choice D⟶more explicit prompt 3 from choice D, and a student's 4th attempt answers incorrect choice E⟶show correct answer

Wu et al. (2017) studied the computerized dynamic error-explanation prompting methods compared with other two methods, namely the instructional prompting dynamic assessment and the traditional classroom instruction. They mentioned that this type of prompt, error-explanation prompting, not only provided individualized instruction to meet students' needs but also assisted students in learning mathematics. Figure 8 showed the general prompt popping up in the interface which occurred when a student responded incorrectly for the first time. It said, "The denominators of these two fractions are different. Use reduction to a common denominator so that we can add or subtract."

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

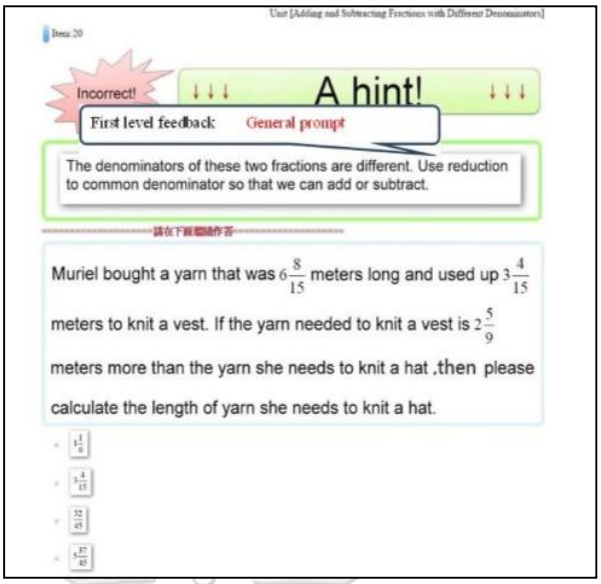


Figure 8 Prompting interface (Wu, Kuo, & Wang, 2017, p.65)

In Table 5 , prompts will be provided based on a number of times each student responded incorrectly and the distractor they chose. For example, if Adam answers choice A for the first time, he will be presented with the popping up prompt, *"Incorrectly! Your answer is 40÷5 = 8. What is an expansion?"* Then, Adam chooses choice D for the second time; prompts showed, *"Incorrectly! Expansion of a fraction is not the numerator minus the denominator. How to expand the denominator?"* For the last time, Adam decides to choose choice B which is also an incorrect answer. The correct answer is provided, *"The expansion refers to multiplying the numerator and denominator of a fraction by the same (bigger than 1) integer, which results in a fraction that is equivalent to the original fraction. Multiplying the numerator and denominator of* $\frac{4}{5}$ *by 8, that is,* $\frac{4}{5} = \frac{4 \times 8}{5 \times 8} = \frac{32}{40}$*, the number in ( ) is 32.*

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 5 Error-explanation prompts (Wu, Kuo, & Wang, 2017)

Prompt	**Item1: 4/5 = ( )/40, what is the number in ( )?**			
	(A) 8	(B) 10	**(C) 32**	(D) 39
Without instructional prompt			Great! Respond correctly	
First level prompt: general prompt	Incorrectly! Your answer is 40÷5 = 8. What is expansion?	Incorrectly! Your answer is 40÷4 = 10. What is expansion?		Incorrectly! Your answer is 40-1 = 39. What is expansion?
Second level prompt: keywords, concrete prompt	Incorrectly! The numerator is 5 ✕ 8 = 40 How to expand the dominator?	Incorrectly! The numerator is 5 ✕ 8 = 40 How to expand the dominator?		Incorrectly! Expansion of a fraction is not the numerator minus the denominator. How to expand the denominator?
Third level prompt: direct instruction	The expansion refers to multiplying the numerator and denominator of a fraction by the same (bigger than 1) integer, which results in a fraction that is equivalent to the original fraction. Multiplying the numerator and denominator of $\frac{4}{5}$ by 8, that is, $\frac{4}{5} = \frac{4 \times 8}{5 \times 8} = \frac{32}{40}$, the number in ( ) is 32			

In the study of Ting and Kuo (2016) , they utilized error-explanation prompting into their studies of finding an area using an integral by comparing five prompting conditions, three conditions of dynamic assessment mediation and two conditions of static assessment. When a student answered an item incorrectly, he or she would receive the appropriate prompt depending on the error pattern that corresponded to his selected response, ranging from implicit to explicit. The example was as the following: The question was *'Find the area enclosed by the left line X = -3 and the right parabola* $y^2 = -x + 1$*'* and the prompt based on the error pattern corresponding to selected choice was provided. For example, choice A, prompt 1 was 'Does the left integrator or the right integrator contain an error?' and prompt 2 was *'Please note that your right integrator was x =*

y-1 and your left integrator was x = y^2-1. Did you do something wrong?'. Then, a student will be allowed to make another choice again.

In summary, from the previous studies, there were two types of prompting, including instructional and error-explanation promptings. However, this study combined the advantages of both instructional prompting and error-explanation prompting called mixed prompting. Three different promptings were utilized in my study to compare the effectiveness of the prompting methods and measured the learning potential of students from different treatment groups. Each prompting benefitted students and examiners in different ways. Instructional prompting helped a student to better understand the text and the question. For the error-explanation prompting, students' error or misunderstanding was disclosed. The mixed prompting explored the content and helped students to realize their misunderstanding.

2.3 Scoring procedures

The system of computerized dynamic assessment can present both scores and profiles. The scoring procedure of computerized dynamic assessment in this review is based on the study of Poehner and Lantolf (2013). In addition, the profile can elaborate not only the individual profile but also the group profiles.

2.3.1 Scoring

Poehner and Lantolf (2013) developed the C-DA as diagnostic assessment and reported the C-DA scores as actual score, mediated score, gain score, and learning potential score (LPS) for each learner. A learner who had the same actual score might have different mediated scores. Figure 9 represents the summary of scores, except the transfer scores is as follows:

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Learner	Actual score*	Mediated score	Gain score	LPS
1	12	42	30	.78
2	12	51	39	.98
3	12	63	51	1.24
4	92	92	0	1.00
5	80	89	9	1.07
6	12	46	34	.87
7	24	52	28	.87
8	24	52	28	.87
9	88	91	3	1.02
10	88	90	2	1.00

Note: Maximum score = 92.

Figure 9 Scores of the Chinese listening test (Poehner, Zhang, & Lu, 2015)

2.3.1.1 Actual score

The score represented a learner performance without mediation or independent performance or the developed ability. It showed nothing about the ZPD. The actual scores ranged from either 4 or 0. If the first response was correct, a learner obtained a score of 4. If a learner answered it incorrectly, he or she received a score of 0 (Poehner et al., 2015).

2.3.1.2 Mediated score

The score represented the total number of mediating prompts a learner used during the test administration. Mediated scores represented how well a learner responded to mediation, but the scores did not reflect learning. The scores ranged from 0 to 4. It was reduced by 1 from each mediating prompt a learner received (Poehner et al., 2015).

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

The study from Poehner et al. (2015) found that two learners who obtained the same actual scores did not receive the same mediated scores. The actual and mediated scores could be compared to show the abilities of learners after the mediation. It was found that learners with high actual scores also had higher mediated scores. Some learners were profitable from mediation, whereas some had less improvement which might be because of a ceiling effect, individuals who got high pretest scores had less room for improvement, or floor effect, individuals who had taken the test beyond their current development level (Lantolf & Poehner, 2004).

2.3.1.3 Learning potential score (LPS)

Previously, Budoff categorized learners depending on pre-training scores, post-training scores, and post-training score adjusted for pretest level. His viewpoint of the learner's performance was based on the difference between pretest and posttest. Poehner and Lantolf (2013) argued that it represents test analysis rather than child analysis. Kozulin and Garb (2002) proposed a formula for calculating LPS in order to capture the degree of change between independent performance and mediated performance in a single score. It measured openness to mediation. Also, it could distinguish between those students who had high learning potential and those who had low learning potential. Following Poehner et al. (2015), the learning potential score was used to predict how well learners responded to instruction in the future. It represented the maximum possible score.

$$\text{LPS} = \frac{(\text{S post} - \text{S pre})}{\text{Max S}} + \frac{\text{S post}}{\text{Max S}} = (2\ \text{S post} - \text{S pre})/\text{Max S}$$

Figure 10 Learning potential score formula (Kozulin & Grab, 2002)

Kozulin and Garb (2002) divided learners into three subgroups based on their LPS scores: "High scorers (LPS ≥ 1.0); Mid-range scorers (.99 ≥ LPS ≤ .72) and Low scorers (LPS ≤ .71)". An LPS of 1 or higher represented the score improved by one standard deviation; on contrary, An LPS under 1 represented a change of less than one standard deviation (Poehner & Lantolf, 2013). Learners in different groups required different supports to enhance their abilities.

Following the study of Poehner et al. (2015), the LPS of learners ranged from .59 to 1.24 in the listening test and from .77 to 1.19 in the reading test. The interesting findings were learners with the same actual scores had different mediated scores and LPS scores. If LPS scores represented the learners' zone of proximal development, learners who had the same independent performance but different LPS scores might require different levels of support in the classroom in the future.

2.3.2 Profiling

The learner profile had been part of studies of C-DA (Poehner et al., 2015; Teo, 2012; Wang & Chen, 2016). It emphasized on independent functioning of learners and responsiveness to mediation. The results from C-DA could be used to generate a learner profile which depicted particular constructs of a learner to show both strengths and weaknesses of their performances as well as to inform instructors about specific areas that needed remedy (Poehner et al., 2015; Teo, 2012). Teo (2012) said that C-DA helped her to capture her students' reading ability by determining the number of attempts and level of prompts. Poehner et al. (2015) showed that learners who had the same actual score, mediated score, and LPS might have different areas of potential or weakness. The learner profile could be formatted in both a group profile and individual profile. A group profile in Figure 12 depicted the number of items, a wide range of constructs and scores. An individual profile in Figure 11 described the scores of each individual on each construct and the mediation learner required or a number of prompts used in each item to emphasize the difficulty of the learner.

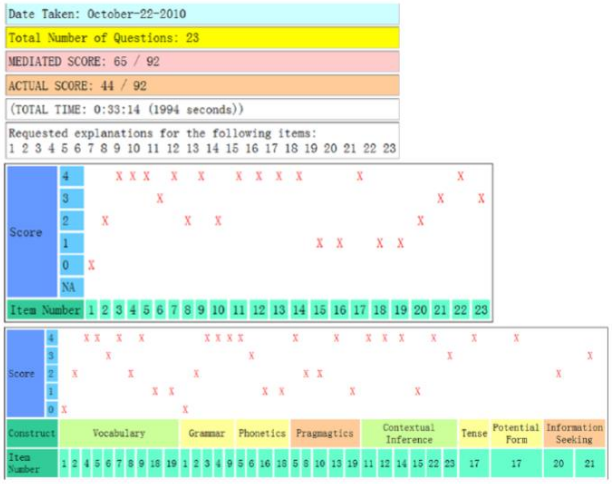


Figure 11 Learning profile of the C-DA Chinese listening test (Poehner, Zhang, & Lu, 2015, p.356)

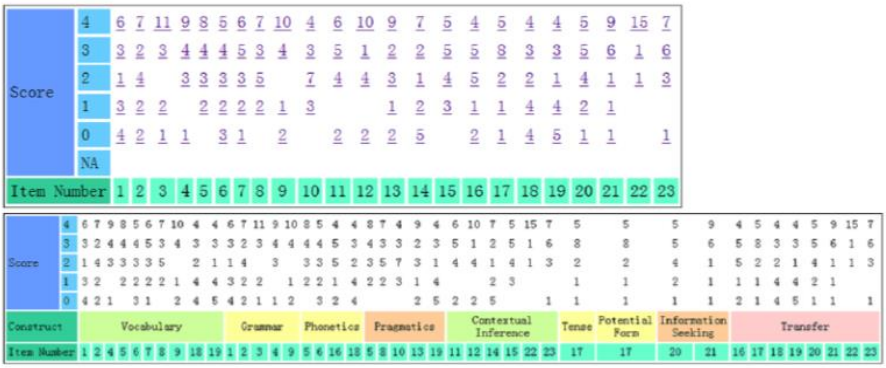


Figure 12 Group profile of the C-DA Chinese listening test (Poehner, Zhang, & Lu, 2015, p.357)

2.4 The application of the two-tier test

This study utilized multiple-choice questions as the measurement of learning potential. A multiple-choice question was the form of assessment that students were asked to respond to one best answer from a list. There were many advantages of using this kind of tests such as validity and reliability, easy scoring, and large-scale administration(Gurel et al., 2015). The validity of the test result is important in educational measurement. One of the concerns of measurement is guessing, which can be considered as error variance and does not reveal the true ability of test-takers. It also reduces the reliability of the test (Gurel et al., 2015; Medawela et al., 2018). According to Yang et al. (2017), they mention the possibility of students' partial understanding or misunderstanding in answering multiple-choice questions. Students who answered correctly may have partial knowledge or even misunderstand the relevant concept. In order to circumvent guessing and examine students' true ability, the two-tier test was applied in this study.

The two-tier test has been widely used for diagnostic testing approaches, which can diagnose students' misconceptions. The multiple-choice questions cannot provide insight into students' idea or reason behind their chosen choice (Gurel et al., 2015). The advantage of this approach is in diagnosing learning problems and providing assistance to overcome the problems (Yang et al., 2017). Two-tier testing comprises a set of two multiple-choice questions. The assessment of the first tier is on students' factual or relevant concept or knowledge, also called the phenomenon

tier, while the second tier, which is called the reason tier, is on the rationale behind the selected choice as shown in Figure 13.

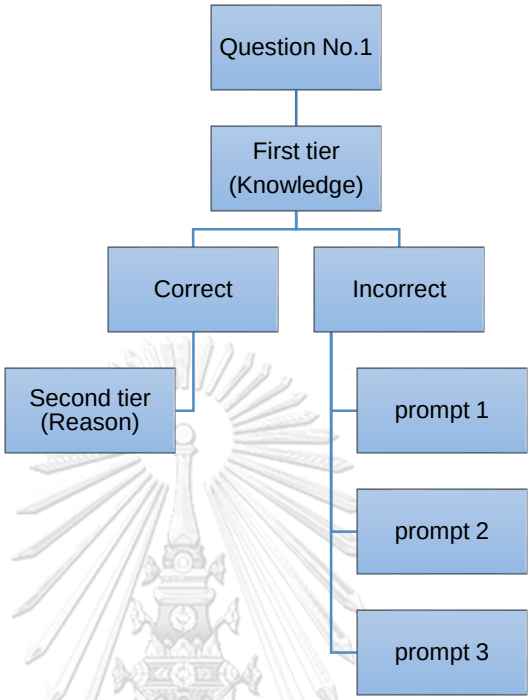


Figure 13 Two-tier testing process adapted in this study

This concept provides a deeper insight in examining students' understanding than traditional multiple-choice tests (Gurel et al., 2015; Lin, 2016). Students' correctly succeeded answer will combine both correct answer choice and reason (Gurel et al., 2015). If students cannot answer correctly, the system can identify the misconception and the reason caused the misconception (Kilic & Saglam, 2009; Yang et al., 2017). With regard to Tamir (1989, as cited in Kilic & Saglam, 2009), the selection of the rationale in two-tier tests is an effective way to measure learning purposefully. The concept of two-tier testing is widely renowned because of its time-efficient, labor-saving, and generalization purposes (Kilic & Saglam, 2009; Lin, 2016).

The two-tier test is more practical and convenient to use for avoiding guessing (Gurel et al., 2015). It has been reported to reduce the chance of guessing. In the task of four answer choices in the multiple-choice questions, students have 25% chance of getting the correct answer from guessing, but it is reduced to only 6.25% when adding the second-tier item (Milenkovic et al., 2016). Milenkovic et al. (2016)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

employed the 15 two-tier items, each comprising four answer choices for high school students and reported the positive results in eliminating the guessing correct answer and the misconceptions of the students.

Yang et al. (2017) employed second-tier items in diagnostic and guiding learning to assess mathematics of undergraduate students as shown in Figure 14. Students' factual knowledge of Calculus was assessed in the first-tier items, whereas the second-tier items revealed students' reasons for their choice chosen in the first tier. Each item comprised three choices. They reported the positive impacts of this approach on correcting students' misunderstanding and guiding students to improve their skills. In addition, Kilic and Saglam (2009) employed two-tier multiple-choice tests on genetics concepts for secondary education students. The first-tier items composed three answer choices, whereas the second-tier items had five optional choices. They also reported the positive two-tier approach in diagnosing students' understanding of the concepts of genetics.

Illustrative example of the second tier options (i.e., identifying the reason for the answer)

Second-tier item	Choices
Why did you choose 3a3-3cos2x sin x?	(1) The derivative of a constant is the same rule as for functions.
	(2) The derivative of a constant is zero.
	(3) The derivative of a constant is still a constant.
Why did you choose 3cos2x?	(1) The derivative of x3 is 3x2
	(2) The derivative of a constant is zero.
	(3) The derivative of Y3 is 3Y2Y".
Why did you choose -3 cos2x sin x?	(1) Because of the chain rule.
	(2) Because (cos 3x)'=3 cos2x
	(3) Because (cos 3x)'=3 cos2x sin x

Chulalongkorn University

Figure 14 Second-tier options (Yang et al., 2017, p.65)

To sum up, in the current study, the two-tier test had been applied to the prevention of guessing. The concept of a two-tier test in the ongoing study had slightly been changed from the original concept. In this test, there was a set of two items. The first-tier item comprised multiple-choice questions with five choices, whereas the second-tier item consisted of three rational choices. Students who answered the first tier correctly were moved to answer the second tier automatically. On the other hand, if some students could not answer the first tier correctly, they would be provided with the

graduated promptings until they responded to the first-tier item correctly and were allowed to move to subsequent items.

2.5 Relevant studies on computerized dynamic assessment (C-DA)

There are numerous findings on computerized dynamic assessment. Some studies have mentioned the performance change. Others have addressed the growth of learning while a few have paid attention to learning potential.

2.5.1 C-DA on performance change

Several studies have found that students increased their posttest scores after the intervention (Ku et al., 2014; Teo, 2012). Ku et al. (2014) developed a computerized dynamic assessment of Accounting subject to measure vocational students. A paired t-test was used to analyze learning performance between pretest and posttest. The computerized dynamic assessment lasted for three weeks. The results indicated that students had higher gain from posttest scores in all three groups (high, medium, and low scores). After using the computerized dynamic assessment, students had better understand of the concepts and improved their learning outcomes. Teo (2012) investigated the progress of students' inferential reading skills before and after receiving the C-DA program by using a paired samples t-test, she found that the participants' performance scores were significantly higher on the posttest than pretest. The participants had metacognition and monitored their reading process actively.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Research showed that the dynamic assessment group outperformed control group or other groups (Stevenson et al., 2016; Wu et al., 2017). The study of Wu et al. (2017) used a pair sample t-test to investigate the student improvement from the three remedial instruction programs, including the direct instruction group, the instructional prompting group, and the error-explanation prompting group. The results showed that the posttest scores were higher than the pretest scores for all three groups which demonstrated that students showed the improvement. The analysis of ANCOVA was used to compare the effectiveness of the treatments. It was found that students with error-explanation prompting group outperformed other groups in learning mathematics. Stevenson et al. (2016) applied the Multidimensional

Rasch model for learning and change (MRMLC) to estimate performance change scores of students. They reported that students improved from the intervention. It was found that those who received the graduated prompts obtained higher gain from pretest to posttest than those who were assigned to practice and attention-control groups. Also, performance change between native and ethnic minority children did not differ after the intervention.

2.5.2 C-DA on growth of learning

Regarding research on growth, DA studies on measurement of growth found that growth of student performance varied and showed different changes over time (Wang & Chen, 2016; Zhang et al., 2017). Wang and Chen (2016) used item response theory (IRT) and a two-way repeated measures analysis of variance to investigate the reading comprehension growth of students through the web-based dynamic assessment. It revealed that the instrument was suitable for most students with a variety of reading abilities. The test information function (TIF) represented that the test was appropriate for students whose ability ranged from -1.08 to 3.0. When analyzing growth, time and initial reading ability (high-, average-, and low-level students) will be used to measure the group of reading ability. It was found that the web-based dynamic assessment had significant effects on the changes or growth in students' reading ability. The results showed that 53% of growth in reading ability contributed by the repeated web-based dynamic assessment. Moreover, students with different initial levels of reading comprehension improved differently. The high-level students dropped from week 1 to week 2 and then had steady growth from week 2 to week 4. On the contrary, low-level students showed growth dramatically from week 1 to week 4. Average-level students grew steadily from week 1 to week 4. Therefore, low-level students benefited the most from the web-based dynamic assessment. Zhang et al. (2017) investigated longitudinal study of computer-based graduated prompting assessment on academic performance over time by testing the growth model of hierarchical linear modelling (HLM). The results showed that students who used the computer-based graduated prompting achieved higher academic performance than those who did not. It found out that prompts provided

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

in a graduated way could improve students' academic performance in 2D graphing. Regarding the effect of graduated prompting over time, although the growth of the experiment group was higher than the control group, the rate didn't reach a significant level. The academic performances of freshmen during three years did not have a higher growth over time compared to the assessment without a computer. Thus, the C-DA did not lead to change over time.

2.5.3 C-DA on learning potential

As regards learning potential, research on dynamic assessment has computed learning potential in different ways (Poehner et al., 2015). Poehner et al. (2015) developed the computerized dynamic testing of L2 Chinese reading and listening comprehension. They reported students' learning potential by calculating a learning potential score for each learner. In their term, learning potential is an openness to mediation that provides insights into the degree of instructional effort required to help learners advance. The pilot study reported that students who had the same current ability in Chinese language had different learning potential scores. Thus, it could be believed that even though students had similar current Chinese skills, they required very different levels of instructional support in a Chinese language class in the future.

It was found that both control and trained groups showed progression patterns from pretest to posttest; however, trained students outperformed control-group students on a number of corrected series completion items. Moreover, they paid attention to the number of prompts needed for overtraining as a good operationalization of student potential for learning. It was found a significant decrease in the number of prompts needed to solve the questions from training 1 to training 2. Students required fewer prompts at the second training. Thus, it indicated that students profited from the prompts especially when taking the more difficult items. Consequently, as shown in Table 6, most studies related to dynamic assessment have studied the performance change or the current performance of the students. Some studies have focused on the growth of learning and few research has examined learning potential.

Table 6 Summary of previous studies on computerized dynamic assessment

Study	Participants	Course	Duration for intervention	Research method	Treatment group	DA format	Learning potential indices	Data analysis
Ku, Shih, & Hung (2014)	34 2nd-year vocational students	Accounting	3 weeks	Experiment (pretest, intervention, and posttest)	One group	Sandwich	Performance change	A paired sample t-test
Poehner, Zhang, & Lu (2015)	68 students for listening test and 82 students for reading test	L2 listening and reading Chinese language	-	Research and development		Cake (graduated prompts)	Learning potential	Learning potential scores
Ting & Kuo (2016)	257 freshmen	Calculus	2 weeks	Experiment (pretest, intervention, and posttest)	Five treatment groups	Sandwich (Pretest/ posttest)	Performance change	One-way ANCOVA
Teo (2012)	68 university EFL students	Inferential reading	8 weeks	Experiment (pretest, intervention, and posttest)	One treatment group	Sandwich (pretest /posttest)	Performance change	A paired samples t-test
Wang and Chen (2016)	213 students	Science	4 weeks	Experiment (initial and 3-week follow-up)	Three groups (initial levels of reading ability)	Cake (graduated prompts)	Pattern of growth	A two-way repeated measures analysis of variance (ANOVA)
Wu, Kuo, & Wang (2017)	118 Grade 5 students	Mathematics	1 period	Experiment (pretest, intervention, and posttest)	Three groups (direct instruction, fixed, mixed)	Sandwich (pretest and posttest)	Performance change	A paired sample t-test and analysis of covariance (ANCOVA)
Zhang et al. (2017)	120 freshmen students	Graphing	Three years (Each cohort for 12 weeks)	Experiment (pretest, intervention, and posttest)	two treatments	Sandwich (pretest and posttest)	Performance change	Hierarchical linear modeling (HLM)
Resing & Elliott (2011)	77 children (mean age 8.9 years)	Inductive reasoning	2 periods	Experiment (pretest, intervention, posttest)	Two treatments	Sandwich	Performance change and learning potential	Repeated measures analysis of ANOVA
Stevenson, Heiser, & Resing (2016)	111 children (56 Native Dutch and 55 ethnic minorities)	Reasoning	1 period	Experiment (pretest, intervention, and posttest)	Three treatments	Sandwich (pretest and posttest)	Performance change	Multidimensional Rasch Model for Learning and Change (MRMLC)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Part 3: Reading Literacy

Reading literacy is the foundation of other subject areas in the educational system; it is an essential skill needed for individual growth, educational and career success, and global citizenship. The attainment in reading literacy successfully leads to the prerequisite for participation in real-life situations in the rapidly changing world (OECD, 2016). Reading literacy is one of the major three domains measured in the Programme for International Student Assessment (PISA), an international survey aimed to evaluate international educational systems. PISA has been developed from the Organization for Economic Co-operation and Development (OECD). The adolescents, 15-year-old students, are assessed for their scholastic performances. The tested contents are not based on the basic curriculum or what is being taught in the classroom, but the wide range of contents necessary for the real-life situations that students may experience in the future are tested. The assessment started in 2000 and then has continued every three years (The Institute for the Promotion of Teaching Science and Technology, 2018).

3.1 Definition of reading literacy

OECD (2016) has defined reading literacy in PISA 2015 as the competency to understand, to use, and to reflect on written texts for developing knowledge and potential to achieve goals and participate in the society. From then on, the term has been updated and extended due to the rapidly changing society. The tomorrow students will be required to use digital tools to succeed with complex and massive information. Students are required to achieve an expanding set of knowledge and skills in various contexts. Reading literacy in 2018 is defined as, “understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one’s goals, to develop one’s knowledge and potential and to participate in society” (OECD, 2016, p.8). The new term has been added is “evaluation” of texts which is considered as an indispensable part of reading that learners should have the ability to investigate the arguments in the text and discover the point of view of the author. Moreover, the “written” has been removed because of the included both fixed and dynamic texts. When considering reading literacy for PIRLS2016, it has been grounded by International Association for Evaluation of Educational Achievement (IEA)’s 1991

study, defining reading literacy as “the ability to understand and use those written language forms required by society and/or valued by an individual.”(Mullis et al., 2015, p.11). PIRLS focuses on the ability to read, which extends to the ability to apply what is read to the new situations or contexts (Mullis et al., 2015). Furthermore, Diowvilai et al. (2012) have defined reading literacy as the ability to retrieve texts in the form of written language, pictures, or symbols, to identify the components of text, and to evaluate the reliability of the text.

Therefore, reading literacy is the constructive process involving interaction between the reader and the texts with the intention of reading for growth and for participation in society. It measures how students understand the text, interpret the meaning of the text, evaluate the text and apply their reading ability into their real-life situations.

3.2 Dimensions of reading literacy

Reading literacy has been categorized into different ways depending on assessment programs or research studies as the following:

3.2.1 PISA 2018

Based on the framework of PISA 2018, it has been adjustably focused on finding, selecting, interpreting, integrating and evaluating information from the full range of texts beyond the classroom context. The emphasis is on the skills required for reading and interacting with non-digital and digital texts in various contexts. The term “cognitive process” has been replaced from “cognitive aspects”, which was used in the previous assessment cycle. In the older version of PISA reading literacy, there were three aspects of reading literacy test including (1) Access and retrieve, learners are required to access and retrieve a particular piece of stated information; (2) Integrate and interpret, learners are required to integrate and interpret on what to read by forming a broad understanding and developing an interpretation; and (3) Evaluate and reflect, learners are required to evaluate and reflect beyond the text in which learner could do on both content of the text and form of the text (OECD, 2016; Thomson et al., 2013). The differences between PISA 2009-2015 and PISA 2018 are shown in Table 7.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 7 The PISA 2009-2015 and PISA 2018 reading framework processes (as adapted from OECD, 2016, p. 13; Thomson, Hillman, & Bortoli, 2013, p. 9)

Reading Literacy PISA 2009-2015		Reading Literacy PISA 2018	
Access and Retrieve	Retrieve information	Locate Information	Access and retrieve information within a text
			Search and select relevant text
Integrate and Interpret	Form a broad understanding	Understand	Represent literal meaning
	Develop an interpretation		Integrate and generate inferences
Reflect and Evaluate	Reflect on and evaluate content of text	Evaluate and Reflect	Assess quality and credibility
	Reflect on and evaluate form of text		Reflect on content and form
			Detect and handle conflict

For the newer version, the processing has been categorized into 3 processes including 1) locate information, 2) understand, and 3) evaluate and reflect (OECD, 2016).

1) Locate Information

Locating information is an important process when a learner uses texts for specific purposes on a daily basis with little or no time to consider the rest of the text. It is the learner's awareness of information needs. There are two main tasks in locating information which is: (1) Access and retrieve information within a text, learners are required to scan a single text for retrieving target information in the form of a few words, phrases or numerical values with no need to comprehend the rest of the text and (2) Search and select relevant text, learners are required to search for information among several texts to select the most relevant text information required for the task by using the help of text descriptors such as headers or source information (e.g. author, medium, date).

2) Understand

Reading involves the understanding of the meaning conveyed in the passage or what the text is about. The two tasks are provided including (1) Representing literal meaning, learners are required to comprehend the literal meaning of sentences or short passages by matching a direct or paraphrasing information in the question with information in a passage and (2) Integrating and

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

generating inferences, in a single text, learners are required to integrate information across sentences or an entire passage by identifying the main idea, summarizing, creating title name. Integrating and generating inferences for multiple texts requires learners to integrate pieces of information located within two or more texts.

3) Evaluate and Reflect

Reading involves reflecting on the content and form of the text as well as assessing the quality and validity of the information provided. This topic is composed of three tasks as follows: (1) Assessing quality and credibility, learners are required to evaluate the quality and credibility of the text whether the information is valid, accurate, unbiased, and reliable. Moreover, the reader is required to identify the source of information, the content, and form of the text or how the author is presenting the information; (2) Reflecting on content and form, learners are required to reflect on the quality and style of the writing such as evaluating the form of writing and how contents are connected that determining the author's purposes and point of view. Students are also required to reflect on their own knowledge and experience to compare, contrast or hypothesize different viewpoints; and (3) Detecting and handling conflict, learners are required to deal with conflict in case of facing multiple texts that contradict each other by identifying each author's stance and selecting one source that has better soundness of the claims and/or more credibility of the source.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.2.2 PIRLS 2016

Progress in International Reading Literacy Study (PIRLS) is organized to measure reading achievement of young students in their fourth year of schooling to provide internationally comparative data on how well children read. PIRLS first assessment was in 2001 and continued every 5 years. Reading literacy framework was based on IEA's Reading Literacy Study. Reading literacy framework for PIRLS 2016, the fourth assessment cycle, is composed of four processes: 1) Focus on and retrieve explicitly stated information, 2) Make straightforward inferences, 3) Interpret and integrate ideas and information, and 4) Evaluate and critique content and textual elements (Mullis et al., 2015) as shown in Table 8.

Table 8 PIRLS 2016 reading literacy framework (Mullis, Martin, & Sainsbury, 2015)

PIRLS literacy	
Focus on and retrieve explicitly stated information	▪ Identifying information that is relevant to the specific goal of reading ▪ Looking for specific ideas ▪ Searching for definitions of words or phrases ▪ Identifying the setting of a story (e.g., time and place) ▪ Finding the topic sentence or main idea (when explicitly stated)
Make straightforward inferences	▪ Inferring that one event caused another event ▪ Concluding what is the main point made by a series of arguments ▪ Identifying generalizations made in the text ▪ Describing the relationship between two characters
Interpret and integrate ideas and information	▪ Discerning the overall message or theme of a text ▪ Considering an alternative to actions of characters ▪ Comparing and contrasting text information ▪ Inferring a story's mood or tone ▪ Interpreting a real-world application of text information
Evaluate and critique content and textual elements	▪ Judging the completeness or clarity of information in the text ▪ Evaluating the likelihood that the events described could really happen ▪ Evaluating how likely an author's argument would be to change what people think and do ▪ Judging how well the title of the text reflects the main theme ▪ Describing the effect of language features, such as metaphors or tone ▪ Determining an author's perspective on the central topic

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.2.3 NAEP

National Assessment of Educational Progress (NAEP) is a mandated project administered by the National Center for Education Statistics (NCES) to measure what students know and can do in their academic achievement. There have been four main subject areas including mathematics, reading, science, and writing. The assessment has been conducted with Grade 4, Grade 8, and Grade 12 American students. NAEP reading framework has been guided by scientifically-based reading research. Reading has been defined as the cognitive process involving to understand

written texts, to develop and interpret meaning, and to use meaning that match with appropriate text types. There has been three reading cognitive targets of NAEP as the following: 1) Locate and recall, 2) Integrate and interpret, and 3) Critique and evaluate (National Assessment Governing Board, 2017) as shown in Table 9.

Table 9 NAEP reading framework (National Assessment Governing Board, 2017)

NAEP Reading Framework	
Locate and Recall	▪ Identify explicitly stated information ▪ Focus on specific elements of a story
Integrate and Interpret	▪ Make complex inferences within and across texts ▪ Explain character motivation ▪ Infer the main idea of an article ▪ Infer and explain the theme of a story
Critique and Evaluate	▪ Consider the text critically by viewing it from numerous perspectives ▪ Evaluate overall text quality or the effectiveness of particular aspects of the text

3.2.4 Thai research

In the context of Thailand, reading literacy has been categorized into five levels, namely decoding, reading fluency, reading comprehension, analytical reading and interpretation, and critical reading (Diowvilai et al., 2012). Five levels of reading literacy are the following:

1) Level one: decoding

Decoding skill is important for a reader to start learning to read. It is the ability of a reader to recognize and apply letter-sound relationships to pronounce written words correctly. A reader should initially have knowledge of letter patterns and should recognize the basic sounds, as well as try to pronounce the word correctly.

2) Level two: reading fluency

Fluency is the ability of a reader to read a text accurately, automatically, and confidently. Their reading tones are natural with the proper speed

and their tones stick to emphasize certain words so that it is easy for the listener to understand what they are reading.

3) Level three: reading comprehension

Reading comprehension is the ability of a reader to comprehend the text and to access the information in the forms of words, sentences, phrases. Moreover, it is the ability to locate the specific information about the five W's (who, what, where, when and why) and identify the main idea of the text.

4) Level four: analytical reading and interpretation

Analytical reading and interpretation is the ability of a reader to identify the components of the texts, make inferences and interpret the meaning of the text, both the overt and hidden messages. The examples of this skill are the following: Identifying facts and opinions within the texts, identifying cause-and-effect relationship within the texts, determining the purpose of the text, determining the viewpoint of the author or what the author trying to say, making prediction while reading, and interpreting author's attitude of writing.

5) Level five: critical reading

Critical reading is the ability of a reader to evaluate and make a judgment about the strengths and weaknesses of the text. A reader should be able to make a judgment by examining evidence given to support the arguments, whether the text is accurate or unbiased. Moreover, a reader needs to form the evaluation based on the judgment he/she made and provide opinions and reasons supporting agreement or disagreement of the text.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Moreover, Praputtakun et al. (2013) have categorized reading literacy into two aspects: retrieving information and interpretation. (1) Retrieving information is used to assess the ability of the reader in identifying the specific information within the text to answer the five W's and one H (who, what, where, when, why, and how). This aspect represents a reader's ability in remembering and understanding. (2) Interpretation is used to assess a reader's ability in interpreting the meaning within the text, which represents a reader's ability in applying the knowledge.

In sum, the reading literacy frameworks of both international projects and Thai research are shown in Table 10 and Table 11. It has been revealed that they have

similar dimensions in terms of 1) the ability to locate and receive information, 2) the ability to understand, integrate and generate inferences, and 3) the ability to evaluate and reflect information. In this way, the current study adopted the reading literacy framework of PISA 2018 as a core framework for developing the reading literacy tests due to its up-to-date framework emphasizing online reading literacy as well as the inclusion of competencies to measure the ability to read multiple texts. Moreover, PISA was considered as one of the indicators for the quality improvement of the educational system in Thailand.

Table 10 Reading literacy dimensions among international programs and Thai research

Reading literacy					
International				Thai	
PISA 2009-2015 (OECD,2016)	PISA2018 (OECD, 2016)	PIRLS	NAEP	Diowvilai et al. (2012)	Praputtakun et al. (2013)
▪ Access and retrieve ▪ Integrate and interpret ▪ Reflect and evaluate	▪ Locate information ▪ Understand ▪ Evaluate and reflect	▪ Focus on and retrieve information ▪ Make inferences ▪ Interpret and integrate ▪ Evaluate and critique	▪ Locate and recall ▪ Integrate and interpret ▪ Critique and evaluate	▪ Decoding ▪ Reading fluency ▪ Reading comprehension ▪ Analytical reading and interpretation ▪ Critical reading	▪ Retrieving information ▪ Interpreting

Table 11 Similarity and discrepancy on reading literacy between PISA2018 and the others

	International			Thai	
PISA 2018	**PISA2009-2015**	**PIRLS**	**NAEP**	**Diowvilai et al. (2012)**	**Praputtakun et al. (2013)**
Locate Information					
▪ Access and retrieve information within a text (in forms of a few words, phrases or numerical values)	✓	✓	✓	✓	✓
▪ Search and select relevant task					
Understand					
▪ Represent literal meaning (a direct or paraphrase type of match between the question and target information within a passage)	✓	✓		✓	✓

Chulalongkorn University

PISA 2018	International			Thai	
	PISA2009-2015	**PIRLS**	**NAEP**	**Diowvilai et al. (2012)**	**Praputtakun et al. (2013)**
▪ Integrate and generate inferences (at the level of an entire passage, such as identifying the main idea, summarizing, or giving a title)	✓	✓	✓	✓	
Evaluate and reflect					
▪ Assess quality and credibility of the text (whether the information is valid, up to date, accurate, unbiased and whether the author is competent, well-informed and benevolent)	✓	✓	✓	✓	
▪ Reflect on content and form (such as the quality and style of writing)	✓	✓	✓	✓	
▪ Detect and handle conflict					

3.3 Reading literacy framework

Reading literacy is composed of three sources of influence including 1) the reader factors, which composed of motivation, prior knowledge, and other cognitive abilities; 2) the text factors, which constitute the format of the text, the complexity of the language used, and the number of texts reader encounters; and 3) the task factors, which contain the purposes of the task and the complexity or number of tasks to be completed. The interaction of the three factors drives a reader to apply reading literacy processes to complete the tasks (OECD, 2016) as shown in Figure 15. In this study, reading literacy is based on the two sources, namely text factors and task factors.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

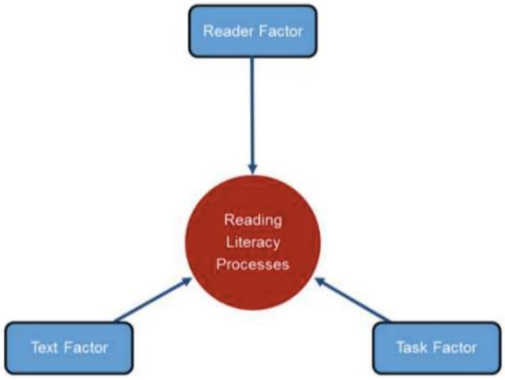


Figure 15 Dimensions of reading literacy (OECD, 2016, p.11)

Framework of PISA reading literacy is composed of three dimensions, including cognitive processes, texts, and the reading situations.

3.3.1 Texts

All texts in reading literacy are based on the daily contents that a student may encounter in their real lives. There are no specific types of texts to assess because learners should encounter a wide range of texts. Thus, the texts will be characterized in several features. In PISA 2009-2015, the texts used in PISA are composed of four classifications: (1) text format, which is continuous texts (e.g., newspaper report, novels) and non-continuous texts (e.g., tables, schedules, indexes and forms, lists, graphs, diagrams, advertisements, catalogues); (2) text type, which represents the purpose of the text (e.g., description, narrative, essay, argumentation or instruction.); (3) medium, which is the form presented in the text including print(paper) or digital(hypertext); and (4) environment, which are divided into an authored environment (the text cannot be modified such as web page) and a message-based environment (the reader can add or change the content such as email, blog) (OECD, 2016; Thomson et al., 2013).

For PISA 2018, the representation of a broader range of texts in the assessment is provided. Texts are both printed and digital medium. There are four dimensions of texts as follows:

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

*(1) **Source*** is either single (e.g. having a definite author, publication date, and reference) or multiple texts (e.g. having different authors, being published at different times, or bearing different titles).

(2) Organization and navigation are static (the simple organization and low density of navigation tools; text) and dynamic (the more complex, non-linear organization and high density of navigation tools).

(3) Format is continuous texts (e.g. newspaper reports, essays, novels, short stories, reviews and letters), non-continuous texts (e.g. lists, tables, graphs, diagrams, advertisements, schedules, catalogues, indexes and forms), and mixed texts (e.g. magazines, reference books and reports where both continuous and non-continuous texts are placed together).

(4) Type

4.1 Description – texts with information related to properties of objects in space. It provides an answer to "what" questions. Examples of description are: a depiction of a place in a travelogue, a catalogue, a geographical map, an online flight schedule or a description of a feature, function or process in a technical manual.

4.2 Narration – texts with information related to time. It provides answers to "when" or "in what sequence". Examples of narration are: a report, a news story, a novel, a short story, or a play, a biography, a comic strip and a newspaper report of an event.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.3 Exposition – texts with information related to how different elements interrelate in a meaningful way. It provides answers to "how" questions. Examples of exposition are: a scholarly essay, a diagram showing a model of memory, a graph of population trends, or a concept map for an entry in an online encyclopedia.

4.4 Argument- texts with information related to relationship among concepts or propositions, expressing opinions or points of view. It provides answers to "why" questions. Examples of argument are: a letter to the editor, a poster advertisement, posts in an online forum or a web-based review of a book or film.

4.5 Instruction- texts with information related to instructions on what to do. Examples of instruction are: a recipe, a series of diagrams showing how to give first aid or guidelines for operating software.

4.6 Transaction- texts with information aiming to achieve a purpose such as requesting that something is done, organizing a meeting or making a social engagement with a friend. Examples of transactions are: a letter, an everyday email, or a text message exchanged between colleagues or friends.

In sum, the texts in PISA 2018 have been extended in the topics of source and organization and navigation, which are different from PISA 2009-2015 to assess the more resourceful and current trends in both digital and non-digital texts.

3.3.2 Tasks

Tasks or situations are involved with the contents related to real-life situations. Reading literacy is scenario-based to help readers accomplish a particular goal. The tasks in the test are provided with a wide range of situations (OECD, 2016, 2019b; Thomson et al., 2013).

(1) Personal – situations that contain the satisfying texts for personal interests in both practical and intellectual ways. Examples of personal texts are personal letters, fiction, diary-style blog, personal emails, instant messages and blogs.

(2) Public – situations that contain texts for activities and concerns of the society. Examples of public texts are official documents, information about public events, message boards, news websites, and public notices.

(3) Occupational – situations that contain texts for the accomplishment of an immediate task. Examples of occupational texts are job advertisements in a newspaper or online job websites, and texts that provide workplace directions.

(4) Educational – situations that contain text for the instruction and are chosen by the instructor rather than the reader. Examples of education texts are printed or electronic textbooks and interactive learning software.

3.4 Reading passages

Reading passages of PISA comprise a wide range of text types, difficulty, and scenarios or situations. The composition of reading passages was based on PISA 2009 when the main emphasis was on reading literacy.

3.4.1 Text types

The reading test covered both the continuous-texts and the non-continuous texts. The non-continuous texts were in the categories of argumentation,

exposition, and transaction. The most common text type in PISA 2009 was exposition, whereas the least common text type was narration.

3.4.2 Difficulty

Readability formulas were used to measure the passage difficulty. The formulas for continuous text were Dale-Chall and Flesch-Kincaid. The average grade level was the 8th grade for Dale-Chall and the Grade 9 for Flesch-Kincaid. The formula for non-continuous text was FORCAST. The average grade level was 10th grade. On the other hand, continuous texts were more difficult than non-continuous texts.

3.4.3 Length

In PISA 2009, there were 2 to 5 items per passage. The test had an average of 354 words per passage and the range of words in passages were 53-758 words.

3.4.4 Scenarios

The reading literacy test was designed to cover a wide range of reading purposes. Passages were selected to represent a wide range of texts and applicability in real-world settings, which covered four scenarios, namely personal, public, occupational, and educational scenarios.

In summary, the passages in this study were based on text types and scenarios. Moreover, the concept of transfer was applied for three testing occasions. Readability formulas were not considered in this study because the formulas had been particularly created for passages in English version, which might not be reliable to be used in Thai language.

3.5 Levels of reading literacy proficiency

Levels of reading literacy proficiency provide the progression of reading literacy demand with the composite scale and each subscale. The scale summarises the proficiency in terms of his or her ability, along with the complexity of an item in terms of its difficulty. The reading proficiency scale represents the idea that students are likely to be able to complete tasks mapped at the same level or lower, and less likely to be able to complete tasks mapped at a higher level on the scale as shown in Figure 16. Higher levels represent the knowledge, skills, and capabilities needed to perform tasks of increasing complexity (OECD, 2016, 2019c). The composite reading

literacy scale and each subscales in reading literacy PISA 2018 were divided into eight levels as follows: Level 1c, Level 1b, Level 1a, Level 2, Level 3, Level 4, Level 5, and Level 6 (OECD, 2019c). The eight levels of reading literacy proficiency are presented in Table 12.

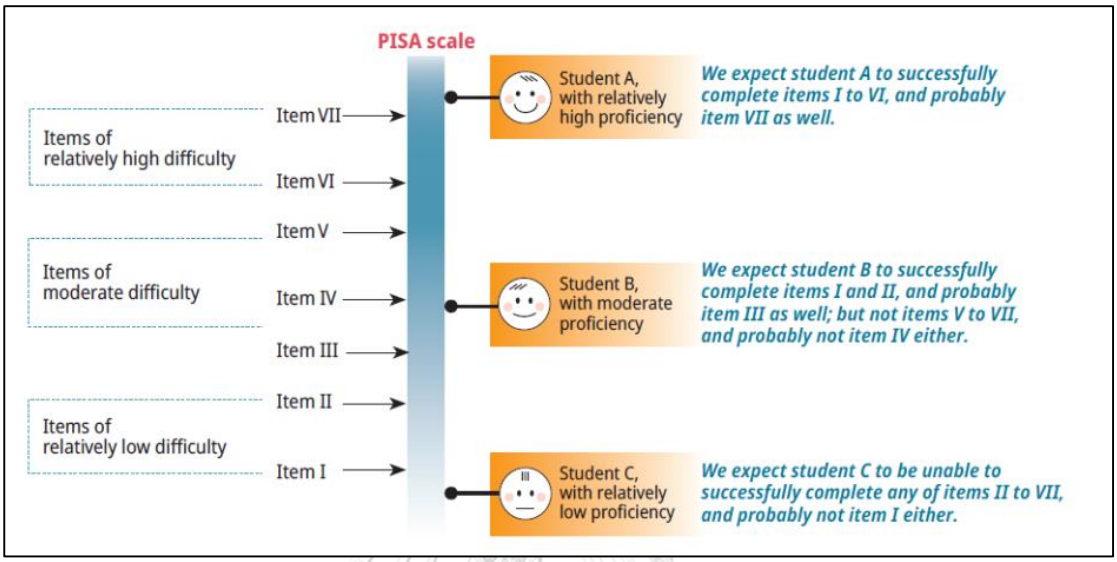


Figure 16 Relationship between items and students on a proficiency scale

Table 12 Levels of reading literacy and descriptions of reading proficiency

Level	Score	Descriptions of reading proficiency	Characteristics of tasks
6	Above 697	• Readers can comprehend lengthy and abstract texts in which the information of interest is deeply embedded and only indirectly related to the task. They can compare, contrast and integrate information representing multiple and conflicting perspectives, using multiple criteria and generating inferences across distant pieces of information to determine how the information may be used. • Readers can reflect deeply on the text's source in relation to its content, using criteria external to the text. They can compare and contrast information across texts, identifying and resolving inter-textual discrepancies and conflicts through inferences about the sources of information, their explicit or vested interests, and other cues as to the validity of the information.	• Tasks require the reader to set up elaborate plans, combining multiple criteria and generating inferences to relate the task and the text. Materials at this level include one or several complex and abstract text(s), involving multiple and discrepant perspectives. Target information may take the form of details that are embedded within or across texts and obscured by competing information.
5	626	• Readers can comprehend lengthy texts, inferring	• Tasks involve dealing with

Level	Score	Descriptions of reading proficiency	Characteristics of tasks
	- 697	which information in text is relevant even though the information of interest may be easily overlooked. They can perform causal or other forms of reasoning based on a deep understanding of extended pieces of text. They can answer indirect questions by inferring the relationship between the question and one or several pieces of information distributed within or across multiple texts and sources. • Readers can establish distinctions between content and purpose, and between fact and opinions as applied to complex or abstract statements. They can assess neutrality and bias based on explicit or implicit cues pertaining to both the content and/or source of the information. They can also draw conclusions regarding the reliability of the claims or conclusions offered in a piece of text.	concepts that are abstract and going through several steps until the goal is reached. Moreover, tasks may required readers to handle several long texts, switching back and forth across texts in order to compare and contrast information.
4	553 – 625	• Readers can comprehend extended passages in single or multiple-text settings. They interpret the meaning of nuances of language by taking into account the text as a whole. Readers can compare perspectives and draw inferences based on multiple sources. • Readers can search, locate and integrate several pieces of embedded information in the presence of plausible distractors. They are able to generate inferences based on task in order to assess the relevance of target information. They can handle tasks that require them to memorise prior task context. • They can evaluate the relationship between specific statements and a person's conclusion about a topic. They can reflect on the strategies that author use to convey their points, based on salient features of texts such as titles and illustrations. They can compare and contrast clains made in several texts and assess reliability of a source based on salient criteria.	• Tasks are often long or complex, and their content or form may not be standard. Many of tasks are situated in multiple-text settings. The texts and the tasks contain indirect or implicit cues.
3	480 – 552	• Readers can represent the literal meaning of single or multiple texts in the absence of explicit	• Tasks require reader to take many features into account

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Level	Score	Descriptions of reading proficiency	Characteristics of tasks
		content. They can integrate content and generate both basic and more advanced inferences. They can also integrate several parts of a piece of text in order to identify the main idea, understand a relationship or construct the meaning of a word or phrase. • They can search for information based on indirect prompts, and locate target information that is not in a prominent position. In some cases, readers can recognize the relationship between several pieces of information based on multiple criteria. • They can reflect on a piece of text and compare and contrast several authos' viewpoints based on explicit information. They may be required to perform comparisons, generate explanations or evaluate a feature of the text.	when comparing, contrasting or categorizing information. The required information is not preminent or there might be a fair amoung of competing information. Texts may include other obstacles, such as ideas that are contrary to expectation or negatively worded.
2	407 – 479	• Readers can identify the main idea in a piece of text of moderate length. they can understand relationship ore meaning within a limited part of the text when the information is not preminent by producing basic inferences. • They can select and access a page and locate one or more pieces of information based on multiple, partly implicit creteria. • They can reflect on overall purpose in texts of moderate length. They can reflect on simple visual or typographical features. They can compare claims and evaluate the reasons supporting them based on explicit statements.	• Tasks involve comparisons or contrasts based on a single feature in the text. Typical reflective tasks at this level require readers to make a comparison or several connections between text and outside knowledge by drawing on personal experience and attitudes.
1a	335 – 406	• Readers can understand the literal meaning of sentences or short passages. Readers can recognize the main theme or the author's purpose in a piece of text about a familiar topic and make a simple connection between the given information and their own prior knowledge. • Readers can select a relevant page based on simple prompts, and locate one or more independent pieces of information within short texts. • Readers can reflect on the overall purpose, and	• Tasks contain explicit cues as regards what needs to be done, how to do it, and where in the texts readers should focus their attention.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Level	Score	Descriptions of reading proficiency	Characteristics of tasks
		on the relative importance of information in simple texts containing explicit cues.	
1b	262 – 334	• Readers can evaluate the literal meaning of simple sentences. They can interpret the literal meaning of texts by making simple connections between adjacent pieces of information in the question • They can scan for and locate a single piece of placed, stated information in a single sentient, a short text or a simple lise. They can access a relevant page from a small set based on simple prompts when explicit cues are present.	• Tasks direct readers to consider relevant factors in the task and in the text. Texts are short and provide support to reader, such as repetition of information, picture of familiar symbols.
1c	189 - 261	• Readers can understand and affirm the meaning of short, simple sentences on a literal level, and read for a clear and simple purpose within a limited amoun of time.	• Tasks involve simple vocabulary and syntactic structures.

3.6 Related studies on reading literacy

There have been two types of research design in studies on reading literacy in Thailand. The first research design is research and development (R&D). Research aimed to develop a model or some methods to improve reading literacy. However, the results of student's reading literacy were still unsatisfactory. Diowvilai et al. (2012) studied the development of the Lampang learning enrichment network to enhance reading literacy of elementary Grade 4-6 students. The main purpose was to develop learning activities to improve reading literacy. The instruments were reading literacy multiple-choice tests for pretest and posttest and the supportive activities. There were supportive activities to develop learning management potential for teachers and supportive learning activities to improve reading literacy for students. The research duration started from September 2009 to March 2012. Reading literacy was categorized into 3 levels as the following: reading comprehension, analytical reading, and critical reading. Reading literacy is measured from 50 multiple-choice questions combined with four levels of proficiency - very good, good, fair, and needs improvement. The results have shown that students' posttest scores improved significantly, but the level of students' reading literacy was

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

below 50%; in other words, they still needed improvement in reading literacy. In addition, Chandai (2016) developed an instructional model based on the scaffolding approach and self-regulated learning for enhancing reading literacy and evaluatin g the effectiveness of the developed model. The sample was 70 Grade 9 students. The experimental lessons were two hours per week and they lasted for 8 weeks. The reading strategies comprised three aspects: access and retrieve, interpret and integrate, and reflect and evaluate. The instrument was a reading comprehension and implementation test. The result revealed that reading literacy of students in the experimental group not only improved from pretest to posttest but was also significantly higher than the control group.

Another research design is survey research to investigate the students' reading literacy or factors affecting reading literacy. Praputtakun et al. (2013) reported the case study of Grade 6 students' scientific literacy and reading literacy. The sample was 175 Grade 6 students in Bangkok. The instrument was a scientific literacy and reading ability test, which was 30 multiple-choice questions. Reading literacy was assessed in two aspects – retrieving information and interpretation. The result revealed that, in terms of reading literacy, the mean score of retrieving information was 4.58 of 7, which counted as 65.5%, whereas the mean score of interpretation was 5.72 of 11, which counted as 52%. Jumnaksarn (2013) studied the development of a causal relationship model of factors affecting the reading literacy of Thai students. Research instruments were secondary data from PISA 2009 international database consisting of 4,866 15-year-old students from 228 schools in Thailand. The results showed that the model fitted between the conceptual model and empirical data. The direct factors affecting reading literacy were reading engagement, metacognition in reading and teaching strategies perceived by students, while the indirect effect was teaching strategies. Nilsawang (2011) investigated the factors affecting reading literacy of Grade 9 students in Srisaket province. The sample was 665 Grade 9 students. The instruments were reading literacy tests and questionnaires on attitude toward reading literacy, learning motivation, learning environment, and family relationship. Factors positively affected reading literacy were time spent on reading and learning environment, whereas factors negatively affected

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

reading literacy were the relationship with a family. The summary of research studies related to reading literacy in Thailand are presented in Table 13.

Table 13 Summary of reading literacy research in Thailand

Researcher	Research objective	Reading literacy	Sample	Instrument	Duration	Result
Diowvilai et al. (2012)	To develop learning activities to enhance reading literary (R&D)	Three levels: (1) Reading comprehension (2) Analytical reading (3) Critical reading	Grade 4-6 students in Lampang province	-reading literacy tests -supportive activities	September 2009-March 2012	(1) The students' posttest improved significantly (2) The students' reading literacy level was need improvement
Praputtakun et al. (2013)	To assess student's scientific literacy and reading ability (Survey)	Two aspects: (1) Retrieving information (2) Interpretation	175 Grade 6 students in Bangkok	-scientific literacy and reading ability test	-	The mean scores of retrieving information was 65.5% and 52% for interpretation.
Chandai (2016)	To develop and assess the instructional model for enhancing reading literacy (R&D)	Three aspects: (1) Access and retrieve (2) Integrate and interpret (3) Reflect and evaluate	70 Grade 9 students	-reading comprehension and implementation test	8 weeks	The experimental group's reading literacy scores was higher than the control group
Jumnaksarn (2013)	To develop causal relationship model affecting reading literacy (Survey)	-	Secondary data from PISA2009 database	PISA instruments (13 booklets and student questionnaire)	-	The model fitted between conceptual model and empirical data
Nilsawang (2011)	To investigate factors affecting reading literacy (Survey)	-	665 Grade 9 students	-reading literacy test -questionnaire	-	Factors positively affected reading literacy were time spending on reading and learning environment.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Part 4: Measurement of Growth

Dynamic assessment is an assessment to estimate the change in multiple testing occasions in order to provide valid information about the potential of an individual or group to develop particular knowledge (Dumas & McNeish, 2017). There are two measurement methods to quantify change: 1) the traditional method of measuring change and 2) the modern method of measuring change.

4.1 The Traditional method of measuring change

4.1.1 Gain scores

Gain score is widely used in the field of dynamic assessment to quantify change or improvement after the intervention. Most of the format of dynamic assessment is Sandwich format from Grigorenko and Sternberg (1998) which is the pretest-training-posttest format. In order to quantify change, the gain scores between pretest and posttest are calculated. The calculation of gain score is as follows

$$Y_2 - Y_1$$

Where, Y_1 = Pretest

Y_2 = Posttest

However, this method has been criticized for measuring change of latent ability. Sternberg and Grigorenko (2002) addressed the problems of using gain scores as follows: (1) lack of reliability of gain scores; (2) the ceiling effects in the subgroup; and (3) the inequality in latent ability gains. They argued that gain scores can be varied in accordance with groups of test-takers and the characteristics of the test, especially item difficulty and discrimination. For example, the high ability group will get the smaller gain on pretest to posttest for the easier test because they perform well on both the pretest and posttest, but they will get the higher gain on the harder one. Unlike the high ability group, the low ability group will get higher gain for the easier test, but they will get smaller gain for the harder one because they perform poorly on both pretest and posttest.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.1.2 Logistic regression analysis

The logistic regression analysis (LRA) extends from the multiple regression analysis in which the outcome is categorical. It is used for testing hypotheses about the relationship between a categorical outcome variable (e.g., success or failure/ improved or not improved) and one or more categorical or continuous predictor variables. However, logistic regression analysis does not model the categorical outcome variable directly, it is assumed to be based on probabilities associated with the value of Y, which is considered as the dichotomous taking on the value 1 (e.g., the positive outcome or success) and the value 0 (e.g., the negative outcome or failure) (Sternberg & Grigorenko, 2002).

4.2 Modern method of measuring change

Nowadays, there are several mathematical models using item response theory or structural equation model to measure change of dynamic assessment as follows:

4.2.1 Multidimensional Rasch-family model for learning and change (MRMLS)

This method has been developed by Embertson as a method integrating Item Response Theory (IRT) to investigate the change in dynamic assessment (Sternberg & Grigorenko, 2002). It has been called the multidimensional Rasch-family model for learning and change (MRMLS) to investigate how much better the learner performs the task at posttest. This model estimates the initial level of ability, indexes of modifiability, and item difficulties indexes. The chance that an item *i* is solved correctly by person *p* (P_{pi}) depends on item difficulty (β_i) and latent ability (θ_p), which included both pretest performance (θ_{p1}), and change performance calculated by posttest minus pretest (θ_{p2}) (Stevenson et al., 2016). The equation of MRMLS is as the following:

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

$$P\left(y_{pi}=1|Z_{p1}\ldots Z_{pJ},\beta_i\right) = \frac{\exp\left(\sum_{j=1}^{J}\varsigma_j Z_{pj}+\varepsilon_p-\beta_i\right)}{1+\exp\left(\sum_{j=1}^{J}\varsigma_j Z_{pj}+\varepsilon_p-\beta_i\right)} \quad \text{where} \quad \varepsilon_p \sim N(0,{\sigma_{\varepsilon p}}^2) \qquad (1)$$

Stevenson et al. (2016) studied the cognitive potential of children with diverse cultural backgrounds by using the multidimensional Rasch-family model for learning and change (MRMLC). MRMLC was calculated to determine the change in ability over time from the initial ability of students (pretest performance) and performance change (posttest minus pretest score called 'performance change scores'). The results revealed that dynamic testing of analogical reasoning was able to reduce the effect of the cultural background. The performance change between indigenous and ethnic minority children did not differ in graduated prompts. Also, there were higher gains from pretest to posttest in children who received the graduated prompt treatment.

4.2.2 Linear logistic test model

One method developed by Fischer (1983, as cited in Grigorenko & Sternberg, 1998) called the linear logistic test model. He has suggested the specified scores for pretest and posttest. The zero scores are for incorrect answers, whereas unity scores are for correct answers. The changes between pre- and posttest answers from 0 to 1 and from 1 to 0 are treated as the change data.

4.2.3 Latent class analysis (LCA)

Another method for investigating dynamic assessment is latent class analysis (LCA). LCA is one of the multivariate analysis methods to classify objects into exclusive groups. The variability within clusters should be minimized, whereas between-cluster variability should be maximized. Membership in a latent class is probabilistic. Each individual can get probabilities to be a member in each class. In dynamic assessment, the unconditional probabilities of being in a latent class can be gainers versus non-gainers. Moreover, the conditional probabilities of making a particular response or demonstrating a particular behavior can be a member in the acquisition of mastery (Sternberg & Grigorenko, 2002).

4.2.4 A Mixed population Rasch model (MIRA)

The combination of item response theory and latent class analysis has successfully been proposed by Rost (1990, as cited in Sternberg & Grigorenko, 2002). He developed a mixed population Rasch model (MIRA). Regarding Sternberg and Grigorenko (2002), an individual will be characterized by only one trait level which is the concept of IRT, and the trait level depends on the class to which the individual belongs. The class is the concept for latent class analysis. Consequently, item difficulties will be interpreted or scored differently based on latent-class membership.

4.2.5 Latent growth curve model (LGM)

Latent growth curve model (LGM) is developed within structural equation modeling (SEM) and is an extended confirmatory factor analysis technique to consider a change. LGM is used to examine changes over time. This approach utilizes both slope (rate of change) and intercept or initial levels (group means) as latent variables to investigate where an individual begins and how the individual changes. According to Duncan and Duncan (2004), this procedure describes an individual's developmental trajectory, capturing individual differences in the trajectories over time and focusing on predictors of individual differences to find out which variables have important effects on the rate of development. The LGM approach provides some of the tools for modeling within-person changes and between-person differences in change (Grimm & Ram, 2009). This approach involves two different analyses which are: (1) the repeated measures of each individual across time (linear or nonlinear) and (2) the individual parameters (slope and intercept values) to determine the growth differences from a baseline measure (Bowles & Montroy, 2013).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

LGM approach is effective for the research design that subjects must be measured repeatedly at least three occasions (Bowles & Montroy, 2013). Moreover, it enables us to test both linear and nonlinear growth models. The data generally

requires large samples, multivariate normal data and equal time intervals for all subjects (Duncan & Duncan, 2004). There are several latent growth curve models as follows: 1) latent curve models, 2) nonlinear latent curve models, and 3) multivariate latent curve models.

4.2.5.1 Latent curve model

Latent curve model is the combination of the basis functions, the level and the slope, plus a set of measurement errors specified to the individual as:

$$y_i = \Lambda_i \eta_i + \varepsilon_i$$

The matrix of basis functions, Λ_i, indicates that the matrix may vary between individuals in terms of its elements. A set of coefficients, η_i, represents change that may vary between individuals. The error of the model, ε_i, shows the discrepancy between an individual's observed score and that based on the fitted trajectory. The error is assumed to be independent with constant variance across time (Blozis et al., 2007). The common formulations of a latent curve model in the present study are no-growth model, linear growth model, quadratic growth model, and unspecified growth model.

1) No-growth latent growth curve model

Latent growth curve with no change can be considered as the baseline model in the examination of change. The model predicts that subjects will maintain the same outcomes or scores across all occasions. In this model, the observed variables, in the form of the squares, are the outcomes or the scores on four occasions. The outcomes also have residual errors which are assumed to be related to other variables in the model. The only latent variable, in the form of the circle, is the level factor or intercept. In the no-growth model, the intercept factor predicts the level of individuals on the outcome variable across all occasions. The intercepts are constrained for an individual across time. The factor loadings are all set to 1 showing that the factor has the same influence on the outcomes at each time point (Bowles & Montroy, 2013). The diagram is shown in Figure 17.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

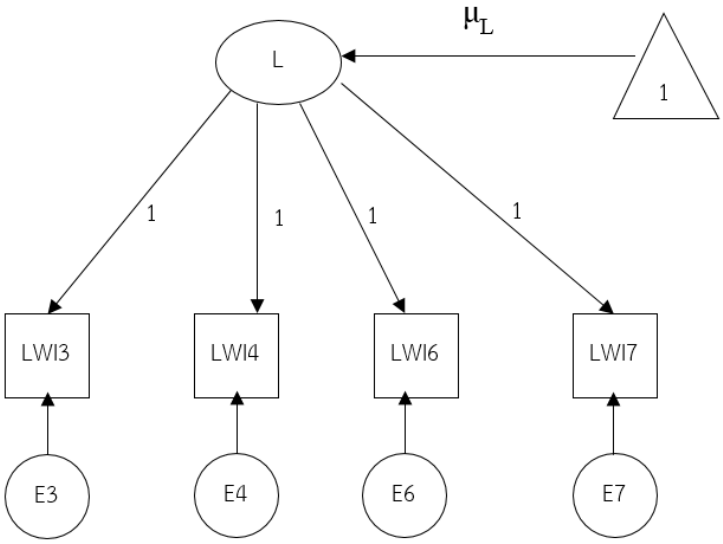


Figure 17 Path diagram for no change latent growth curve (adapted from Bowles & Montroy, 2013, p. 268)

The example is from Bowles and Montroy (2013), using the data from the Letter-Word Identification (LWI) subscale of the Woodcock-Johnson III Tests of Achievement. There were a total of eight occasions, but this example limited the data only from four occasions, namely Occasions 3, 4, 6, and 7, which were labeled as LWI3, LWI4, LWI6, and LWI7, respectively. The outcomes or the scores of individuals on any occasion are equal to the individual's unobserved value on the intercept factor plus random error. The formula of this no-growth model is expressed as

$$Y_n = L_n + e_n$$

Where, Y_n = the outcome variable

L_n = the Level factor or intercept

e_n = the random residual error

In Figure 17, the outcomes were LWI scores. The Level was all set to 1 indicating that the factor had the same influence on the outcomes at each time point. The random residual error also included that is assumed to be unrelated to other variables in the model (Geiser et al., 2013).

2) Linear latent growth curve model

The widely used form of change in the latent growth curve model is linear change. The construct is assessed by a single variable measured at each time point (Geiser et al., 2013). There are two latent variables which are the level or intercept factor and the slope factor. The newly second factor labeled *S* or slope is added in the model representing the growth or rate of change per unit of time. The intercept and slope are latent because they are not variables that exist in the dataset. They are estimated based on the collection of trajectories obtained for each individual. Interestingly, the factor loadings on the slope factor are called the basis coefficients which are used to define the shape of change and interpret the level and slope factors (Bowles & Montroy, 2013). The model of linear growth curve is shown in Figure 18.

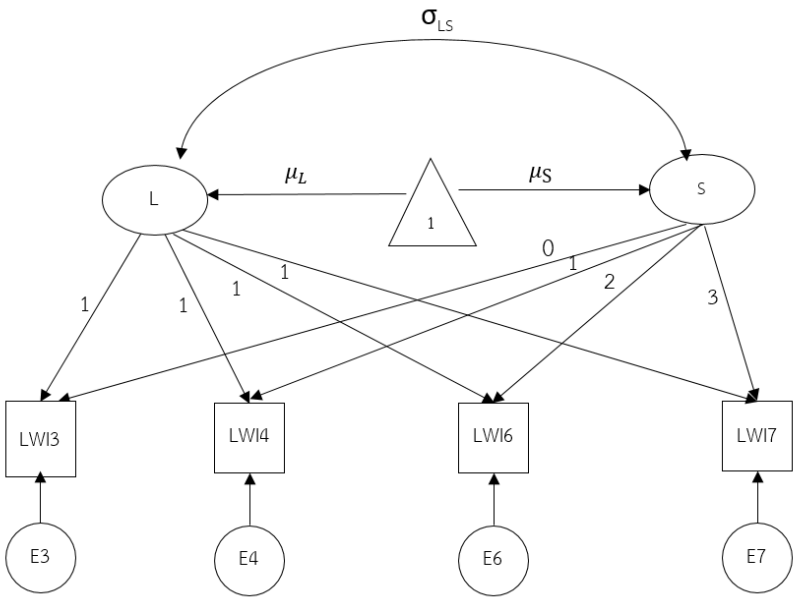


Figure 18 Path diagram of linear latent growth curve (adapted from Bowles & Montroy, 2013, p. 270)

In Figure 18 diagram, it shows the basis coefficients are linearly set to 0, 1, 2, and 3. The intercept or the initial value is still constant across time, but there are linear growths of change in the slope (Bowles & Montroy, 2013). The

diagram shows that the outcomes were LWI at each time point. The intercept parameters were coded 1, 1, 1, and 1 to represent the same starting value. Because this study focused on a test of the linear rate of growth, the slope parameters were linearly coded 0, 1, 2, and 3. The zero was the starting point (Bowles & Montroy, 2013).

The correlation between the intercept and slope should be interpreted carefully. The positive correlation represents the high initial status had a greater rate of change, whereas the negative correlation determines the high initial status had a lower rate of change (Schumacker & Lomax, 2010). The equation of linear latent growth curve model is influenced by a first-order intercept factor, a first-order slope factor, and an error term as follows:

$$Y_{t,n} = L_n + B_t S_n + e_n$$

Where, Y_n = the outcome variable

L_n = the level factor or intercept when time=0

B_t = the basis coefficients or the factor loadings on the slope factor

S_n = the slope or rate of change per unit of time

e_n = the random residual error

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

The means and variances of intercept and slope are used to estimate growth. The mean of the intercept factor represents the average value of the initial status parameter, and its variance represents individual differences in initial status. The mean of the slope factor, that is different from zero, indicates the linear growth trajectory or the group growth parameter, and its variance represents individual differences in change over time (Schumacker & Lomax, 2010).

3) Quadratic growth model

The model is specified to include a linear change rate and an acceleration rate. The linear slope represents the rate of change and the quadratic slope shows the rate of change is changing rapidly. The factor loadings of the quadratic factor are the squared loadings of the linear factor. The intercept, linear

slope, and quadratic slope have means (μ_i, μ_{sl}, and μ_{sq}) variances (σ^2_i, σ^2_{sl}, and σ^2_{sq}), and covariances ($\sigma_{i,sl}$, $\sigma_{i,sq}$, and $\sigma_{sl,sq}$). The residual scores have a mean of zero and an estimated variance (σ^2_e) (Grimm & Ram, 2011). The formula of the quadratic growth model is expressed as (Bowles & Montroy, 2013):

$$Y_{t,n} = L_n + S_n{*}t + Q_n{*}t^2 + e_{tn}$$

Where, Y_n = the outcome variable

L_n = the level factor or intercept when time=0

S_n = the slope or rate of change per unit of time

Q_n = the quadratic factor

t = time

e_n = the random residual error

The basis coefficients of the slope are equal to time (t), and the basis coefficients of the quadratic are equal to t^2. The example path diagram of quadratic is shown in Figure 19.

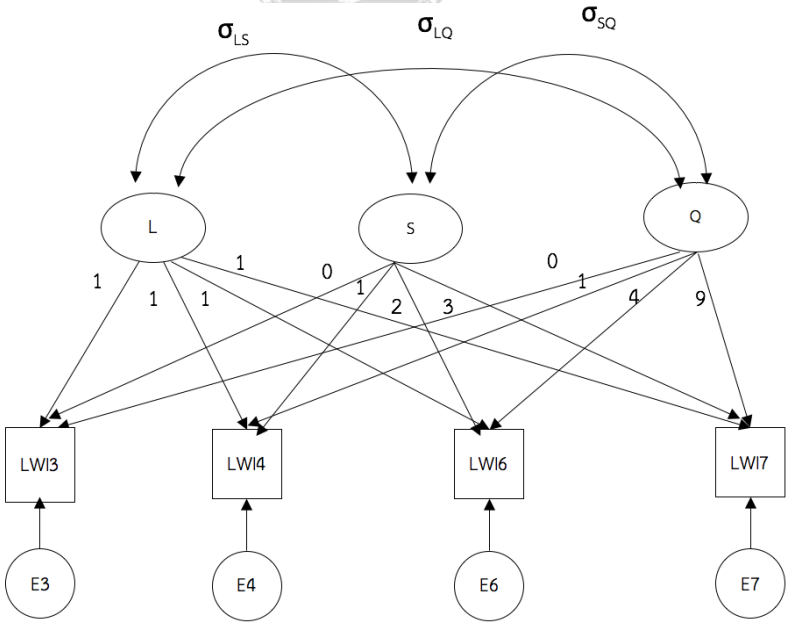


Figure 19 Path diagram of quadratic latent growth curve (adapted from Duncan, Duncan, & Strycker, 2006, p. 27)

In Figure 19, there were two slope factors for linear change and curvilinear change. The factor loadings of intercept for three measurement occasions were similarly fixed to 1. The factor loadings of the linear factor were fixed to 0, 1, 2, and 3 showing the rate of growth in the trajectory, whereas the quadratic factor estimates the squared loading of the slope by fixing the factor loadings to the specific values, that is, 0, 1, 4, and 9. The model estimated 3 means, 3 variances, and 3 covariance for the intercept, linear, and quadratic latent factors.

One of the drawbacks of the quadratic model is that the parameters of the model are difficult to interpret because the linear and quadratic slopes affect the rate of change. Also, the quadratic model cannot present the asymptotic level of development.

4) Unspecified growth model

The basis coefficients can be either specified or unspecified. If the shape of trajectory is not known, the data can determine the shape of the growth curve. The basis coefficients are equivalent to factor loadings and they will be estimated rather than being predefined. The factor loading in the slope factor will be allowed to be freely estimated in the unspecified growth model to explore what growth process is the most likely (Duncan, Duncan, & Strycker, 2006). In order to estimate a linear or a nonlinear growth process, three measurement occasions are required. The unspecified growth model can be called a latent basis model. The straightforward constraints are to select one basis coefficient to be 0 to identify the level factor and another to be 1 to identify the slope factor. The basis coefficient for the first measurement occasion is often fixed to 0 and the last occasion to 1. Other basis coefficients are estimated freely as the proportion of the total change over time (Bowles & Montroy, 2013). The path diagram of the unspecified growth model is shown in Figure 20.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

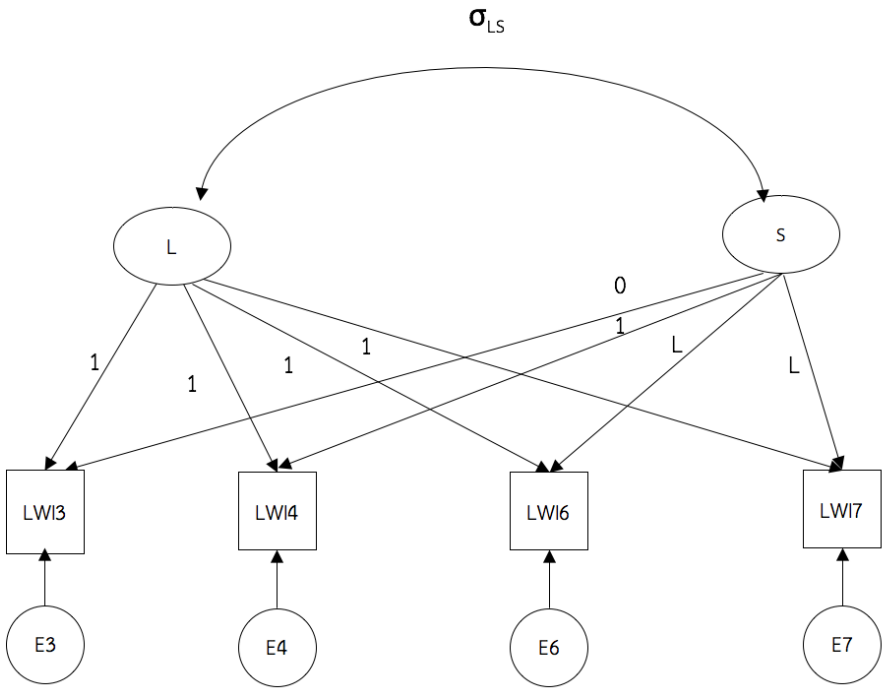


Figure 20 Path diagram of unspecified latent growth curve (adapted from Duncan, Duncan, & Strycker, 2006, p. 31)

In Figure 20, the factor loadings of the intercept factor were fixed to 1 in every measurement occasion. Two factor loadings on the slope factor were fixed to 0 and 1 for the first and second measurement occasions, but the third and fourth loading were estimated freely.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

4.2.5.2 Nonlinear latent curve model

The well-known model of the latent growth curve is linearity; however, the actual change or growth may not be linear in some conditions. Latent growth curve can be tested as nonlinearity for the measurement of three or more time points (Duncan & Duncan, 2004). There are several models to be tested as nonlinearity. According to Harring et al. (2012), selection of a nonlinear function can be justified based on previous research or based on expected behaviors of the underlying change process. Even if other curves perform better, one function might still be preferred and selected because it replicates the development in important ways or captures interesting scientific characteristics.

The nonlinear models describe a series of growth that have specific nonlinear patterns characterized by mathematical functions such as the Gompertz, Richards, logistic, and exponential curves (Grimm & Ram, 2011). They can be categorized into two types as follows: 1) Additive nonlinear latent curves or nonlinear structured latent curve models and 2) Multiplicative nonlinear latent curves (Grimm & Ram, 2011).

1) Additive nonlinear latent curves or nonlinear structured latent curve model

Browne (1993, as cited in Blozis et al., 2007) proposed a structured latent curve model which allows individuals to differ in terms of a nonlinear change rate. A structured latent curve model differs from a latent curve model in terms of trajectory. Under a latent curve model, all individual-level trajectories are assumed to follow the same form of change; however, the mean response is assumed to follow a particular function under a structured latent curve model. It is limited with regard to the kind of functions that may be specified (Blozis et al., 2007) . For the additive nonlinear model, it has been referred to as partially nonlinear random coefficient models (RCM) because of its random coefficients entering the model linearly, whereas the parameters enter the model nonlinearly. Additive models follow specific nonlinear functions and have additive random coefficients. Predicted scores are the sum of the participant's scores on random coefficients multiplied by a function of time that does not vary across participants. Because of the additive random coefficients, the estimation of models requires the ability to impose nonlinear constraints. The models have interpretable parameters and are able to model complex patterns with few parameters (Grimm & Ram, 2011).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2) Multiplicative nonlinear latent curves

The model is referred to as fully nonlinear random coefficient models (RCM). They cannot be estimated directly in SEM because the random coefficients enter the model. The random coefficient, or the rate of change, is

distinct from the total amount of change at the individual level. Individual predicted scores are the sum of the participant's scores on random coefficient, the rate of change, multiplied by another random coefficient, the total amount of change, varying across time and participants. This model is needed if the research focuses on the separating variation in rates of change, timing, and total amounts of change (Grimm & Ram, 2011).

Moreover, the nonlinear model can be divided into two groups regarding growth trajectories including: 1) J-shaped growth trajectories and 2) S-shaped growth trajectories (McNeish & Dumas, 2017).

2.1) J-Shaped Growth Trajectories

The J-shaped growth trajectories feature the rapid growth at the beginning and then slow down after the time progresses to approach the asymptote. This particular feature is appropriate with a wide variety of constructs such as vocabulary. They do not have a lower asymptote and the inflection points. The parameters in these trajectories are initial value, upper asymptote, and the rate parameter. The models in these trajectories are Michaelis-Menten and Von Bertalanffy.

2.2) S-Shaped Growth Trajectories

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

They feature the lower growth at the beginning and get more rapid in the middle and eventually grow slowly before approaching the asymptote. These features are used to measure the field of second language acquisition. The parameters in S-shaped trajectories are 1) a lower asymptote; 2) rate parameter, 3) an inflection point, and 4) a rate-of-approach parameter. The inflection point in an S-shaped trajectory is the point at which the rate of growth no longer is increasing and the outcome variable continues to grow at a decreasing rate. The rate-of-approach is how quickly an inflection point occurs. The models in these trajectories are Schnute curves, the Richards curve, Gompertz, and Logistic.

4.2.5.3 Multivariate latent curve model

A multivariate latent curve model can provide estimates of individual change in measured constructs and provides information on the degree to which change in different constructs are associated. It is useful in measuring relationships in terms of covariances among intercepts, linear slopes, and occasion-specific residuals across different outcomes. A multivariate model for two measures can be specified as

$$\begin{bmatrix} y_{1i} \\ y_{2i} \end{bmatrix} = \begin{bmatrix} \Lambda_{1i} & 0 \\ 0 & \Lambda_{ki} \end{bmatrix} \begin{bmatrix} \eta_{1i} \\ \eta_{2i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{ki} \end{bmatrix}$$

Where, n_{1i} and n_{2i} are the random coefficients for the two measures

Λ_{1i} and Λ_{ki} are the common matrices corresponding to two measures

The benefits of multivariate models are the intraindividual change and interindividual differences in the level and slope. Also, it can answer how one variable is associated with change in another. There are several approaches to multivariate latent growth models. The first-order multivariate latent growth model is associative LGM and the second-order multivariate latent growth models are factor-of-curves and curve-of-factors models.

1) Associative latent growth model

The model examines the correlations among parameters of the measured constructs. It is necessary to estimate the univariate model of each construct before formulating an associative latent growth model in order to examine intraindividual and interindividual variation whether intercept and slope factor variances are significantly different from zero. Then, each model will be combined to estimate simultaneously to examine relationships among the individual parameters for measured constructs (Duncan et al., 2006) as shown in Figure 21.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

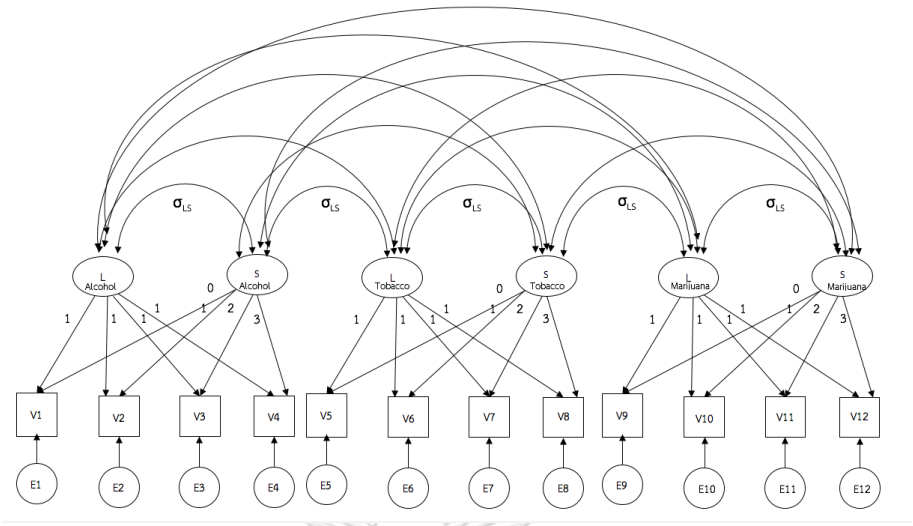


Figure 21 Path diagram of associative latent growth model (adapted from Duncan, Duncan, & Strycker, 2006, p. 66)

Figure 21 displays the associative form of growth among the growth parameters of adolescent alcohol, tobacco, and marijuana use. The factor loadings are constrained at 0,1,2, and 3 to show linear growth over time.

2) factor-of-curves model

The model examines whether a higher order factor describes relationships among lower order functions. The higher order model has a similar structure as the first-order associative model, whereas the covariance among the first order constructs are fixed at 0 and are changed to be explained by the higher order factor instead. The factor loading between the second-order factor and the first-order factor are restricted to be equal over time for each measured construct. Hence, the first-order factor describes individual differences within each univariate series, whereas the second-order factor describes individual differences among the first-order latent growth model (Duncan et al., 2006).

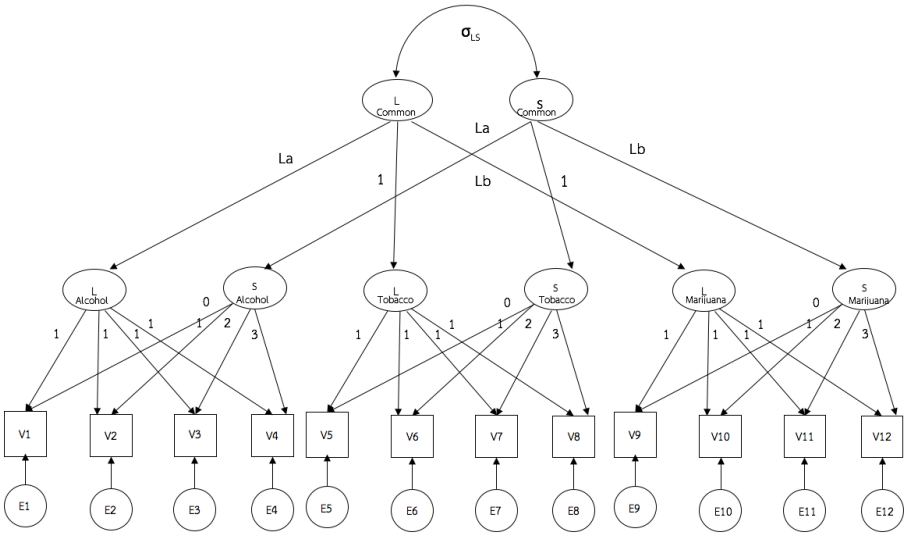


Figure 22 Path diagram of the factor-of-curves latent growth model (adapted from Duncan, Duncan, & Strycker, 2006, p. 70)

Figure 22 shows whether the higher order common factor, which is substance use construct, could describe the relationships among the growth factors of alcohol, tobacco, and marijuana use. The loadings between the second-order for substance use and the first-order tobacco factors were fixed to 1 for all measurement occasions. For the loadings of alcohol factors and the marijuana factors with second-order substance use are restricted to be equal over time.

3) curve-of-factors model

The model represents what measured variables have in common at each time point. The observed variables at each time point are factor-analyzed to produce factor scores that will be used for modeling growth curves. Because the model requires factor pattern invariance, the common factor elements must be equal over time (Duncan et al., 2006).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

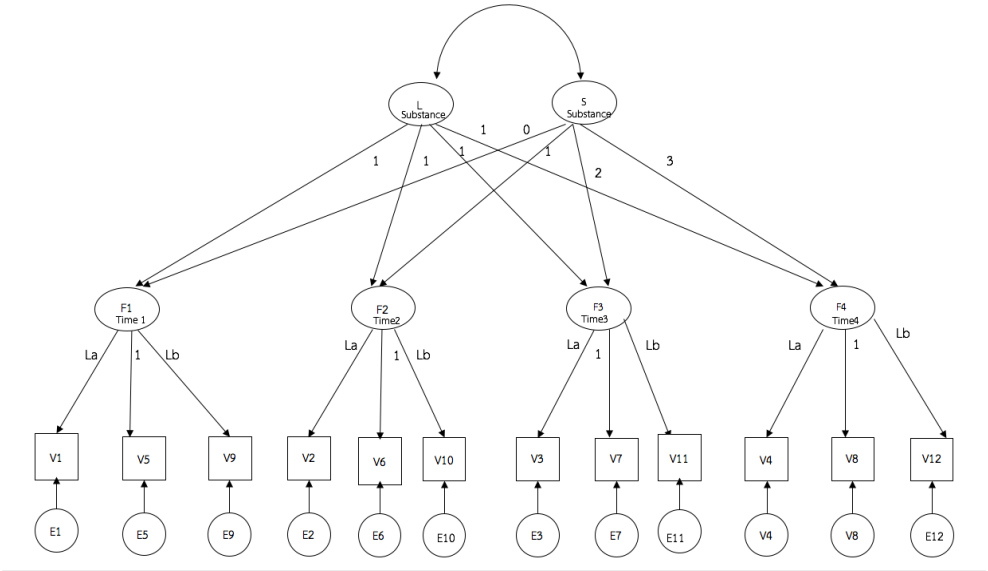


Figure 23 Path diagram of the curve-of-factors latent growth model (adapted from Duncan, Duncan, & Strycker, 2006, p. 72)

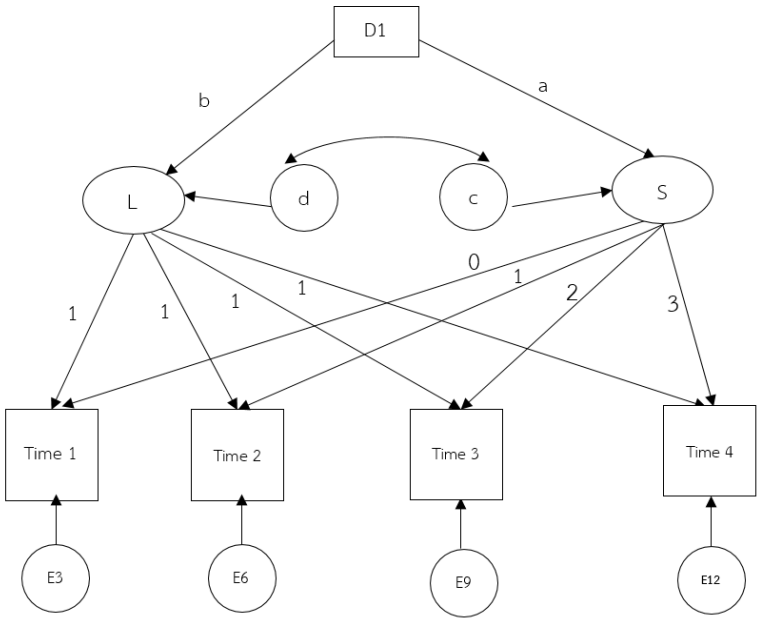
In Figure 23, it represents what the three substance use behaviors have in common at each time point in order to analyze substance use factor scores. Tobacco use (V5 to V8) was used as the scaling reference, fixing to be 1 over time. The alcohol use and marijuana use were constrained to be equal across time.

4.2.5.4 Latent growth curve model in multiple populations

The LGM model can be used to investigate group differences. The group differences can be fixed (e.g. race or gender) or can be varied by time (e.g. memory retention or cognitive ability). Data from different individuals may provide comparable information (Duncan & Duncan, 2004).

The example of Schumacker and Lomax (2010) tested gender as group differences on the cell phone customer satisfaction. The gender was tested as a dummy variable (1=male; 0=female) across four fixed time points (3, 6, 9, 12 months). The intercepts and slopes were tested across four time periods to investigate changes over time; group differences across time periods were tested. The slope coefficients were 0, 1, 2, and 3 to indicate a linear test across the time periods. The intercept coefficients were 1,

1, 1, and 1 to represent the means for the time periods. The diagram of the latent growth curve in multiple populations is shown in Figure 24.



Where, a = slope, b= intercept, c = error terms for slope, d = error terms for intercept, and D1 = group difference variable (dummy variable)

Figure 24 Latent growth curve model with gender differences (adapted from Schumacker & Lomax, 2010, p. 214)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

The formulas for measuring latent growth curve model for group differences were as follows:

The level 1 model:

$$Y_{iT} = \alpha_i + \lambda_T \beta_i + \varepsilon_{iT}$$

The level 2 model:

$$\alpha_i = \mu_\alpha + \gamma_{\alpha D1} D_{1i} + \varsigma_{\alpha i}$$
$$\beta_i = \mu_\beta + \gamma_{\beta D1} D_{1i} + \varsigma_{\beta i}$$

Where,

$\gamma_{\alpha_{D1}} D_{1i}$ = the slope parameter estimate for the dummy coded variable (Male = 1)

$\gamma_{\beta_{D1}} D_{1i}$ = the intercept parameter estimate for the dummy coded variable (Male = 1)

ς_{α_i} = the error term for slope

ς_{β_i} = the error term for intercept

4.2.5.5 Second-order latent growth curve model

The second-order latent growth model or curve-of-factors model uses multiple observed variables, whereas the first-order latent growth curve models (FGMs) are used to estimate change from a single observed variable. In SGMs, the construct is assessed by at least two observed variables at each time point. The intercepts are constrained to be equal to present that change over time. The growth process is modeled at the level of latent rather than observed variables (Hancock et al., 2001). The benefits of SGMs include: separating measurement error from true trait change, enabling to test the assumption of measurement invariance across time, having the greater statistical power to detect individual differences in change, and using multiple indicators (Geiser et al., 2013). The diagram of the second-order latent growth curve is shown in Figure 25.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

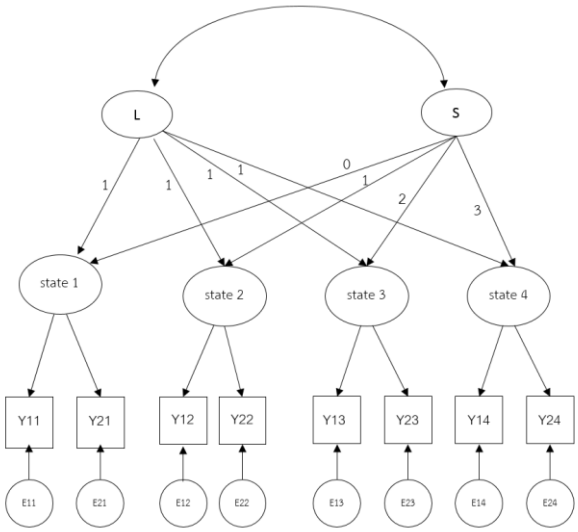


Figure 25 Linear second-order latent growth curve model for four measurement occasions (adapted from Geiser, Keller, & Lockhart, 2013)

The formulas of second-order latent growth curve model are as follows (Hancock et al., 2001):

The level 1 model:

$$Y_{it} = \tau_t + \Lambda_t \eta_i + \varepsilon_{it}$$

The level 2 model:

$$\eta_i = \alpha + \Gamma x + \varsigma_i$$

Where,

Y = the observed variable measured at time *t*

τ = a vector of variable intercepts

Λ = a matrix of loading relating each η_j construct to its measured variable indicators

η = the vector of growth factors

ε = a vector of random normal error

Γ = a matrix of second-order factor loading reflecting the growth pattern underlying the η_j constructs

x = a vector of covariate values

ς = a vector of random effects

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

In conclusion, there are many statistical techniques to measure change in dynamic assessment studies as presented in Table 14. The traditional method of measuring gain scores is the widely popular method; on the other hand, little DA research has utilized the modern statistical methods. In the present study, multivariate latent growth curve model was used to indicate change or growth of a student's performance of three constructs of reading literacy because the model enabled the relationships among the measured constructs.

Table 14 Summary of the methods for measuring change

	Method	Uniqueness	Advantages	Disadvantages	Related DA studies
Traditional Methods	Gain scores	The tradition method for measuring change from pretest to posttest	-Easy to calculate	-The ceiling effect -Re-testing effect	
	Logistic regression analysis	The relationship between outcome variable and predictor variables	-Use predictor variables to predict the outcomes	-The outcome variable is only categorical.	Shrestha (2017)
Modern Methods	Multidimensional Rasch-family model for learning and change (MRMLC)	The investigation of how learner performs the task at posttest	-Estimate the initial level of ability, indexes of modifiability, and item difficulties indexes. -Use small samples	-The complexity of the approach	Stevenson, Heiser, & Resing (2016)
	Linear logistic test model	The changes between pre- and posttest answers	-	- The limited estimation only dichotomous score (0,1)	-
	Latent class analysis	The classification of objects into exclusive groups.	-Focusing on structure of the cases	-Latent variable and manifest variable are only categorical	-
	Mixed population Rasch model (MIRA)	The combination of item response theory and latent class analysis	-The combination of item response theory and latent class analysis	-The complexity of the approach	-
	Latent growth curve	The measurement of individual growth	-Estimating both linear and nonlinear curves - Detecting group and individual differences in changes	- The complexity of the approach -The selected model should match with the empirical data	McNeish and Dumas (2017)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Part 5: Measurement of Learning Potential

In the educational field, measurement of learning potential gradually gains popularity because of the increasing focus of students' progress. However, in the field of dynamic assessment, it is unclear on how to define the true meaning of learning potential and how to measure it. In the current study, there will be two topics relating to the measurement of learning potential as follows: 1) concept of learning potential and 2) measurement of learning potential.

5.1 Concept of learning potential

Dynamic assessment is capable of estimating students' potential for developing a particular skill or ability (Dumas & McNeish, 2017). The goal of dynamic assessment is to evaluate, to intervene, and to change (Grigorenko & Sternberg, 1998) According to Grigorenko and Sternberg (1998), learning potential is something developing and modifiable rather than something developed; it is the means of assessing students' ability to learn by teaching them something and observing their learning to predict future performance. Learning potential of students can be quantified during the acquisition of new cognitive operations. This is coherent with Stevenson et al. (2016), addressing that learning potential provides information on one's ability to learn from instruction. Resing and Elliott (2011) said that potential for learning, or learning potential, is defined as successful outcomes and strategies employed by an individual student while solving tasks.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

As regards to Poehner and Lantolf (2013), learning potential is defined as the openness to mediation that enables individuals to take over the mediational process and to use it for their own learning independently. This predicts future responsiveness to instruction and provides insights into the degree of instructional effort that will likely be required to help learners learn in their future learning. Therefore, several scholars in the field of dynamic assessment believe that learning potential is the students' ability to learn through the assistance received during learning and assessment that can predict their future performances.

Furthermore, Dumas and McNeish (2017) have defined learning potential as the conceptualization of three components which are: ability, capacity, and availability as follows:

1.) Ability is the amount of skill that a student has at a particular time or the capability of measuring a student's current ability at that time point. Ability provides limited information about an individual's potential for future growth because of its focus on developed ability, not developing ability. In order to understand the potential for growth in ability, multiple testing occasions are necessary and ability should be measured several times.

2.) Capacity is the maximum amount of ability that a student can be expected to attain. Dynamic assessment framework normally assumes the capacity to grow in linear function over time, but some abilities are unlikely to grow linearly.

3.) Availability is the discrepancy between capacity and ability (i.e. availability = capacity - ability) which resembles Vygotsky's concept of the zone of proximal development (ZPD). Research on dynamic assessment has paid attention to gain insight into the zone of proximal development (McNeish & Dumas, 2017)as shown in Figure 26.

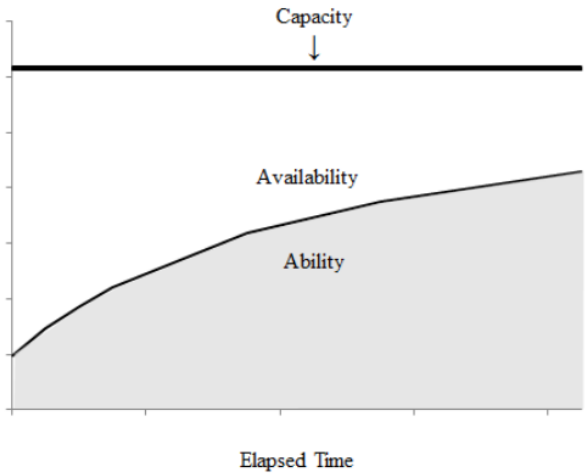


Figure 26 Three components of potential (adapted from McNeish & Dumas, 2017, p. 62)

From McNeish and Dumas's viewpoint, if instructor says that someone has potential, it may mean that the individual has already developed skills and can presently show in their tasks (high ability) and the individual has a high ceiling for developing relevant skills (high capacity), or that the individual could develop in the future but has yet to do presently (high availability), or some combination of those three aspects (McNeish & Dumas, 2017).

5.2 Measurement of learning potential

Learning potential is the capability hidden inside which is beyond the measurement of a static test. In the field of dynamic assessment, learning potential has been measured by calculating the learning potential scores and quantifying the efficiency of prompts as the following:

5.2.1 Calculating the learning potential scores

5.2.1.1 Learning potential score (LPS)

Kozulin and Garb (2002) proposed a formula for calculating Learning Potential Score (LPS) in order to capture the degree of change between independent performance and mediated performance in a single score by including the maximum scores in the formula. Following Poehner et al. (2015), the learning potential score is used to measure an openness to mediation and to predict how well a learner responds to instruction in the future as follows:

$$\text{LPS} = \frac{(\text{S post} - \text{S pre})}{\text{Max S}} + \frac{\text{S post}}{\text{Max S}} = (2\ \text{S post} - \text{S pre})/\text{Max S}$$

Kozulin and Garb (2002) categorize learners into three subgroups based on their LPS scores as follows: 1) High scorers (LPS ≥ 1.0), 2) Mid-range scorers (.99 ≥ LPS ≤ .72), and 3) Low scorers (LPS ≤ .71). The result of LPS scores shows that learners in different groups may need different supports to enhance their learning abilities.

Following the study of Poehner et al. (2015), the LPS of learners ranged from .59 to 1.24 in the listening test and from .77 to 1.19 in the

reading test. The interesting findings were learners with the same actual scores had different mediated scores and LPS scores. If LPS scores represented the learners' zone of proximal development, learners who had the same independent performance but different LPS scores might require different levels of support in a classroom in the future.

5.2.1.2 Dynamic measurement modeling

Dumas and McNeish (2017) proposed dynamic measuring modeling to measure potential using the second-order nonlinear growth model. When considering the concept of learning potential, capacity and availability which are the developing components, cannot be measured by using the single-timepoint measure. The capacity is the upper limit of ability growth that cannot be directly measured by the linear trajectory. A suitable model for measuring these components needs to be longitudinal and nonlinear with a functional form to reach an upper asymptote (McNeish & Dumas, 2017). Moreover, the ability and capacity should be measured on the same scale as one another that can be directly comparable. The model should include random effects on each of its parameters to provide the subject-specific information rather than the average information and allows the data of each individual capacity or availability to be estimated (McNeish & Dumas, 2017). Therefore, the dynamic measuring modeling (DMM) will be selected to analyze learning potential because of its capability in estimating subject-specific ability and availability at each time point as well as a subject-specific time-invariant capacity asymptote. DMM incorporates individualized growth trajectories for every student in a data set. DMM conceptualizes nonlinear growth models in a novel way as a method to measure student capacity (McNeish & Dumas, 2017).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Dynamic measurement modeling has applied the Michaelis-Menten nonlinear growth function to analyze the data. Nonlinear models truly benefit trends in the data and also the growth parameters can vary among individuals (Cameron et al., 2015). The Michaelis-Menten model is the conditionally linear mixed effects

model that can enter the function in a nonlinear fashion or called an inverted J-shaped growth (Harring et al., 2012). It is appropriate to investigate the nonlinear change in a continuous latent variable that is repeatedly measured over time (Harring et al., 2012). The model is composed of three parameters- initial value, upper asymptote, and mid-point. The mid-point or rate parameter is the midpoint at which the outcome variable is halfway between the initial value and the asymptote. The upper asymptote is the maximum capacity that a student can attain (Dumas & McNeish, 2017).

In order to analyze the nonlinear growth trajectory using the dynamic measurement modeling, the procedures are based on the study of McNeish and Dumas (2017) as the following:

1) Because of the longitudinal data, the longitudinal invariance model will be established using confirmatory factor analysis. Moreover, the best fitting nonlinear growth model will be selected from several nonlinear curves (e.g., Gompertz, Michaelis-Menten, Exponential, von Bertalanffy etc.). McNeish and Dumas (2017) suggested how to select the most appropriate model by choosing a small number of models that align well with the theoretical modeling goals and then comparing the empirical fit of these models without subject-specific growth parameters.

2) It was found that the Michaelis-Menten trajectory yielded a lower Bayesian information criterion (BIC) compared to other trajectories. Moreover, the rate parameter, showing the point at which growth is halfway between the intercept and asymptote, gives more rich information than other models. Hence, the Michaelis-Menten (MM) model, which specifically emphasizes information about the growth in a students' ability over time to estimate capacity, seems to fit the type of trajectory most. The capable program of estimating this model is structural latent curve model from Mplus and non-linear mixed effect model from SAS Proc NLMIXED. The formulas of this model are as follows:

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

$$(1) \qquad Y_{it} = \tau_t + \Lambda_t \eta_{it} + \varepsilon_{it}$$

where

Y_{it} = a *p x 1* vector of manifest indicator variables at time for the *i*th individual

τ_t = a *p x 1* vector of manifest indicator variable intercepts at time *t*

Λ_t = a *p x q* matrix of factor loading at time *t*

η_{it} = a *q x 1* vector of latent variable for the *i*th person at time *t*

ε_{it} = a *p x 1* vector of residuals for the *i*th individual at time *t*

p = the number of manifest variables

q = the number of latent variables

$$(2) \qquad \eta_{it} = \beta_{0i} + \frac{(\beta_{Ui} - \beta_{0i})}{\beta_{Ri} + t} t + d_{it}$$

where

β_{0i} = the initial value which captures ability when time is 0 ($\alpha_0 + \varsigma_{0i}$)

β_{Ui} = the maximum value of the outcome as time reaches infinity ($\alpha_U + \varsigma_{Ui}$)

β_{Ri} = the point in time when ability is halfway between the initial value and the asymptote ($\alpha_R + \varsigma_{Ri}$)

d_{it} = a residual term that captures the difference between model-implied and observed values for the *i*th person at time *t*

α = a population-averaged fixed effect that represents the growth factor means

ς_i = a student-specific random effect which allows each person to have their own unique growth curve

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

According to Equations (1) and (2), they show the statistical equations to measure the latent growth model in general. Equation 1 is the first-order measurement model to measure the overall scores or sum scores of students' performance on the measured construct across each time point. Whereas, the Equation 2 points out the second-order growth model that measures each aspect of the measured construct across each time point. A Michaelis-menten equation is used in Equation 2. The three beta parameters are composed of a population-averaged

fixed effect (α) and a student-specific random effect (ς). The random effects allow each person in the data to have their own unique growth curve.

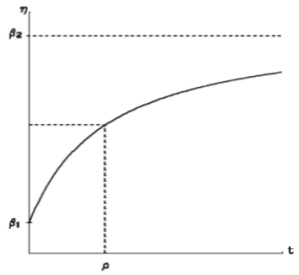


Figure 27 Michaelis-Menten function (Harring et al., 2012)

As shown in Figure 27, the complex system can be treated with the same estimation strategies used in linear growth models (Harring et al., 2012). The three parameters of the Michaelis-Menten functions are: 1) the initial value (β_{0i}), which captures ability when time is 0; 2) the rate parameter (ρ), which represents the point in time when ability is halfway between the initial value and the asymptote; and 3) the asymptote (β_{U_i}), which represents the maximum value of the outcome as time approaches infinity. The upper capacity asymptote is modeled with a subject-specific random effect that makes each individual student in the data set receive a unique estimate; it represents the amount of the measured ability that a student is predicted to develop in the future (McNeish & Dumas, 2017).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Because the Michaelis-Menten model has parameters that enter the model in a nonlinear fashion, namely the rate parameter, that makes the model cannot be directly fit in structural equation software. The structured latent curve model (SLCM) can be linearized by setting constraints on the loading from the second-order growth factors to the first-order factors. The constraints are based upon the first partial derivatives of the target function of each parameter. The parameters need to be constrained to follow a pre-specified nonlinear function (Dumas & McNeish, 2017) as shown in Equations 4-6.

The mean response of the Michaelis-Menten target function is shown in Equation 3.

(3) $$f(\alpha_0, \alpha_U, \alpha_R, t) = \frac{(\beta_{Ui} - \beta_{0i})}{\beta_{Ri}+t}\ t$$

The partial derivatives are as follows:

(4) $$\frac{\partial}{\partial \alpha_0} f(\alpha_0, \alpha_U, \alpha_R, t) = 1 - \frac{t}{t+\alpha_R}$$

(5) $$\frac{\partial}{\partial \alpha_U} f(\alpha_0, \alpha_U, \alpha_R, t) = \frac{t}{t+\alpha_R}$$

(6) $$\frac{\partial}{\partial \alpha_R} f(\alpha_0, \alpha_U, \alpha_R, t) = -\frac{(\alpha_A - \alpha_0)t}{(\alpha_R+t)^2}$$

In order to measure the learning potential, three components of learning potential are composed of **(1) *Ability***, the realized ability at the specific time; **(2) *Availability***, the untapped potential at time *t*; and **(3) *Capacity***, the upper asymptote of ability. Each of the growth model coefficients has a random effect; each individual has a unique initial value, upper asymptote, and growth. Estimation of these parameters can lead to quantifying the ability, capacity, and availability. Ability and capacity are directly estimated from the model, but availability is calculated by subtracting the estimated ability from the subject-specific capacity estimate (McNeish & Dumas, 2017).

(7) Ability = η_{it}

(8) Capacity = $\alpha_U + \varsigma_{ui}$

(9) Availability = $\beta_{Ui} - \eta_{it} = \alpha_U + \varsigma_{ui} - \eta_{it}$

Where,

α = a population-averaged fixed effect that represents the growth factor means

ς_i = a student-specific random effect which allows each person to have their own unique growth curve

Based on Equations (7) - (9), they address how to measure learning potential regarding ability, capacity, and availability. The ability and capacity will be directly estimated from the model. The η_{it} represents ***the estimated ability*** or the scale score for the *i*th person at time *t*. The Equation (8) shows ***capacity*** or the upper

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

asymptote for the *i*th person modeled with a subject-specific random effect to make sure that each individual receives a unique estimate. Random effects (ς_{ui}) are defined to have a mean of 0 as the average capacity for person *i*. The availability is represented in the Equation (9) to calculate the differences between capacity and ability at time *t* (McNeish & Dumas, 2017).

5.2.2 Quantifying the efficiency of prompts

Several computerized dynamic assessment studies have utilized a graduated prompting approach in their studies. Computing the mediating prompts are essential to better understanding how much help or assistance a student relies on during the test.

The prompts are statements, hints, leading questions that are provided gradually from most implicit to most explicit. If a student answers incorrectly, he or she is provided with the most implicit prompt and then a student will be able to re-attempt the question. If a student' s second answer is incorrect, the more explicit mediating prompt will be offered. This process continues until either a student answers correctly or a student gets all four mediating prompts. For the multiple-choice questions with five answers, there will be a total of four mediating prompts in each item.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Efficiency of prompts can be measured by counting the total number of prompts that students require. It is essential to indicate the amount of external mediation or support a student needs to attain his or her learning goal (Poehner et al., 2015). Resing and Elliott (2011) said that the amount of prompts is an indication to predict future school success.

Several studies have investigated the amount of prompts. Resing and Elliott (2011) investigated the minimum amount of prompts a student needed for solving the test items. The prompts will be arranged as the instructional prompting method, ranging from implicit to explicit. The results revealed that students needed fewer prompts in their second training than their first training which could indicate

that students clearly profited from the prompts. The students who scored poorly on the pretest tended to need more prompts than others as shown in Figure 28. Stevenson et al. (2016) investigated the number of prompts required from the graduated prompting approach. The study revealed that both students with constructed-response condition and with multiple-choice condition required fewer prompts during the training. The students with multiple-choice conditions required fewer prompts on average than the other group. Poehner et al. (2015) investigated the difference between the independent and mediated performance. The results showed that student scores improved as a result of mediation. Students with higher actual scores had higher mediated scores.

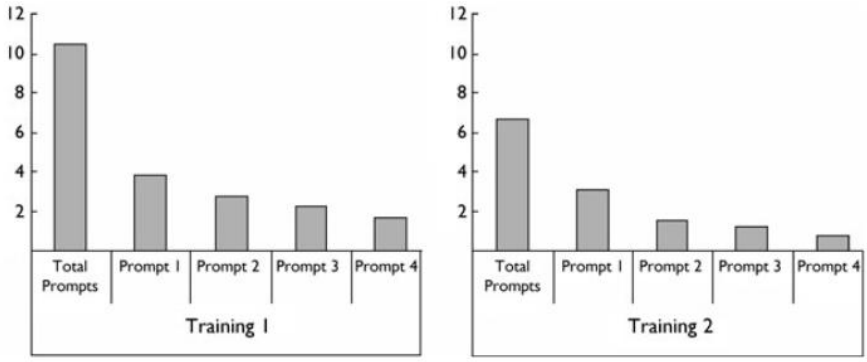


Figure 28 Number of prompts per training session (adopted from Resing and Elliott, 2011, p.591)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Hence, this study applied the dynamic measurement modeling as the statistical technique to compute capacity and availability scores of each individual. Moreover, the efficiency of prompts was measured in terms of mediated score and levels of promptings. Counting the sum of mediated scores determined the amount of mediating prompts required to finish the test and levels of promptings determined what kind of mediation students required whether it is explicit or implicit.

Part 6: Test Validation Using Item Response Theory

6.1 Concept of item response theory

Item Response Theory (IRT) is based on the notion of Item Characteristic Curve (ICC), estimating the probability of a correct response on a test item as a function of the item characteristics (e.g. difficulty, discrimination, and guessing) and the ability level of the test-takers. IRT estimates and interprets item statistics referred to parameters. In IRT, parameters of the items are invariant across individuals, and parameters of the persons are invariant across items (Kanjanawasee, 2012).

Item parameters in IRT are estimated directly using three logical unidimensional models (Kanjanawasee, 2012). The unidimensional models will be presented as follows:

The one- parameter logistic (1-PL) model differs only in difficulty (b); the slopes (a) are constant and no guessing (c). 1PL model assumes equal discrimination parameter (a) across items. The equation of one-parameter logistic (1-PL) model is presented as:

$$P_i(\theta) = \frac{1}{1+ e^{-(\theta - b_i)}}$$

The two-parameter logistic (2-PL) model presents that items are different in terms of difficulty (b) and slopes (a), without guessing (c). The equation of two-parameter logistic (2-PL) model is presented as:

$$P_i(\theta) = \frac{1}{1+ e^{-Da_i(\theta - b_i)}}$$

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

The three-parameter logistic (3-PL) model presents that items differ in terms of difficulty (b) and slopes (a), and pseudo-guessing (c). The equation of three-parameter logistic (2-PL) model is presented as:

$$P_i(\theta) = C_i + \frac{1-C_i}{1+ e^{-Da_i(\theta - b_i)}}$$

6.2 Assumptions of IRT

IRT assumptions that need to be met are unidimensionality, local independence, and non-speeded tests. Unidimensionality shows that each test item should measure a single latent trait. The common IRT models assume that a single underlying ability accounts for the performance on the test. Local independence

emphasizes the probabilities that each student's correct answer is not affected by other examinees and other items on the test. Students' responses to test items are affected solely by the trait intended to measure. Also, IRT assumption is that tests are non-speeded because speeded tests provide additional dimensions that are not related to the measured ability (Kanjanawasee, 2012). Because the passage-based assessment may violate these first two assumptions, the testlet effect should be checked.

6.3 Testlet effect

Testlet effect is the study of the effect of item grouping. Testlet is a group of items sharing the same stimuli such as shared reading passages, scenarios, figures, or tables (Baghaei & Ravand, 2016). Wainer and Kielly (1987, as cited in Ha, 2017)and Byun and Lee (2016) mentioned that several IRT studies have considered passage-based reading tests as testlet-based tests. The items sharing the same subskills or passages which lead to correlated responses to other items within the same testlet are called local item dependence (LID).

Local dependence included in the testlet-based test violates the local independence assumption, which is the fundamental IRT assumption. The violation of using the standard item response theory (IRT) models when LID is present causes significant problems such as biased item parameter estimates, underestimated standard errors (SE) and overestimated reliability and test information (Baghaei & Ravand, 2016; Byun & Lee, 2016). As a result, ignoring the presence of LID probably leads to problems in the valid test score and interpretation. In order to investigate the local dependence of the unidimensional model, the best fitted unidimensional model will be selected. The standardized local dependence (LD) χ^2 statistics will be computed, representing the extent of item dependence for each item pair (Byun & Lee, 2016). Min and He (2014) propose a guideline for the interpretation of LD-χ^2 values as follows: if the values exceed 4, it shows clear local dependence between items and if the values exceed 10 shows the extreme local dependence.

The solutions to addressing local item dependence are as follows: 1) using score-based polytomous IRT models such as the graded response model and polytomous logistic regression and 2) using the testlet-effects models. Score-based

approaches have been criticized for reporting biased parameter estimates and substantial overestimation of reliability and test information values (Ravand, 2015). The application of testlet-effects models, namely Bifactor and Testlet Response Theory (TRT) models, are preferred in order to account for LID without loss of information. TRT models are multidimensional Item Response Theory (MIRT) (Ha, 2017). The Testlet-effects models have been proposed by several researchers. The two-parameter testlet model and three parameter model have been proposed by Bradlow, Wainer, and Wang in 1999 and Wainer, Bradlow, and Du in 2000, respectively (Kim, 2017).

1) The testlet model

The uniqueness of the testlet model is the inclusion of the random effect parameter γ, which describes the interaction between persons and items, representing the local dependency within each testlet $g(i)$ (Bradlow et al., 1999). The larger variance indicates a greater effect for testlet g (DeMars, 2006, 2012). The model yields two ability parameters, a general ability θ and a testlet specific ability $\gamma_{g(i)}$, which is fixed across items and random across persons (Baghaei & Ravand, 2016) When $\gamma_{g(i)}$ is zero, the model reduces to the standard 2PL IRT model as there is no testlet effect (Eckes, 2014). The two-parameter testlet model is presented as:

$$P_i(\theta) = \frac{1}{1+ e^{D\boldsymbol{a_i}(\boldsymbol{\theta}-b_i-\boldsymbol{\gamma_{g(i)}})}}$$

where $P_i(\theta)$ is the probability that an examinee answers item i correctly, (θ) is the latent trait representing person ability, a_i is the item discrimination parameter, b_i is the item difficulty parameter, e is 2.71828, D is 1.7, and $\gamma_{g(i)}$ is the random testlet effect parameter for person j on testlet g(i).

2) The bifactor model

Within the bifactor analysis, the model assigns separate discrimination parameters to the general and group dimensions. Each item response is a function of the primary trait θ_P and one testlet trait θ_{Tg}, in which g is the testlet. Each item has a nonzero value for the discrimination in the general dimension and for one of the testlet dimension. The mean and variance of each θ are set to 0 and 1 (DeMars,

2006). The other discriminations in the vector are fixed to zero. Using d_i as item difficulty, easier items have higher values for d, as opposed to unidimensional notation. The formula of the two-parameter bifactor model equation was presented (DeMars, 2006, 2012).

$$P_i(\theta) = \frac{1}{1+ e^{D(a_{Pi}\theta_P + a_{Ti}\theta_T - d_i)}}$$

where $P_i(\theta)$ is the probability that an examinee answers item i correctly given the θ and the item parameters, (θ) is composed of θ_P and the vector θ_T of the testlet traits, a_{Pi} is the item discrimination on the primary trait, a_{Ti} is a vector of testlet discrimination parameters for item i, d_i is the item difficulty parameter, e is 2.71828, and D is 1.7.

In order to investigate the testlet-effects model, the most fitted unidimensional model was compared with the testlet-effects model to estimate the best-fitting model by comparing the goodness-of-fit indices such as -2loglikelikhood values, which is known as a kind of deviance statistic indicating model fit, as well as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The model which had the lowest -2likelihood, AIC or BIC values would be considered to have the best fit and the most suitable model for the dataset (Byun & Lee, 2016; Ha, 2017). According to Eleje and Esomonu (2018), the analysis of model fit was a check on internal validity; if the model was acceptable, the items were valid.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

6.4 Item parameter analysis

The item parameters are estimated using the data from students' responses to select the good items that have appropriate values. There are three item parameters to be selected, namely the item discrimination (a-parameter), item difficulty (b-parameter), and guessing item (c-parameter). The selection of item parameters will be as the following:

6.4.1 The a-parameter

The measure of the steepness of the item characteristic curve (ICC) known as item discrimination (slope). If a parameter value is positive, it refers that students with higher ability have a high probability of answering an item correctly

and students with lower ability have a low probability of answering an item correctly. When a-parameter is negative, it means that students with higher ability have a low probability of answering an item correctly, but students with lower ability have a high probability of answering an item correctly. A high discrimination value indicates that the item discriminates well between low-and high-level students. It typically ranges from 0 to 3. The values above 1.0 are considered highly discriminant. The acceptable discrimination values that are good for item selection ranges from 0.50 to 2.50 (Kanjanawasee, 2012).

6.4.2 The b-parameter

The location index along the x-axis that tells how easy or how difficult an item is, which is known as item difficulty. The negative item difficulty indicates that the item is easy, whereas a positive item difficulty shows that the item is difficult. The difficult item functions for the high-ability students and the easy one function for the low-ability students. According to Baker (2001), the difficulty values usually range from -3 to +3. An item is considered good for selection if the item has b-value between -2 to +2 (Kanjanawasee, 2012)

6.4.3 The c-parameter

It is known as a pseudo-guessing parameter to estimate the likelihood that an examinee with very low ability can guess the correct answer. The c-value should be 0.20 or lower to be considered a desirable value. The probability of correct guessing for 0.25, showing that student who randomly selects the item from four options can answer correctly about 1 out of 4 times (Kanjanawasee, 2012)

6.5 Item and test information functions

The information functions represent the precision of measurement at each ability level for either an individual test item or the entire test. The greater information presents the more precise or reliable measurement at each ability level (Baker, 2001). The information functions reveal the reliability of the scale at the different levels of the measured dimensions. (1) Item information function (IF) is the

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

amount of information of a single item can be computed at any ability level and is denoted by $I_i(\theta)$, where *i* indexed the item. The amount of item information corresponds to the item's difficulty parameter. The highest amount of information shows the greatest precision at the ability level. (2) Test information function (TIF) is the amount of information yielded by the test at any ability level, in other words, the sum of the item information at that level. It is denoted by $I(\theta)$. The more items in the test, the greater the amount of information (Baker, 2001; Kanjanawasee, 2012).

The equation of test information function will be:

$$I(\theta) = \sum_{i=1}^{k} I_i(\theta)$$

The equation of item information function will be:

$$I_i(\theta) = \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Part 7: Equivalence Analysis

Parallel test forms are needed in this research study. According to Rogers (1995, as cited in Bordens & Abbott, 2011), parallel forms should have the same number of items and the same response format, along with the same issues and items are equally difficult. Moreover, the parallel test forms need to be equivalent in terms of statistical and content related properties so that the test scores are comparable. In short, the parallel versions of a test must be as equivalent as possible. The equivalence analysis will be assessed as follows:

(a) Content-related property: The item-objective congruence (IOC) index

Test contents are assessed before the administration by the group of experts to check the parallelism of the test contents. The item-objective congruence (IOC) index is used to depict the agreement of experts on the equivalent tests. Experts are asked to rate their opinions by using three-point rating scales, giving a rating of 1 (items are in parallel forms), -1 (items are not in parallel forms) and 0 (uncertainty). The index score of each item ranges from -1 to +1. If the IOC is higher than .5, the items are parallel forms and are appropriate to use in the main study. If the IOC is .5 or less than .5, the items are not in parallel forms and need revisions (Kantiwong, 2017).

(b) Statistical-related property: Item response theory (IRT)

Item response theory (IRT) is used to check parallel test forms. Items are parallel if they have equal parameters. In this sense, the TCC and TIF are the same (Ali & van Rijn, 2016). There are two methods to analyze parallel forms of the test, namely using the test characteristic curve (TCC) and using the test information function (TIF). The test characteristic curve focuses on the parallel forms in terms of difficulty, whereas the test information function addresses the parallel forms in terms of precision. The test information function is commonly used as the statistical function rather than the test characteristic curve. According to Ali and van Rijn (2016), differences in difficulty can be lessened by equating, but differences in precision cannot. Checking the parallel forms by using the test information function (TIF) and item information function (IIF) are more appropriate. The item information function provides the amount of information at specific ability values and the test information

function is a summation of the all item information functions. In order to compute the parallel forms, the root mean square deviation (RMSD) statistic is used to quantify the average difference between the test scores of the assembled test form and the reference test form (Debeer et al., 2017; Kantiwong, 2017). If $RMSD_{IF}$ and $RMSD_{TIF}$ values are smaller than .50, it is considered to be a parallel test form (Debeer et al., 2017). The formation of IIF- and TIF-parallel test forms are presented below.

The IIF-parallel test forms will be computed as follows:

$$RMSD_{IF} = \sqrt{\frac{\sum_{j=1}^{q}[i(\theta j)-i'(\theta j)]^2}{q}}$$

The TIF-parallel test forms will be computed as follows (Ali & van Rijn, 2016):

$$RMSD_{TIF} = \sqrt{\frac{\sum_{j=1}^{q}[I(\theta j)-I'(\theta j)]^2}{q}}$$

Where, $RMSD_{IIF}$ = Root mean square deviation of item information

$RMSD_{TIF}$ = Root mean square deviation of test information

q = The number of ability levels at which the functions are evaluated

$I(\theta)$ = Test information

$i(\theta)$ = Item information

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Part 8: Research Framework

This research aimed to develop the computerized dynamic assessment to measure reading literacy, growth of learning, and learning potential of Grade 9 students. The conceptual framework was drawn upon the theories and the research studies related to the construction of the computerized dynamic assessment, the reading literacy framework, the test validation, and the measurement of current reading literacy, growth, and learning potential as the following:

1. Psychometric properties of the reading literacy instruments

1.1 Content validity was how well the reading literacy instruments measured what it was intended to measure and the prompts matched the prompt guidelines defined.

1.2 Model fit and item analysis was determined to measure the most fitted model to the reading literacy data. Both unidimensional IRT models and multidimensional IRT models were computed and compared. Item analysis was the selection of the test items in terms of item discrimination and item difficulty.

1.3 Confirmatory factor analysis was determined to measure whether the reading literacy instruments matched with theorectical constructs.

1.4 Reliability was the degree to which the instrument produced consistent results. Omega coefficient was used for internal consistency.

1.5 Test and item parallelism was the equivalence of test forms in terms of test information function (TIF) and test characteristic curve (TCC). Moreover, the item-by-item parallelism was measured by item information function (IF) and item characteristic curve (ICC).

2. Computerized dynamic assessment for reading literacy

Computerized dynamic assessment for reading literacy was developed to measure reading literacy, growth, and learning potential using the graduated prompting approach. Dynamic assessment was the procedure integrating teaching and assessment simultaneously by offering the appropriate forms of mediation to students during the assessment process (Davin, 2013; Elliott, 2003; Lantolf & Poehner, 2004). In this procedure, mediation was provided during the assessment

procedure. Mediation in this study was the use of prompts which could determine the amount and type of assistance students required to perform well in classrooms in the future (Poehner et al., 2015). It is noted that different types of prompting methods could lead to different types of responses and probably benefited students in different ways (Lin, 2016; Wu et al., 2017). In this study, there were four different prompting methods to assess and mediate students as follows: 1) instructional prompting, 2) error-explanation prompting, 3) mixed prompting, and 4) verification prompting (control condition) (Poehner et al., 2015; Ting & Kuo, 2016; Wu et al., 2017).

2.1 Instructional prompting was the prompt or mediation to help a student better understand the text or the question. The prompt was arranged from most implicit to most explicit. The most implicit prompt was provided when a student answered incorrectly at the first try. If a student answered incorrectly again, the second prompt was more explicit. The pattern of prompting was based on the incorrect attempt of students.

2.2 Error-explanation prompting was the prompt or mediation based on the answer choice a student selected and the incorrect attempt of the student in order to provide the reason why the selected choice was wrong. If a student selected the incorrect answer choice, the prompt was provided depending on the error of his or her selected response, ranging from implicit to explicit.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.3 Mixed prompting was the prompt based on the number of times a student answered incorrectly and a choice he or she selected incorrectly. The prompt showed the reason why the chosen answer was not the correct answer and would help a student to better understand the text or the question.

2.4 Verification prompting was the prompt given only to tell a student that the response was either correct or incorrect.

3. Reading literacy was the measurement of a student's reading literacy posttest score that a student had at the current time without the assistance or mediation provided. Reading literacy measured in this study comprised three

subscales followed the reading literacy framework of PISA 2018 proposed by OECD (2016), including 1) locate information, 2) understand, and 3) evaluate and reflect.

3.1 Locate information was the ability of a student to quickly read the texts for specific purposes. The tasks required a student to scan a single text and retrieve target information in the form of words, phrases or numerical values and to search information from one or several texts by using the help of text descriptors such as headers, source information etc.

3.2 Understand was the ability of a student to understand the meaning of the text and draw a conclusion of the text's meaning and purpose. The tasks required a student to comprehend sentences by finding the main idea or making a summary and to integrate and generate inferences from simple connecting inferences to more complex coherence relationships.

3.3 Evaluate and reflect was the ability of a student to reflect on the content and form of the text and to evaluate the quality of the information. The tasks required a student to evaluate the quality and credibility of the text, to reflect on the quality and style of the writing, to detect the soundness of the claims and handle conflict of the contradicting multiple texts.

4. Growth of learning was the measurement of growth trajectories from the computerized dynamic assessment for reading literacy. Reading literacy growth was estimated from univariate latent growth and multivariate latent growth curves.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.1 Univariate latent growth was the measurement of growth in overall reading literacy and each reading literacy subscale over three time points.

4.2 Multivariate latent growth was the measurement of interrelation among the reading literacy subscales over three time points.

5. Learning potential was the measurement of the hidden reading literacy ability of an individual student that could not do it presently but needed help or assistance to develop their potential in order to do it independently in the future. Learning potential was measured in terms of availability score, mediated score, and levels of prompting.

5.1 Availability score was the measurement of distance between the predicted capacity and the current ability of a student over three time points. It indicated how far a student would reach their attainment.

5.2 Mediated score was the measurement of efficiency of prompting. It measured the amout of support a student needed to answer the questions correctly or to attain their goals.

5.3 Levels of prompting was the measurement of the prompts counted in each level that a student required to answer items correctly. The first level of prompting was the most implicit that a student needed for the first time when he/she answered incorrectly. A student was obtained prompting until he/she answered correctly or received the last prompt as the correct answer.

As a result, the research framework of the present study was shown in Figure 29.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

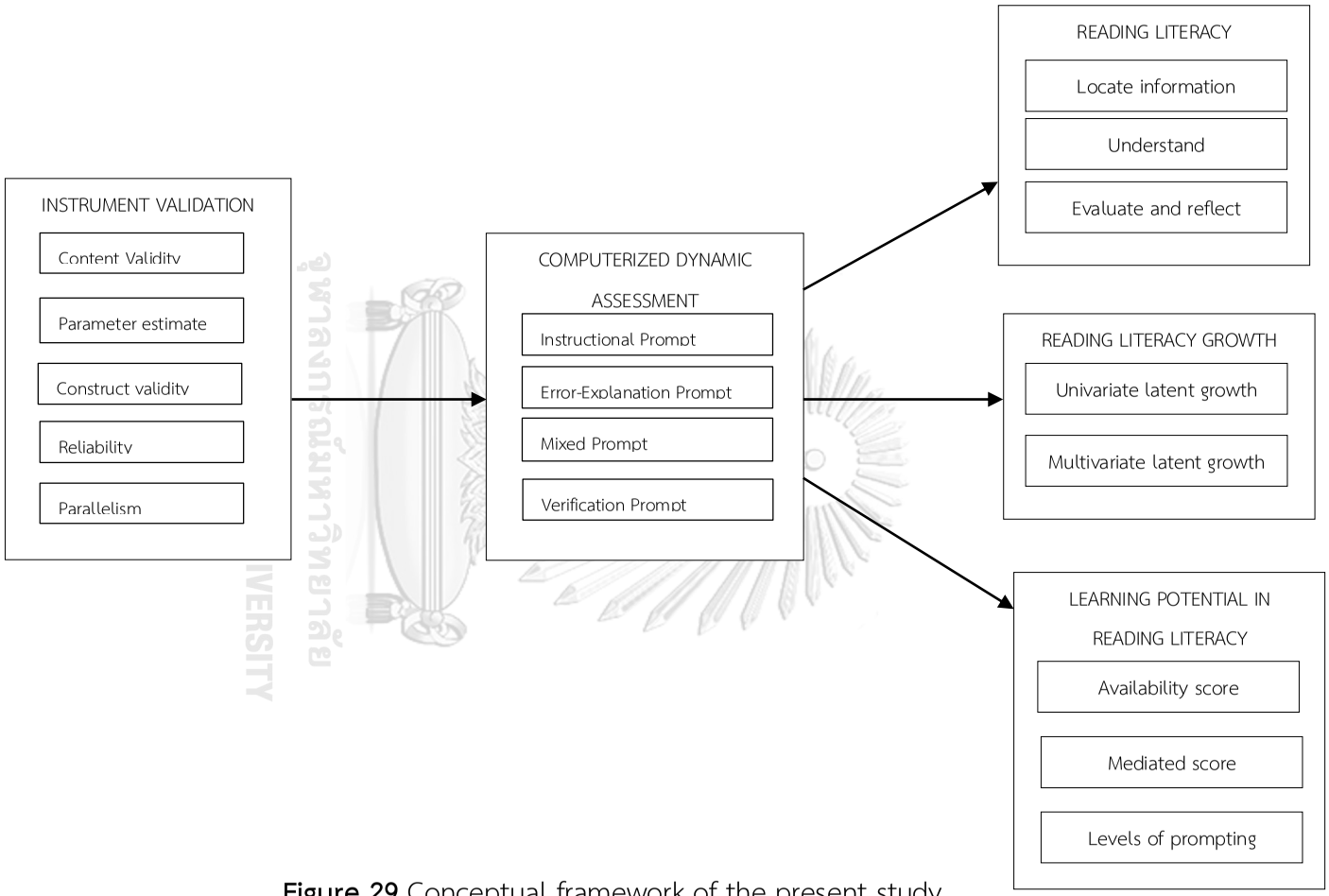


Figure 29 Conceptual framework of the present study

CHAPTER THREE
RESEARCH METHODOLOGY

This chapter discussed the research design and methodology that were used in the study. The population and participants, research procedure, research instrument, data collection, and data analysis were presented in this chapter.

1. Participants

1.1 Groups of participants

1.1.1 Participants for the large-group pilot study

A group of Grade 9 students at schools under the Secondary Educational Service Area Office 1 were recruited from those who had similar characteristics to the participants but did not take part in the experiment.

There were a total of 277 Grade 9 students for pilot study for the validation of reading literacy pretest and posttest. The multi-stage random sampling was used to select the participants. In the first stage of sampling, the researcher used stratified random sampling to select schools based on school size as follows: medium, large, and extra large schools. For the second stage, random sampling was used to select schools. There were a total of 6 schools under the Secondary Educational Service Area Office 1. For the third stage, students were randomly selected to take part in the pilot study. The majority of the participants were female (n = 168, 60.64%) and 39.35% (n = 109) were male. Most of them were studying at extra-large schools (n = 186, 67.14%), followed by medium schools (n = 47, 16.96%) and large schools (n = 44, 15.88%). The data was carried out in January-March 2020.

Moreover, 525 Grade 9 students of six secondary schools in Bangkok were recruited to validate the reading literacy instruments for computerized dynamic assessment. The selection of the participants was designed by using a multi-stage random sampling procedure. Of these individuals, 52.19% (n = 274) were female and 47.81% (n = 251) were male. The majority of the participants (63.05%, n = 331) were recruited from extra-large schools. These samples span a broad range of their grade point average(GPA). Half of the participants (55.23%, n = 290) indicated having a GPA

ranging from 3.00-4.00. Approximately 42% (n = 220) of the participants had a GPA of 2.00 to 2.99. Based on this data, they were divided into three groups as sub-samples. The demographic data of sub-samples are presented in Table 15.

Table 15 Characteristics of the samples

Characteristic	Sub-sample 1 (n_1 = 181)	Sub-sample 2 (n_2 = 178)	Sub-sample 3 (n_3 = 166)	Total Sample (n = 525)
Gender: (%)				
Female	81 (44.75)	105 (58.98)	88 (53.02)	274 (52.19)
Male	100 (55.25)	73 (41.02)	78 (46.98)	251 (47.81)
School size: (%)				
Extra large	111 (61.32)	110 (61.79)	110 (66.26)	331 (63.05)
Large	30 (16.58)	30 (16.86)	24 (14.46)	84 (16.00)
Medium	40 (22.10)	38 (21.35)	32 (19.28)	110 (20.95)
GPAX: (%)				
3.00-4.00 (High)	93 (51.38)	107 (60.11)	90 (54.22)	290 (55.23)
2.00-2.99 (Mid)	86 (47.51)	65 (36.52)	69 (41.57)	220 (41.90)
0.00-1.99 (Low)	2 (1.11)	6 (3.37)	7 (4.21)	15 (2.85)

1.1.2 Participants for the main study

The participants were Grade 9 students in eleven secondary schools under the Secondary Educational Service Area Office 1 and the Secondary Educational Service Area Office 2 in the academic year 2019. The participants were selected using a multi-stage random sampling method. Stratified random sampling was used to select schools in accordance with school size. Moreover, schools in each school size were randomly selected. For the final stage, ninth-grade classrooms in each school were randomly selected to participate in this study. The participants were 600 ninth grade students selected from 11 secondary schools in Bangkok. Of these students, 541 students had valid data and had no missing data on the dependent variables. As a result, a sample size of 541 was used for answering research questions. They were randomly assigned into the instructional-based group (*n*=148), the error-explanation group (*n*=139), the mixed prompting group (*n*=131), and the control group (*n*=123). Of these students, 56.93% (n = 308) were female and 43.06% (n = 233) were male. They were studying at medium schools (40.11%, n = 217), followed by extra-large

schools (27.35%, n = 148) and large schools (21.07%, n = 114). Most of them were very high achievers (38.81%, n = 210), followed by high achievers (29.39%, n = 159), mid achievers (17.37%, n = 94), and low achievers (14.41%, n = 78), respectively.

In terms of levels of learning achievement, the participants were classified into groups by using their grade point averages (GPAs) on last semester as 1) very high achievers (those who received 3.51-4.00 of their GPAs), 2) high achievers (those who received 3.01-3.50 of their GPAs), 3) mid achievers (those who obtained 2.51-3.00 of their GPAs), and 4) low achievers (those who obtained 0.00-2.50 of their GPAs). The participants' demographics are shown in Table 16.

Table 16 The demographics of the participant

Variables	Group				Total
	Group A	Group B	Group C	Group D	
Gender: (%)					
Female	86 (15.89)	74 (13.67)	76 (14.04)	72 (13.30)	308 (56.93)
Male	62 (11.46)	65 (12.01)	55 (10.16)	51 (9.42)	233 (43.06)
Size of schools: (%)					
Extra-large	38 (7.02)	42 (7.76)	36 (6.65)	32 (5.91)	148 (27.35)
Large	31 (5.73)	29 (5.36)	27 (4.99)	27 (4.99)	114 (21.07)
Medium	61 (11.27)	52 (9.61)	55 (10.16)	49 (9.05)	217 (40.11)
Small	18 (3.32)	16 (2.96)	13 (2.40)	15 (2.77)	62 (11.46)
Learning achievement: (%)					
Very high	61 (11.27)	56 (10.35)	48 (8.87)	45 (8.31)	210 (38.81)
High	42 (7.76)	39 (7.20)	41 (7.57)	37 (6.83)	159 (29.39)
Mid	23 (4.25)	27 (4.99)	21 (3.88)	23 (4.25)	94 (17.37)
Low	22 (4.06)	17 (3.14)	21 (3.88)	18 (3.32)	78 (14.41)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

1.2 Sample size

The sample size was based on statistical analysis. In this study, the latent growth curve model was used to analyze the data, which required larger sample sizes than other approaches. The acceptable sample size for structural equation modeling was varied depending on many factors, especially the complexity of the model. Most importantly, the sample size must be adequate to represent the population of interest (Hair et al.,

2010). Bentler and Chou (1987, as cited in Kelloway, 2015)and Hair et al. (2010) suggest the minimum sample size of SEM based on the ratio of cases per estimated parameters. The widely accepted sample size is a 10:1 ratio or a 5:1 ratio. In the present study, the multivariate model had a total of 39 model parameters. Thus, the minimum number of samples was 390 for a 10:1 ratio. Moreover, the sample size estimate was calculated using the G*power for the analysis of covariance (ANCOVA) of four group means. The estimated samples were 107 students to achieve a power of 0.95 with the moderate effect size (d=0.50). Accordingly, in order to avoid the missing data, the total number of 541 samples in this study was appropriate and met the minimum required samples of the parameter estimation and the analysis of power from G*power.

1.3 Ethical issues

This research was approved by the Research Ethics Review Committee for Research Involving Human Subjects: The Second Allied Academic Group in Social Sciences, Humanities and Fine and Applied Arts at Chulalongkorn University with the COA No. 068/2563. The researcher informed the participants about the right of any individual to refuse to participate or to withdraw at any time. Also, the participants were informed that they were protected from physical and mental harm and danger that may arise during the research study. As a result, their information was protected and confidential.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2. Research Procedure

The research procedure consisted of three phases. The first phase associated with the construction of the instrument. The second phase involved psychometric validation of the instrument. Then, the third phase involved the implementation of computerized dynamic assessment for reading literacy in order to measure current performance of reading literacy, growth of learning, and learning potential. The research procedure is shown in Figure 30.

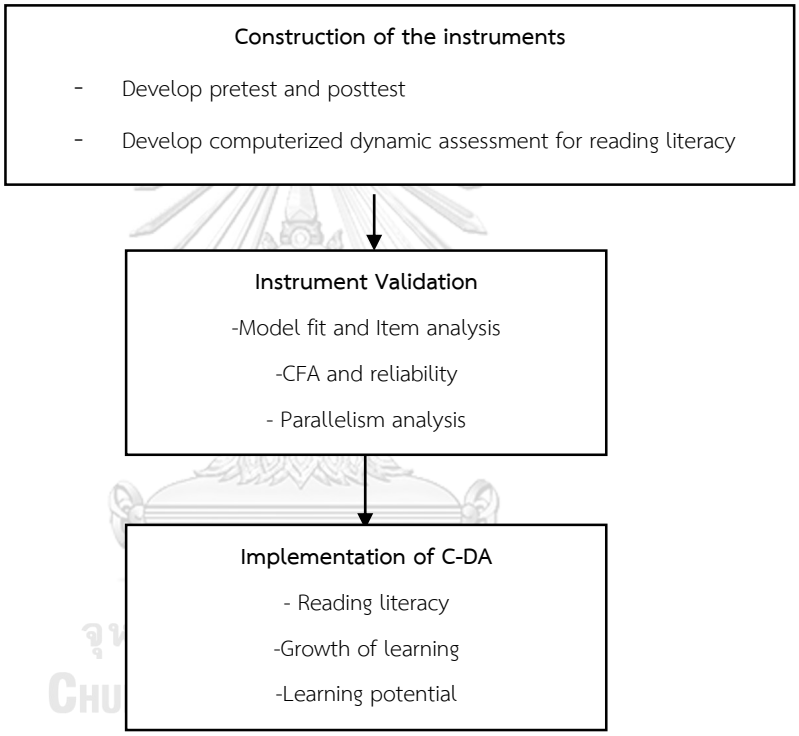


Figure 30 Research procedure of the present study

2.1 Phase 1: Construction of the instrument

2.1.1 Review literature and related studies

The researcher studied the basic concepts and related documents related to reading literacy, especially the aspects of reading literacy, test components, and examples of reading literacy tests. Moreover, the researcher studied the theory and concepts about dynamic assessment, especially the notion of the zone of proximal development, dynamic assessment approaches, the graduated prompting approach, and the prompting

as mediation. Those related documents were synthesized and applied to develop the newly computerized dynamic assessment for reading literacy. The aspects of the reading literacy test relied upon the research foundation of reading literacy. As a result, the test content areas were based on the PISA 2018 reading literacy framework from OECD (2016). The reading literacy framework comprised three aspects as shown in Table 17.

Table 17 Aspects of the reading literacy test

Aspects	Meaning	Sub-aspects
Locate information	It is the ability of the student to read the texts for some specific purposes, not read the whole passage.	1) access and retrieve information within a text The task requires a student to scan a single text to retrieve target information in the form of words, phrases or numerical values.
		2) search and select the relevant task The task requires a student to search and select information from one or several texts to retrieve the most relevant information given the demands of the tasks
Understand	The ability of the student to understand the meaning of the text and draw a conclusion of the text meaning and purpose.	3) represent literal meaning The task requires a student to comprehend sentences or short passages by matching a direct or close paraphrasing of information
		4) integrate and generate inferences The task requires a student to integrate and generate inferences from simple connecting inferences to more complex coherence relationship across sentences or an entire passage in order to create a main idea, produce a summary or make inferences within the text
Evaluate and reflect	The ability of the student to reflect on the content and form of the text and to evaluate the quality and validity of the information.	5) assess quality and credibility The task requires a student to evaluate the quality and credibility of the text in terms of validity, accuracy, and bias and requires the student to evaluate the source of information such as author, content, and form
		6) reflect on content and form The task requires a student to reflect on the quality and style of the writing from their own experience to evaluate the form of writing and how the content are connected that determine the author's purpose of writing and the author's point of view
		7) detect and handle conflict The task requires a student to detect the soundness of the claims and handle conflict of the contradicting information within a text or multiple texts.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.2 Construct the instruments

The researcher developed computerized dynamic assessment for reading literacy tests as well as a pretest and a posttest for Grade 9 students. The instruments drew upon existing theoretical information on reading literacy. The procedures of test construction were as follows: 1) review the literature on reading literacy, 2) develop the table of specification, 3) select reading passages, and 4) write test items (Creswell, 2012). The test content areas were based on three dimensions, namely 1) locate information, 2) understand, and 3) reflect and evaluate. There were two research instruments in this study which included 1) the pretest and posttest and 2) the computerized dynamic assessment for reading literacy as the following:

2.1.2.1 The reading literacy pretest and posttest

Each reading literacy pretest and posttest comprised 20 items with five multiple-choice questions based on four reading passages. The pretest and posttest were parallel items. The items were written in Thai. The tests were online test administered without mediation. Reading literacy pretest was used to measure students' prior knowledge before receiving the dynamic assessment intervention, and the posttest was used to measure students' independent performance after receiving the dynamic assessment intervention. The table of specification was shown in Table 18.

Table 18 Table of specification of reading literacy pretest and posttest

Aspects	Sub-aspects	Percentage	Pretest	Posttest
Locate information	Access and retrieve information within a text	15%	3	3
	search and select the relevant task	15%	3	3
Understand	represent literal meaning	20%	4	4
	integrate and generate inferences	20%	4	4
Evaluate and reflect	assess quality and credibility	10%	2	2
	reflect on content and form	10%	2	2
	detect and handle conflict	10%	2	2
Total		100	20	20

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.2.2 Computerized dynamic assessment for reading literacy

Computerized dynamic assessment for reading literacy comprised 20 items with four reading passages. The test items consisted of five multiple-choice questions with one correct answer written in Thai. There were three sets of tests, which were parallel to each other. In each set of items, there were two items, the first-tier item and the second-tier item. The researcher developed both the first- and second-tier items for reading literacy tests. The first-tier item measured the intended construct and the second-tier item measured the rationale behind the correctly selected choice. Students were permitted to answer the second-tier item if they answered the first-tier item correctly for the first attempt. Students would obtain a point when they answer a set of two items correctly. Table of specification covered the three subscales of reading literacy. The number of items were proportionally arranged in accordance with the framework of PISA 2018 from OECD (2016). The distribution of each subscale was presented as shown in Table 19.

Table 19 Table of specification of reading literacy instruments

Subscale		Reading literacy Form C-DA1	Reading literacy Form C-DA2	Reading literacy Form C-DA3	Percentages
Locate information	Access and retrieve information within a text	3	3	3	15%
	search and select the relevant task	3	3	3	15%
Understand	represent literal meaning	4	4	4	20%
	integrate and generate inferences	4	4	4	20%
Evaluate and reflect	assess quality and credibility	2	2	2	10%
	reflect on content and form	2	2	2	10%
	detect and handle conflict	2	2	2	10%
Total		20	20	20	100

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.3 Select reading passages

The researcher chose four reading passages for a set of reading literacy tests. The passages were selected based on tasks (personal, public, occupational, and educational), texts, and level of complexity. The level of complexity of passages related to the level of transferability (near transfer, far transfer, and very far transfer) in the concept of dynamic assessment proposed by Campione and Brown (1985). In each assessment time point, the first passage was near transfer passages. The second and third passages were far transfer passages and the fourth was very far transfer passage. There were a total of 12 passages for three sets of reading literacy tests for three testing timepoints. Each passage was taken from non-curriculum books, newspapers, magazines, posters, brochures, or website texts. For the pretest and posttest, there were four passages selected in this study.

The researcher selected the reading passages in accordance with task, text, and the complexity of the passages as the following:

2.1.3.1 Task

The reading literacy passages were based on real-life situations. The tests were scenario-based for some reading purposes to achieve the particular goals as proposed in OECD (2019a). The reading situations in personal, public, occupational, and educational situations were shown in Table 20.

Table 20 The examples of tasks

Task	Example
personal	personal letters, fiction, diary-style blog, personal emails, instant messages and blogs
public	official documents, information about public events, message boards, news websites, and public notices
occupational	job advertisement in a newspaper or online job websites, and texts that provide workplace directions
educational	education texts are printed or electronic textbooks and interactive learning software

2.1.3.2 Text

Dimensions of text comprised: type, source, and format (OECD, 2019a). The researcher selected four text types presented in the reading texts including

description, narration, exposition, and argument. There were two sources including single and multiple texts. The format combined continuous texts, non-continuous texts, and mixed format as shown in Table 21.

Table 21 Three dimensions of text taxonomies

Text		Explanation	Example
type of text	description	texts with information related to properties of objects in space	a depiction of a place in a travelogue, a catalogue, a geographical map, an online flight schedule or a description of a feature, function or process in a technical manual
	narration	texts with information related to time	a report, a news story, a novel, a short story, or a play, a biography, a comic strip and a newspaper report of an event
	exposition	texts with information related to how different elements interrelate in a meaningful way	a scholarly essay, a diagram showing a model of memory, a graph of population trends, or a concept map for an entry in an online encyclopedia
	argument	texts with information related to relationship among concepts or propositions, expressing opinions or points of view	a letter to the editor, a poster advertisement, posts in an online forum or a web-based review of a book or film
source	single	texts have a definite author (or group of author), publication date, and reference	traditional printed books, the pseudonyms in a blog post, or the sponsors of a website
	multiple	texts have different authors, being published at different times, or bearing different titles	printed newspapers, textbooks, forums, customer reviews, or question and answer websites
format	continuous texts	texts formed by sentences that are organized into paragraphs	newspaper reports, essays, novels, short stories, reviews and letters
	non-continuous texts	texts composed of a number of lists or elements	lists, tables, graphs, diagrams, advertisements, schedules, catalogues, indexes and forms
	mixed texts	texts combining both continuous and non-continuous elements	magazines, reference books and reports where both continuous and non-continuous texts are placed together

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.3.3 Level of complexity

The level of complexity was based on the transferability. The concept of transfer aimed to measure student's ability to transform learning into a new context by increasing levels of complexity in reading passages. Students were assessed reading literacy over three timepoints. In each timepoint, the complexity of passages increased from near transfer passage to far transfer and very far transfer passages. As a result, there were one near transfer passage, two far transfer passages, and one very far transfer passage.

1) Near transfer was the transfer of knowledge and skill to the context or task that was similar to the previous state so that the near transfer passage was locally related to students' experience and students were familiar with particular topics well.

2) Far transfer was the transfer of knowledge and skill to the context or task that combined the new problem with the same previous state. The far transfer passage was generally a passage combining previous knowledge and new knowledge to students.

3) Very far transfer was the transfer of knowledge and skill to more complex and challenging contexts or tasks. The very far transfer passage was a passage that students hardly encountered in their daily lives but necessary to know in order to become a global citizen.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

As a result, there were four reading passages for each test form. Each test form had reading passages with different tasks, texts, and levels of complexity based on the concept of transfer as shown in Table 22.

Table 22 Selection of passages based on task, text, length, and transferability

Assessment	Passage	Transferability	Reading material	Context/task	text type	text source	text format
first assessment timepoint	1	near transfer	a short story	personal	narration	single	Continuous text
	2	far transfer	public notices	public	exposition	multiple	Mixed text
	3	far transfer	online job-related website	occupational	description	multiple	Mixed text
	4	very far transfer	Featured article	educational	argument	single	Continuous text
second assessment timepoint	5	near transfer	a short story	personal	narration	single	Continuous text
	6	far transfer	public notices	public	exposition	multiple	Mixed text
	7	far transfer	online job-related website	occupational	description	multiple	Mixed text
	8	very far transfer	Featured article	educational	argument	single	Continuous text
third assessment timepoint	9	near transfer	a short story	personal	narration	single	Continuous text
	10	far transfer	public notices	public	exposition	multiple	Mixed text
	11	far transfer	online job-related website	occupational	description	multiple	Mixed text
	12	very far transfer	Featured article	educational	argument	single	Continuous text

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

When considering the measured constructs and reading passages, items were matched with each measure construt and reading passage, test blueprint of the computerized dynamic assessment for reading literacy was developed as shown in Table 23. There were a total of three test forms. Each form comprised 20 multiple choice items, along with four passages. In locate information subscale, there were a total of six items (i.e., Items 1, 6, 7, 11, 12, and 16). Understand subscale was consisted of eight items (i.e., Items 2, 3, 8, 9, 13, 14, 17, and 18). Evaluate and reflect subscale had six items (i.e., Items 4, 5, 10, 15, 19, and 20).

Table 23 Test blueprint of instruments for computerized dynamic assessment for reading literacy

Subscale	Measured construct	Reading literacy form (Total items)			Percent	Item no.	Passage			
		C-DA1	C-DA2	C-DA3			One (Personal)	Two (Public)	Three (Occupational)	Four (Educational)
Locate information	Access and retrieve information within a text	3	3	3	15%	1	✓			
						6		✓		
						16				✓
	search and select the relevant task	3	3	3	15%	7		✓		
						11			✓	
						12			✓	
Understand	represent literal meaning	4	4	4	20%	2	✓			
						8		✓		
						13			✓	
						17				✓
	integrate and generate inferences	4	4	4	20%	3	✓			
						9		✓		
						14			✓	
						18				✓
Evaluate and reflect	assess quality and credibility	2	2	2	10%	5	✓			
						15			✓	
	reflect on content and form	2	2	2	10%	4	✓			
						19				✓
	detect and handle conflict	2	2	2	10%	10		✓		
						20				✓
Total		20	20	20	100	20				

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.4 Write test items

2.1.4.1 Test format

The researcher constructed the test that was in line with the table of specification, combining three aspects: 1) locate information, 2) understand, and 3) evaluate and reflect. The items were written in Thai language and formatted as multiple-choice questions requiring students to select one correct answer from a list of five options. The test integrated mediation during the assessment to support students and the two-tier test to prevent guessing. The tests consisted of 20 items, with five items related to each passage.

2.1.4.2 Test characteristics

The computerized dynamic assessment for reading literacy combined the second-tier item in order to prevent student's guessing. The first tier asked about the knowledge intended to measure. If a test-taker was able to solve the first-tier question correctly, he or she was allowed to answer the second-tier item related to the rationale behind the correct answer. Furthermore, because of the repeated measure in the intervention to measure growth, students were assessed over three time points. There were three sets of parallel tests, each consisting of 20 items.

2.1.4.3 Scorings

Scorings were divided into two scores based on the work of Poehner and Lantolf (2013): actual score (0,1) and mediated score (4-0). The actual score and mediated score were adapted for the computerized dynamic assessment for reading literacy in the present study.

1) Actual score was the score that a student received either one point or zero points depending on the correct first response. For those who answered a set of two items correctly at the first attempt, they received one point. For those who did not answer either first-tier item or second-tier item correctly, they obtained zero points. Student's answer score was marked as either correct (1 point) or incorrect (0 points) to represent their independent, unassisted performance.

2) Mediated score is the score represents the amount of mediating prompt students received. Mediated score ranged from 4 to 0 based on how much mediation students required to complete the test. The mediated score

was fully 4 points in case a student answered the first-tier item correctly at the first try without using any prompts. The 3 point for the mediated scores was provided in case a student answered the first-tier item correctly after receiving the first prompt. The 2-point was provided when a student could respond to the first-tier item correctly when the second prompt was provided. The 1-point score was provided in case a student answered the first-tier item correctly after receiving the third prompt. The 0-point score was provided when a student received the third prompt but they still could not answer the first-tier item correctly. The scoring was shown in Table 24.

Table 24 Scoring procedure example

First-tier attempt	Answer	Second-tier answer	Prompt needed	Actual score (0,1)	Mediated score (1-4)
First try	Right	Right		1	4
		Wrong	-	0	4
	Wrong	-	First prompt		
Second try	Right	-	-	0	3
	Wrong	-	Second prompt		
Third try	Right	-	-	0	2
	Wrong	-	Third prompt		
Fourth try	Right	-	-	0	1
	Wrong	-	Show answer	0	0

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.5 Develop the prompts

The researcher studied the theoretical concept and related documents on dynamic assessment and the concept of graduated prompting approach, which was one of the interventionist approaches of dynamic assessment, and developed the specification of prompts for reading literacy. There were four prompting methods in this study as follows: 1) instructional prompting method, 2) error-explanation prompting method, 3) mixed prompting method, and 4) verification prompting method (control condition). Each prompt differed in terms of the pattern of prompts and the content of prompts. They were developed individually for each item following the guideline of the specification of prompts.

The table of specification of prompts were developed from three different types of prompts for reading literacy. Each was different in terms of pattern of prompts and the content or prompts. The table of specification was shown as follows:

2.1.5.1 The instructional prompting

The prompt relied on the instruction to guide students to answer the questions correctly. It was arranged in a graduated way based on attempts students chose the incorrect answer, ranging from most implicit to most explicit as shown in Table 25.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 25 Table of specification of instructional prompts

Type of prompt	Levels of prompting		Dimension of reading literacy	Content of prompt
Instructional prompt	First level	General prompt: ask students the general questions related to the measured dimensions	Locate information	ask students to identify the information within the text or to select relevant information from texts
			Understand	
			Reflect and evaluate	
	Second level	Concrete prompt: guide students to the basic concepts of what to be measured	Locate information	guide students how to identify the information within the text or how to select relevant information from texts
			Understand	guide students to better understanding of the key concept within the text, to interpret the meaning of the text, to detect the main idea of the text , or to generate inferences from the text
			Reflect and evaluate	guide students to evaluate the quality and credibility of the text, to make judgments from the text, to detect and handle conflict
	Third level	Direct instruction: direct students to the specific sentence or explain students the main concept of the text	Locate information	direct students to focus on certain paragraphs or sentences
			Understand	direct students to focus on certain paragraphs or sentences, give explanation about the meaning of the text, state the key concept of the text, or state relationships within the text
			Reflect and evaluate	state the judgement of the information within the text, state the evaluation of the quality or credibility of the text, or state the conflict of the texts
	Fourth level	Explain the answer: demonstrate how to solve the problem	Locate information	provide the correct answer and demonstrate how to solve the problem
			Understand	
			Evaluate and reflect	

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.5.2 The error-explanation prompting

The prompt was provided in accordance with the choice of incorrect answer students chose, ranging from implicit to explicit as presented in Table 26.

Table 26 Table of specification of error-explanation prompts

Type of prompt	Levels of prompting		Dimension of reading literacy	Content of prompt
Error-explanation prompt	First level	**General prompt:** ask students to answer again	Locate information Understand Reflect and evaluate	▪ tell students that the chosen choice is not correct answer and ask them to answer again
	Second level	**Concrete prompt:** tell students the reason why their chosen choice is not correct	Locate information Understand Reflect and evaluate	▪ tell students the implicit reason why their chosen choice is not correct answer
	Third level	**Specific prompts:** tell students the explicit reason why their chosen choice is not correct by stating textual information or certain sentences or phrases related to the choice	Locate information Understand Reflect and evaluate	▪ tell students the explicit reason why their chosen choice is not correct by stating textual information or referring to certain sentences or phrases related to that choice
	Fourth level	**Explain the answer:** demonstrate how to solve the problem	Locate information Understand Evaluate and reflect	▪ provide the correct answer and demonstrate how to solve the problem

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.5.3 The mixed prompting

The mixture between instructional prompting and error-explanation prompting was operated. Prompting was based on the instruction to guide students to answer the question correctly and the reason why the selected choice was incorrect as depicted in Table 27.

Table 27 Table of specification of mixed prompts

Type of prompt	Levels of prompting		Dimension of reading literacy	Content of prompt
Mixed prompt	First level	**General prompt:** ask students the general questions related to the measured dimensions	Locate information	▪ tell students that the chosen choice is not correct answer
			Understand	▪ ask students to identify the information within the text or to select relevant information from texts
			Reflect and evaluate	
	Second level	**Concrete prompt:** guide students to the basic concepts of what to be measured	Locate information	▪ tell students the implicit reason why their chosen choice is not correct answer ▪ guide students how to identify the information within the text or how to select relevant information from texts
			Understand	▪ tell students the implicit reason why their chosen choice is not correct answer ▪ guide students to better understanding of the key concept within the text, to interpret the meaning of the text, to detect the main idea of the text , or to generate inferences from the text
			Reflect and evaluate	▪ tell students the implicit reason why their chosen choice is not correct answer ▪ guide students to evaluate the quality and credibility of the text, to make judgments from the text, to detect and handle conflict
	Third level	**Direct instruction:** direct students to the specific sentence or explain students the main concept of the text	Locate information	▪ tell students the explicit reason why their chosen choice is not correct by stating textual information or referring to certain sentences or phrases related to that choice ▪ direct students to focus on certain paragraphs or sentences

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Type of prompt	Levels of prompting	Dimension of reading literacy	Content of prompt
		Understand	▪ tell students the explicit reason why their chosen choice is not correct by stating textual information or referring to certain sentences or phrases related to that choice ▪ direct students to focus on certain paragraphs or sentences, give explanation about the meaning of the text, state the key concept of the text, or state relationships within the text
		Reflect and evaluate	▪ tell students the explicit reason why their chosen choice is not correct by referring to certain sentences or phrases related to that choice ▪ state the judgement of the information within the text, state the evaluation of the quality or credibility of the text, or state the conflict of the texts
	Fourth level **Explain the answer:** demonstrate how to solve the problem	Locate information	▪ provide the correct answer and demonstrate how to solve the problem
		Understand	
		Evaluate and reflect	

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.6 Validate the test content and test parallelism

Before administration, the reading literacy tests were validated by a group of experts to check the content validity and test parallelism. There were five groups of experts for the content validation. The first group of experts were asked to review the content validity and test parallelism of the reading literacy pretest and posttest. The other four groups of experts were invited to review the content validity and the parallelism of the computerized dynamic assessment for reading literacy. The instruments were separated into four sets of tests based on the similarity of tasks (personal, public, occupational, and educational). Each set of reading literacy tests comprised five experts, four of them were experts in the field of Thai language and reading literacy and the other expert was in the field of educational measurement and evaluation. They were asked to validate the test contents as follows: (1) the description of each text, items, and prompts

were correct, clear and understandable; (2) the statement of each item matched the reading literacy components defined; and (3) the prompts clearly matched the prompt guidelines defined. Moreover, they were asked to depict their agreement on each item pair of the equivalent tests in terms of the measured constructs and language use.

There were a total of twenty-five experts taking part in the content validation. The item-objective congruence (IOC) was used to conclude the experts' opinions. For the validity of content, items with the IOC higher than .5 was appropriate. Items with the IOC lower than .5 needed revision in accordance with experts' comments (Rovinelli & Hambleton, 1977). In terms of parallelism of the tests, items are parallel forms if the IOC is higher than .5. If the IOC is .5 or less than .5, the items are not in parallel forms and need revisions (Debeer et al., 2017; Kantiwong, 2017).

The equation of the item-objective congruence was as follows:

$$\text{IOC} = \frac{\sum R}{N}$$

Where, $\sum R$ = The sum scores of experts' rating opinions

N = The number of experts

With respect to the IOC index of reading literacy pretest, the results suggested that all items had an IOC index ranging from 0.6 - 1.0, except for item 3 and item10 that needed to be revised. The experts also suggested that some items needed minor changes in language use. Some Thai words should be added or changed to make it more comprehensible. In addition, the result of the IOC index of the reading literacy posttest revealed that all items had an IOC index ranging from 0.6 – 1.0, except for item 30. The experts also suggested that some items had to be revised in terms of language use. Thus, some items were revised in terms of language use. A list of experts was presented in Appendix A. The result of the IOC index of the reading literacy pretest and posttest was shown in Appendix B.

In terms of content equivalence, the results suggested that all item pairs had IOC index higher than 0.5, except for the pair of item 3 pretest and item 3 posttest that needed to be revised.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.7 Revise the test items

After reviewing the content validity and parallelism, the tests were revised based on the suggestion and comments of experts and were ready for the pilot study.

2.2 Phase 2: Validation of the research instrument

2.2.1 Administer a pilot study

Both instruments were piloted with a large group of Grade 9 students who had the similar characteristics to the participants but did not take part in the experiment. There were two groups of pilot study. The first group comprised 277 grade ninth students to validate reading literacy pretest and posttest, and the second group was 525 grade ninth students recruited to validate computerized dynamic assessment for reading literacy. The pilot study was carried out in January-March 2020. The main purpose of this pilot study was easy-to-read test content, readiness of equipment, and task difficulty. Participants were asked if any of the items were problematic, or had confusion or misinterpretation. Participants were also asked to provide direct feedback about the clarity of the items. Some items would be deleted or revised.

2.2.2 Analyze the model fit and item parameter estimates

Data were analyzed in terms of model fit as well as item parameter estimates using the "mirt" package of the freeware R (Chalmers, 2012). For model fit and model comparison, several statistical tests were applied, including the likelihood ratio test, AIC (Akaike information criterion), BIC (Bayesian information criterion) indices, deviance statistic (G^2), and RMSEA. The model with the smaller statistical tests was preferable (Bock & Aitkin, 1981; Gibbons & Hedeker, 1992). Moreover, the item parameter analysis was determined in terms of the item discrimination (a-parameter) and item difficulty (b-parameter).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

The formula of the two-parameter bifactor model equation was presented (DeMars, 2006, 2012).

$$P_i(\theta) = \frac{1}{1 + e^{D(a_{Pi}\theta_P + a_{Ti}\theta_T - d_i)}}$$

where $P_i(\theta)$ is the probability that an examinee answers item i correctly given the θ and the item parameters, (θ) is composed of θ_P and the vector θ_T of the testlet traits, a_{Pi} is the item discrimination on the primary trait, a_{Ti} is a vector of testlet discrimination parameters for item i, d_i is the item difficulty parameter, e is 2.71828, and *D* is 1.7.

For the multidimensional IRT model, the multiple item discrimination estimates are turned into multidimensional item discrimination (MDISC) estimates (Cai & Kunnan, 2018) presented as:

$$\text{MDISC} = A_i = \frac{-d}{B_i}$$

For the multidimensional IRT model, the multiple item difficulty estimates are transformed into multidimensional item difficulty (MDIFF) estimates (Cai & Kunnan, 2018) presented as:

$$\text{MDIFF} = B_i = \frac{-d}{\sqrt{a_0{}^2 + a_s{}^2}}$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Where, d is the item intercept, a_0 is the discrimination parameter on the general factor, and a_s is the discrimination parameter for the specific-related factors.

For the result of the reading literacy pretest and posttest, the 2PL bifactor model was used to examine the psychometric properties of reading literacy pretest and posttest. For the item parameter estimates, a majority of the pretest items had good multidimensional discrimination estimates. It was found that most items were at moderate level and 5 items were highly MDISC items (i.e., Items 1, 6, 7, 13, and 16). Only 2 items (i.e., Items 5 and 10) needed to be improved as its MDISC was very low discriminating value. Regarding the multidimensional difficulty estimates, all

pretest items were produced in the acceptable range, except for two items (i.e., Items 10 and 20) that had highly MDIFF values and needed to be improved.

In addition, for the multidimensional discrimination estimates of the posttest, a majority of the posttest items had good MDISC. It was found that most items were at moderate and relatively high MDISC. None of them needed to be improved. Regarding the multidimensional difficulty estimates, all posttest items were produced in the acceptable range, except for the item 24 and item 25 which had highly MDIFF values and needed to be improved.

2.2.2 Check the construct validity

Construct validity of reading literacy instruments were analyzed using the second-order confirmatory factor analysis to support whether the reading literacy instrument matched with the measured constructs. Reading literacy was considered as functions of locate information, understand, and evaluate and reflect subscales. Goodness-of-fit statistics were investigated.

2.2.3 Check the reliability

The reliability of reading literacy instruments were analyzed with pilot groups. Internal consistency was calculated using Cronbach alpha coefficient and Omega coefficient. Reliability coefficients range between 0.00 and 1.00, with a higher value indicating a higher degree of internal consistency. The desirable values should be .70 as recommended by Nunnally (1978) for the scale in the initial development and .80 for basic research scale.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Internal consistency using Cronbach's alpha coefficient and Omega coefficient were calculated for reading literacy pretest and posttest. The Cronbach's alpha reliability estimates for pretest and posttest were 0.72 and 0.77, respectively. For the Omega coefficient, the reliability estimates of pretest and posttest were 0.76 and 0.80, respectively. This indicated that the highly acceptable reliability values for the overall reading literacy pretest and posttest, which yielded highly internal reliability.

2.2.4 Check the parallelism of the reading literacy instruments

The parallel forms of the computerized dynamic assessment for reading literacy were analyzed by using the test information function (TIF) and test characteristic curve (TCC). The researchers used a root mean square deviation (RMSD) statistic to quantify the closeness among the tests. For the test form parallelism, If $RMSD_{TIF}$ and $RMSD_{TCC}$ values are smaller than .50, it is considered to be parallel test forms (Debeer et al., 2017; Kantiwong, 2017). For the item parallelism, item pairs were analyzed in terms of item information function (IF) and item characteristic curve (ICC). They were presented below.

The test parallelism was computed as follows (Ali & van Rijn, 2016; Debeer et al., 2017):

$$RMSD_{TIF} = \sqrt{\frac{\sum_{j=1}^{q}[I(\theta j)-I'(\theta j)]^2}{q}}$$

$$RMSD_{TCC} = \sqrt{\frac{\sum_{j=1}^{q}[T(\theta j)-T'(\theta j)]^2}{q}}$$

The item parallelism was computed as follows:

$$RMSD_{IIF} = \sqrt{\frac{\sum_{j=1}^{q}[i(\theta j)-i'(\theta j)]^2}{q}}$$

$$RMSD_{ICC} = \sqrt{\frac{\sum_{j=1}^{q}[t(\theta j)-t'(\theta j)]^2}{q}}$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Where, q = The number of ability levels at which the functions are evaluated

$I(\theta)$ and $I'(\theta)$ = Test information functions at a particular θ level of the assembled test and the reference test

$i(\theta)$ and $i'(\theta)$ = Item information function at a particular θ level of the assembled test and the reference test

$T(\theta)$ and $T'(\theta)$ = Expected test scores of the assembled test and the reference test

$t(\theta)$ and $t'(\theta)$ = Expected item scores of the assembled test and the reference test

For the parallelism of reading literacy pretest and posttest, the analysis of reading literacy pretest and posttest were calculated using information functions and characteristic curves along with seven ability levels (i.e. -3.0, -2.0, -1.0, 0, 1.0, 2.0, 3.0). The data contained responses from 277 students. With respect to item parallelism, it can be seen that $RMSD_{IIF}$ estimates range from 0.00 to 0.13 and the $RMSD_{ICC}$ range from 0.00-0.17, showing satisfactory closeness in terms of item information functions and item characteristic curves between the two assembled forms. Moreover, when considering the closeness of test forms, $RMSD_{TIF}$ and $RMSD_{TCC}$ were 0.44 and 0.18, respectively, representing the conformity of the two forms. Thus, both pretest and posttest were parallel in terms of test forms and individual items.

2.2.5 Select the test items for the main study

After validating the test items and checking the parallel forms of the tests, the item parameters, a-parameter (item discrimination) and b-parameter (item difficulty), were estimated from the suitable IRT model. Test items with good discrimination, along with good difficulty estimates were selected for the main study. Hence, the reading literacy instruments were revised based on the results of content validity, reliability, and item analysis.

2.2.6 Develop the computer-based instrument called computerized dynamic assessment for reading literacy

The researcher then developed computerized dynamic assessment (C-DA) for reading literacy using iSpring Suite. The program was designed to examine the students' reading literacy ability by assessing a set of reading literacy tests, combining the two-tier test and a series of graduated prompts.

2.3 Phase 3: Implementation of the computerized dynamic assessment for reading literacy

2.3.1 Allocate students into groups

After sampling a group of subjects from the general population, students in each classroom were assigned to four conditions, including 1) instructional prompting condition, 2) error-explanation prompting condition, 3) mixed prompting condition, and 4) controlled condition. Participants were assigned to groups on a random basis by using their student ID numbers so that each participant had an equal chance of appearing in any group of experiments. They received different codes to enter into the classes in Google classrooms. Each code represented different prompting methods. Each individual student could be able to join only one class based on their randomly Google classroom code. All of them were involved in the study for three testing sessions at different time points as shown in Figure 31.

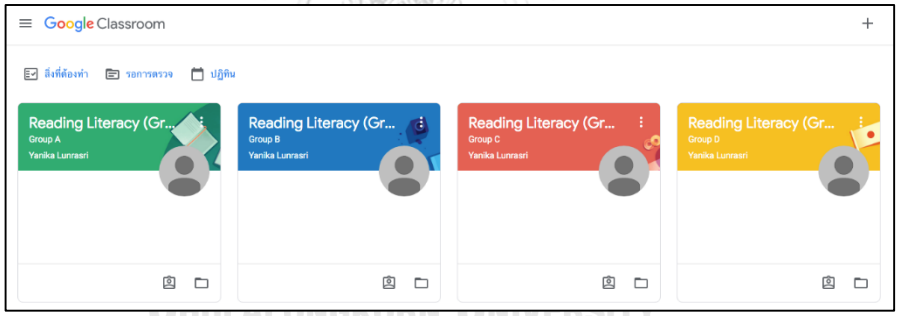


Figure 31 Classes in Google classroom for different prompting groups

2.3.2 Conduct the pretest

Reading literacy pretest was used to assess Grade 9 students' reading literacy. The test was 20 multiple-choice questions with four choices. Students were given the pretest in the first week of the study to help researchers know the level of reading literacy skills. Students had to work independently without receiving any mediation during the test. The researcher constructed a reading literacy pretest in Google Form as presented in Figure 32.

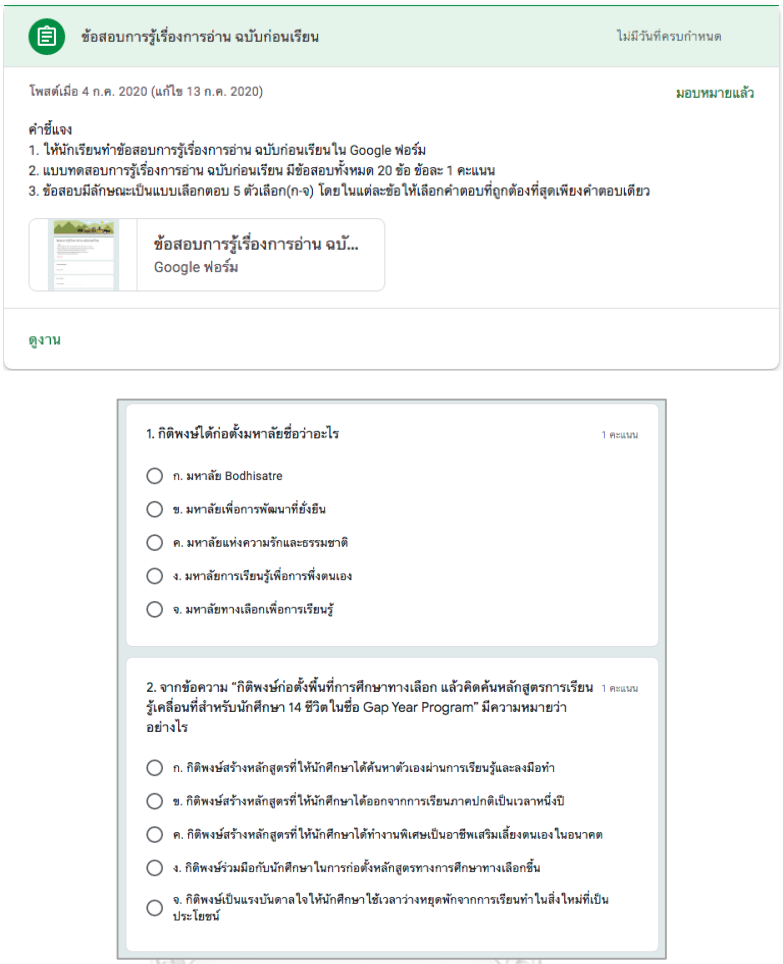


Figure 32 Reading literacy pretest in Google Form

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.3.3 Conduct the computerized dynamic assessment for reading literacy

This study adopted a quasi-experimental design that aimed to determine whether discrepancy existed between two or more treatment conditions in the educational settings (Campbell & Stanley, 1963). In order to account for threats of external or confounding variables to internal validity, statistical technique using analysis of covariance (ANCOVA) was performed by examining for differences in reading literacy scores between groups, while controlling for the effect of pretest as the covariate. Moreover, this study analyzed two-way analysis of variance (ANOVA) to investigate the effects of group differences and characteristics of samples, including gender and levels of learning achievement, on reading literacy.

Before the experiment, Students were instructed to use the computerized dynamic assessment and practiced until they felt more acquainted with the system. During the experiment, they were administered a computerized dynamic reading literacy test with 20 multiple-choice items with five choices. When they responded incorrectly, he/she received the prompt based on their group conditions. They were assessed three parallel sets of reading literacy tests for three testing timepoints. The study lasted eight weeks. The following experimental design is shown in Table 28.

Table 28 The experimental design of this study

Group A	O_{Pre}	O_{x1A}	O_{x2A}	O_{x3A}	O_{Post}
Group B	O_{Pre}	O_{x1B}	O_{x2B}	O_{x3B}	O_{Post}
Group C	O_{Pre}	O_{x1C}	O_{x2C}	O_{x3C}	O_{Post}
Group D	O_{Pre}	O_{x1D}	O_{x2D}	O_{x3D}	O_{Post}

Where,

O_{Pre} = The pretest

O_{Post} = The posttest

O_{XA} = The group received the instructional prompting condition

O_{XB} = The group received the error-explanation prompting condition

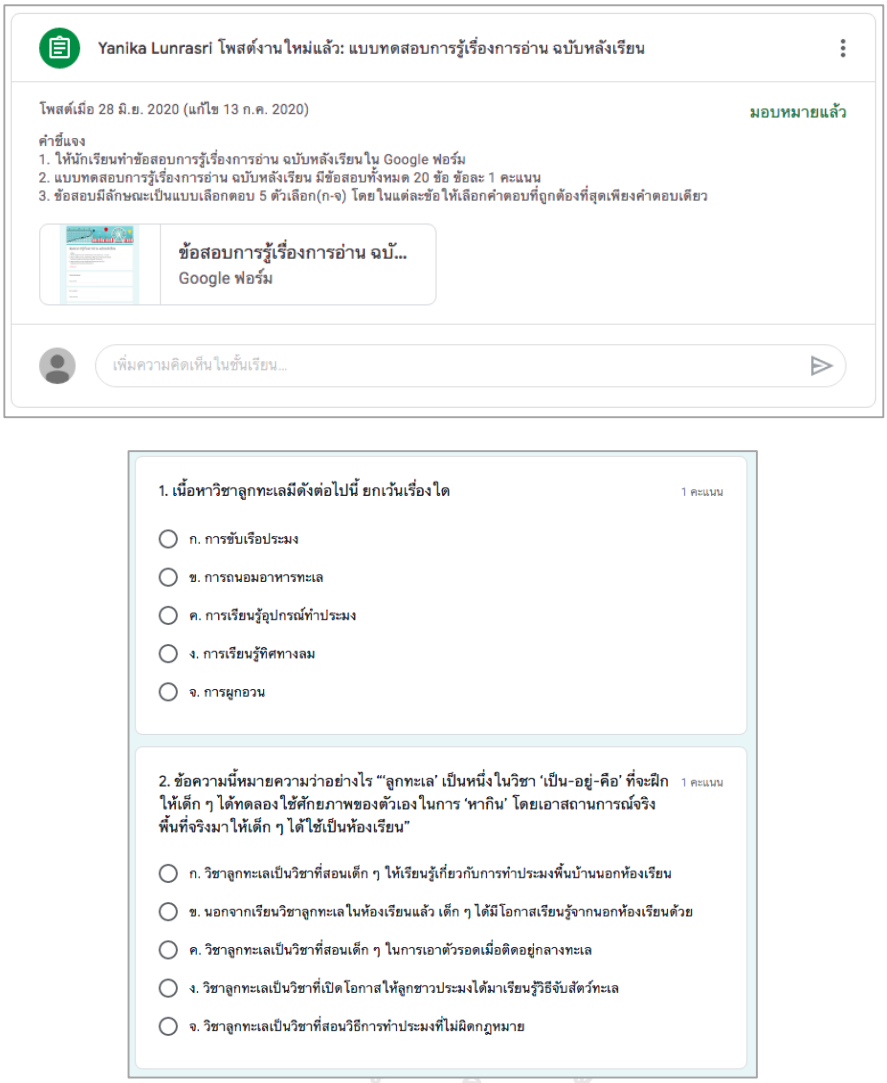
O_{XC} = The group received the mixed prompting condition

O_{XD} = The group received the verification prompting condition

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.3.4 Conduct the posttest

After finishing the training sessions, the students took a reading literacy posttest, which was parallel with the reading literacy pretest. The results of the experiment were prepared for the analysis in terms of current reading literacy ability, growth in reading literacy, and learning potential. The researcher constructed a reading literacy posttest in Google Form as presented in Figure 33.



จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Figure 33 Reading literacy posttest in Google Form

3. Data Collection

In the present study, the researcher collected the data in three periods: before the experiment, during the experiment, and after the experiment. The following section explained the data collection of each step in detail.

3.1 Before the experiment

During this procedure, the reading literacy pretest was administered to Grade 9 students at the first week of the testing session. The test lasted 20 minutes. Then, the participants were briefly introducted to the computerized dynamic assessment. Before taking the test, students were instructed to read or listen to instructions carefully on the computer and in order to make them acquainted with the system, they were allowed to practice the computerized dynamic assessment prior to the real assessment.

3.2 During the experiment

Students' reading literacy was assessed through the computerized dynamic assessment for three testing periods. The whole experiment lasted 8-9 weeks. Each time point was the four-week interval. Computerized dynamic assessment was administered in classroom context. The program recorded the number of correct and incorrect responses as well as the number of prompts required during training for each student. Participants were administered in their school computer lab or in their classrooms with computers. There was no time limit for taking the test, depending on students' reading ability, difficulties encountered, and number of prompts used. Students were allowed to work at their own pace. Researcher administered the test by herself.

3.3 After the experiment

The reading literacy posttest was administered in the final week of the study to assess the students' current ability after receiving the computerized dynamic assessment and compare each group's conditions. The posttest lasted 20 minutes. Then, the results of the experiment were further analyzed to measure growth and learning potential.

4. Data Analysis

In the present study, data were analyzed as the following:

4.1 Data analysis for instrument validation

4.1.1 Content validity of the instruments

Content validity was checked on how well the reading literacy instrument measured a theoretical construct using the item-objective congruence (IOC) index.

4.1.2 Psychometric properties of the reading literacy instruments

4.1.2.1 Model fit and model comparison

Data were analyzed in terms of model and item fits as well as item parameter estimates using the "mirt" package of the freeware R (Chalmers, 2012). The preliminary analysis was conducted to select the best fitted unidimensional IRT model for the data. The items were calibrated using 1PL and 2PL models, whereas the 3PL model was disregarded because of medium-sized data and the two-tier instruments that had lower guessing concern. Then, the better fitting model was compared with competing more complicated models, multidimensional 2PL model, testlet model and bifactor multidimensional models.

The models were compared to estimate the best-fitting model by using the goodness-of-fit indices such as -2Log likelihood values, its associated root-mean-square error of approximation (RMSEA) values as well as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The model which had the lowest -2Log likelihood, RMSEA, AIC or BIC values was considered to have the best fit and the suitable model for the dataset (Bock & Aitkin, 1981; Gibbons & Hedeker, 1992).

4.1.2.2 Item parameter analysis

The best fitted model was applied for the item analyses. For the multidimensional model, the direct scores were transformed into composite scores before interpretation by transforming the multiple item discrimination estimates into multidimensional item discrimination (MDISC) and the multiple item difficulty estimates into multidimensional item difficulty (MDIFF) (Cai & Kunnan, 2018). Based on guidelines available in Baker (2001), items that have discrimination values ranged

from 0.35- 1.69, indicating that these items have good discrimination between low and high ability students. In addition, items that have difficulty values higher than 2.00, presenting that these items have more difficulty for the test-takers as presented in Table 29.

Table 29 Labels for item discrimination and item difficulty estimates

Discrimination values		Difficulty values	
Label	Range of values	Label	Ranges of values
Very low	0.01-0.34	Very easy	< -2.00
Low	0.35 – 0.64	Easy	-2.00 - -0.51
Moderate	0.65 – 1.34	Medium	-0.51 – 0.49
High	1.35 – 1.69	Difficult	0.51 – 1.99
Very high	≥ 1.70	Very difficult	≥ 2.00

4.1.3 Construct validity

The researcher aimed to test the theorectical structures of reading literacy using second-order confirmatory factor analysis (CFA). Goodness of fit was assessed using the χ^2 statistic, the comparative fit index (CFI), the non-normed fit index (NNFI), the standardized root mean square residual (SRMR), and the root mean square error of approximation (RMSEA). Values close to 0.06 for the RMSEA and .08 for SRMR indicates a good model fit. Values of CFI and NNFI are greater than .95 as indicative of good fit (Hu & Bentler, 1999). Values greater than .90 are typically interpreted as an acceptable fit (Schumacker & Lomax, 1996). Data were analyzed using "Lavaan" package of the freeware R.

4.1.4 Reliability

The reliability was estimated based on Cronbach's alpha coeffieicnt and Omega Coefficient (ω) to measure the consistency of the test scores. In order to be acceptable, the coefficients of reliability should be greater than 0.70 (Nunnally, 1978).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.1.5 Parallelism of reading literacy instruments

Regarding the test forms, the statistical evidence for being parallel are TCC and TIF. A root mean square deviation (RMSD) statistic was used as an index to quantify the closeness of TCC and TIF between the test forms. If RMSD values were smaller than .50, both test forms were considered to be parallel (Debeer et al., 2017).

In examining parallelism at the level of the individual item, item characteristic curve (ICC) and item information function (IIF) were calculated for each individual item. Similar to the equivalence of test form, the RMSD was also calculated to determine the closeness of ICC and IF of item pairs. The items were equivalent when RMSD values were smaller than .50 (Debeer et al., 2017).

4.2 Data analysis for reading literacy, growth, and learning potential

The analysis measured current performance of reading literacy, growth of learning, and learning potential of Grade 9 students. All statistical analyses were analyzed using SPSS software version 22 and R software version 4.0.4. The flowchart of data analysis was presented in Figure 34.

4.2.1 Current performance of reading literacy

The researcher compared the current performances of Grade 9 students from reading literacy pretest and posttest scores. One-way ANCOVA was computed to determine whether the various prompting conditions had a significant effect on the posttest scores when controlling the pretest. The pretest scores served as a covariate and the posttest scores served as the dependent variable. The groups served as the independent variable. Furthermore, the researcher investigated the effect of different prompting conditions of computerized dynamic assessment (C-DA) and students' demographic data (gender and levels of learning achievement) on reading literacy posttest score when the pretest score was treated as the covariate using two-way ANCOVA. The statistical significance level was set at $p < .05$.

4.2.2 Growth

The researcher compared growth of learning of Grade 9 students receiving different prompting methods by using univariate latent growth curve and multivariate latent growth curve models from Lavaan package in R software. The

researcher estimated the univariate latent growth curve model for each measured construct and then estimated the multivariate latent growth curve model for all three constructs simultaneously. The following details disclosed how to analyze the growth in this present study.

4.2.2.1 Univariate latent growth curve model

The collecting data was divided into 4 datasets which were: 1) data about locate information, 2) data about understand, 3) data about reflect and evaluate, and 4) data about overall reading literacy. Each construct was estimated from univariate growth curves to measure how each construct changed over time.

In each construct, the univariate growth curves differed in terms of the shape of the growth curves or the basis coefficients were compared to determine which model better described changes over time. as follows: 1) the intercept only, the basis coefficients equal to 0 in the slope factor and 2) the linear growth model, the loading slope factors were fixed at values of 0, 2, and 3 in the base model. Full information maximum likelihood estimation was conducted.

Moreover, the χ^2 values and the fit measures were directly comparable for every model. For the model evaluation criteria, absolute fit statistics were used to determine how well a statistical model reflected the data which included the closeness-of-fit measures, badness-of-fit measures, and the χ^2 goodness-of-fit test. Closeness-of-fit measures comprised the Comparative Fit Index (CFI) and Tucker-Lewis index (TLI). The values of CFI and TLI should be above .90 to indicate good fit (Marsh et al., 2004, as cited in Tomas et al., 2017). Badness-of-fit measure was Root-Mean-Square Error of Approximation (RMSEA). The RMSEA should be below .05 to indicate close fit (Brown & Cudeck, 1993, as cited in Hong & Ho, 2005). The χ^2 goodness-of-fit test indicated that the model reflected the data adequately. A small Chi-square was preferred (Brown & Cudeck, 1993, as cited in Hong & Ho, 2005).

In order to analyze the data of the fitted model, multi-group analyses were conducted. In this analysis, latent growth curve models for each variable were compared by different prompting groups. The means of intercept and slope factors

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

were computed to estimate the growth as well as the variances of intercept and slope were computed to estimate the inter-individual differences. It should be noted that variance of slope should significantly be different from zero to estimate change in reading literacy.

4.2.2.2 Multivariate latent growth curve model

The multivariate latent growth curve model was used to analyze the association of the developmental trajectories of the three measured constructs. The associations between three variables were measured over time. The associative latent growth model was applied in this study to estimate 1) locate information, 2) understand, and 3) reflect and evaluate simultaneously. Each univariate latent growth model was combined to estimate the full model. In order to analyze the data, the covariance among the intercepts, the covariance among the slopes, and covariance between intercepts and slopes were estimated.

The multivariate latent growth curve model for multiple group analysis was used for the group comparison. The estimation of growth was based on the full model of each group. The covariance among intercepts, covariance among slopes, and covariance between the intercepts and slopes were compared among four groups.

4.2.3 Measurement of learning potential

Measurement of learning potential was measured in terms of availability scores, mediated scores, and levels of promptings. The researcher compared the learning potential of Grade 9 students from different group conditions.

4.2.3.1 Availability score

Availability score was applied to estimate learning potential by using the computation of the Michaelis-Menten growth model from NLME package in R.

The Michaelis-Menten growth model measured upper capacity and rate performance in reading literacy over three time points. Capacity was estimated as the random-effect parameter, allowing each individual to have their estimated capacity. Both ability and capacity scores were directly estimated from the model, but availability score was calculated by subtracting the estimated ability from the subject-specific capacity estimate.

As a result, in order to compare availability scores from different prompting conditions, the researcher analyzed the differences in availability scores over three time points using one-way MANOVA. The independent variable was groups received different promptings, and the dependent variables were the availability score measured at the first time, the availability score measured at the second time, and the availability score measured at the third time.

4.2.3.2 Mediated scores

The sum total of the mediated score of each student was counted to investigate the amount of mediation students needed to reach their learning attainment. As a result, the mean differences of mediated scores among four group conditions were estimated by using two-way mixed ANOVA at the significant level of .05. The group conditions and time were counted as independent variables and the mediated score was a dependent variable.

4.2.3.3 Levels of promptings

Levels of promptings of reading literacy as well as levels of promptings of reading literacy subscales were investigated using one-way MANOVA at the significant level of .05. The independent variable was prompting condition, and the dependent variables were first level of prompting, second level of prompting, third level of prompting, and fourth level of prompting.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

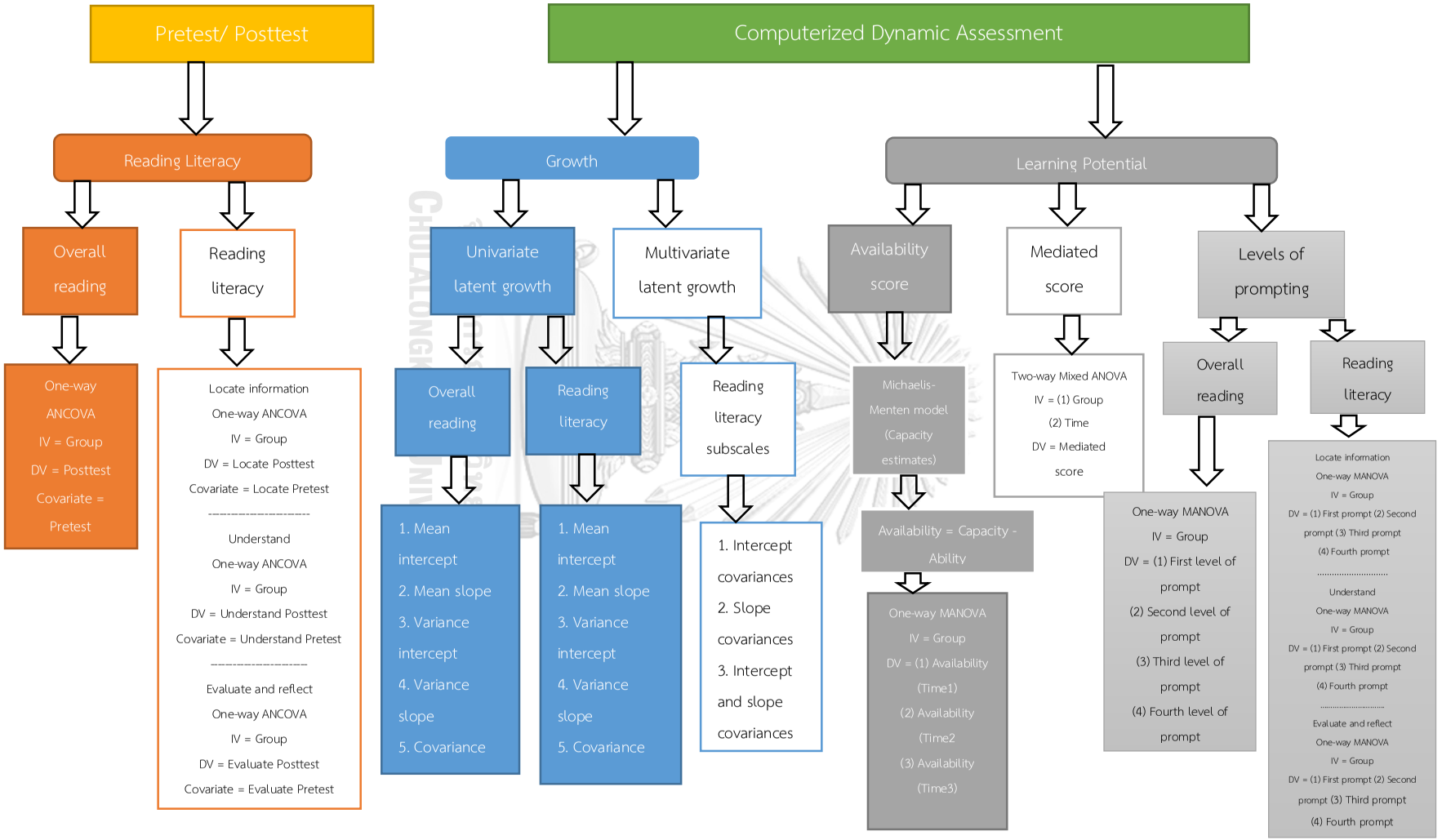


Figure 34 Flowchart of data analysis of the study

CHAPTER FOUR
RESULTS

This study aimed to develop and validate the research instruments as well as to measure reading literacy, growth of learning, and learning potential of Grade 9 students by comparing different promptings of the computerized dynamic assessment on reading literacy. The results were presented as the following:

Part 1: Results of the development and validation of the computerized dynamic assessment for reading literacy of Grade 9 students

1.1 Results of the development of the computerized dynamic assessment for reading literacy of Grade 9 students

1.2 Results of the validation of the computerized dynamic assessment for reading literacy of Grade 9 students

1.3 Results of the parallelism of the computerized dynamic assessment for reading literacy of Grade 9 students

Part 2: Results of the effects of different promptings of the computerized dynamic assessment on reading literacy ability, growth of learning, and learning potential of Grade 9 students

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1 Results of the effects of different promptings of the computerized dynamic assessment on reading literacy ability of Grade 9 students

2.2 Results of the effects of different promptings of the computerized dynamic assessment on growth of learning of Grade 9 students

2.3 Results of the effects of different promptings of the computerized dynamic assessment on learning potential of Grade 9 students

Part 1: Results of the Development and Validation of the Computerized Dynamic Assessment for Reading Literacy of Grade 9 Students

1.1. Results of the development of the computerized dynamic assessment for reading literacy of Grade 9 students

The instrument was developed in accordance with the conceptual theory of Zone of Proximal Development from Vygotsky and Graduated prompt approach. The test format was five-multiple-choice questions with one correct answer. The content was adapted from the framework of PISA reading literacy 2018 proposed by OECD (2019a). The tested contents were composed of three dimensions of reading literacy, including 1) locate information, 2) understand, and 3) evaluate and reflect. There were 20 items combined with four reading passages in each testing session. The test was constructed as three parallel tests to measure in three different time points.

The instrument was constructed with the combination of the graduated prompting approach and the two-tier test. In each question, there were a set of two items. The first-tier item was the general question to measure student's reading literacy and the second-tier item was used to avoid guessing. When a student answered the first question correctly, he or she was allowed to answer the second-tier question automatically which asked about the rationale why this choice was the correct answer. This item identified how a student clearly understood the first-tier question and was able to find the correct answer without guessing. Those students who answered a set of two items correctly would get a point and moved to another question. If a student answered the first-tier question incorrectly, he or she would be provided the prompting patterns based on their group conditions. The graduated prompting approach was used as the mediation providing to help students learn and find the correct answer during the assessment procedure. After receiving the first prompt, students were allowed to answer the first-tier question again for the second attempt and if they answered correctly, they would continue to do the second-tier item. However, if they still could not answer correctly, the second prompt was provided. The system continued until the students answered a set of two items and were able to move to the next question. The structure of computerized dynamic assessment was presented in Figure 35.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

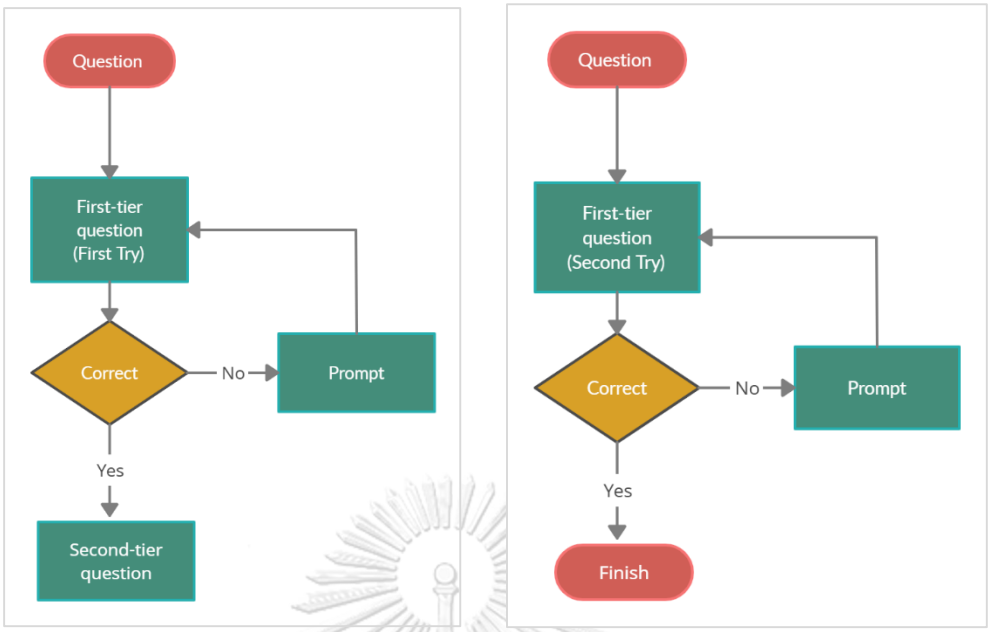


Figure 35 Structure of computerized dynamic assessment

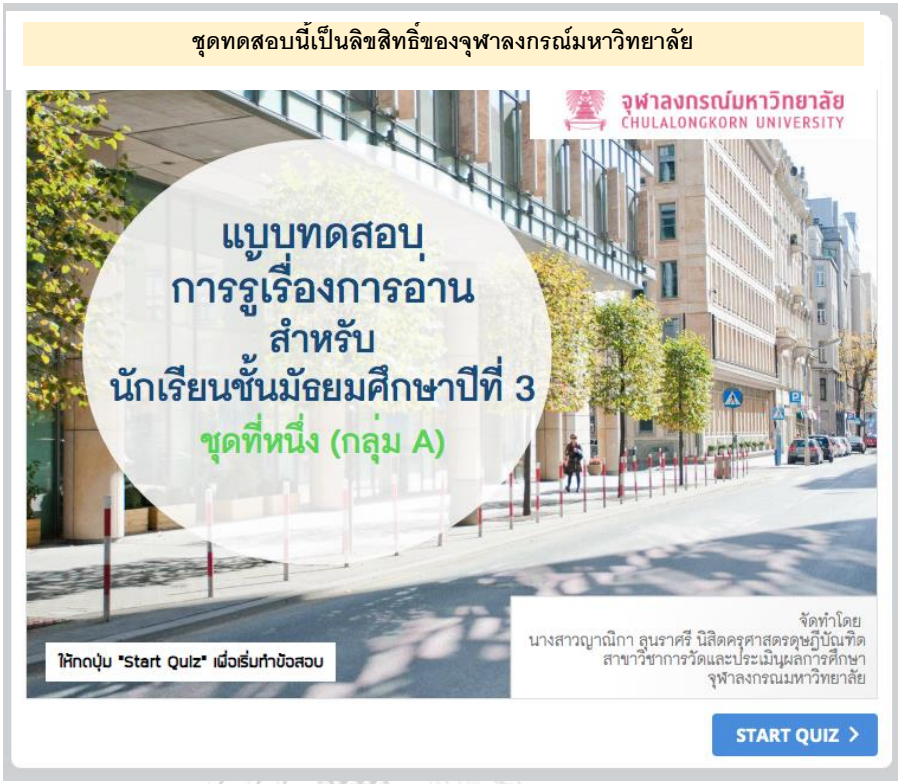
1.1.1 The design of computerized dynamic assessment for reading literacy

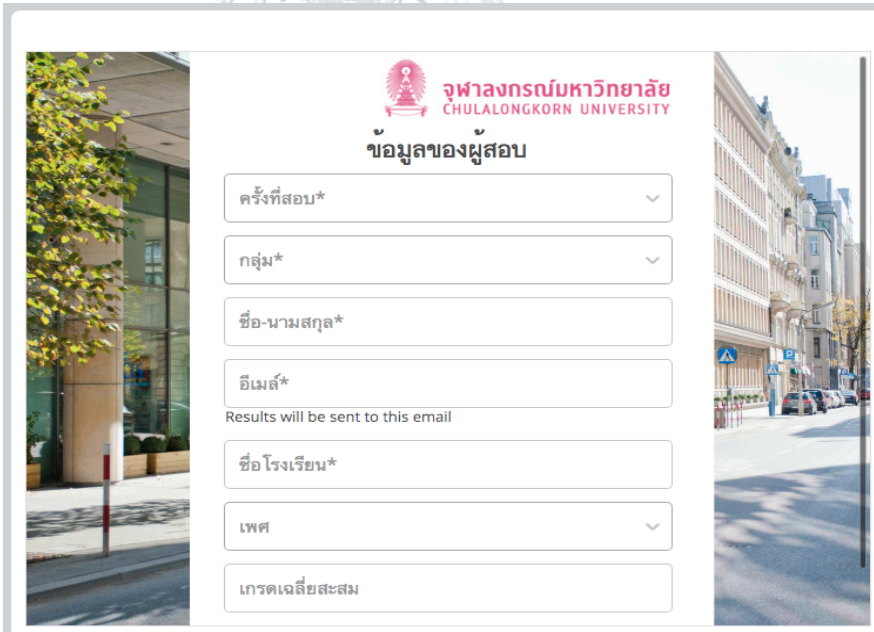
The design of computerized dynamic assessment was constructed and designed from the iSpring Suite 9. The program was interactive and online. The researcher generated shared links that limited the accessibility and published the link in Google classroom for each prompting group. The program consisted of 1) user profile interface, 2) dynamic assessment interface, and 3) scoring report interface.

1.1.1.1 User profile interface

Students were required to fill their names, school names, emails, and groups before entering the main assessment interface as presented. Moreover, students were required to follow the directions of the computerized dynamic assessment as depicted in Figure 36.

ชุดทดสอบนี้เป็นลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย





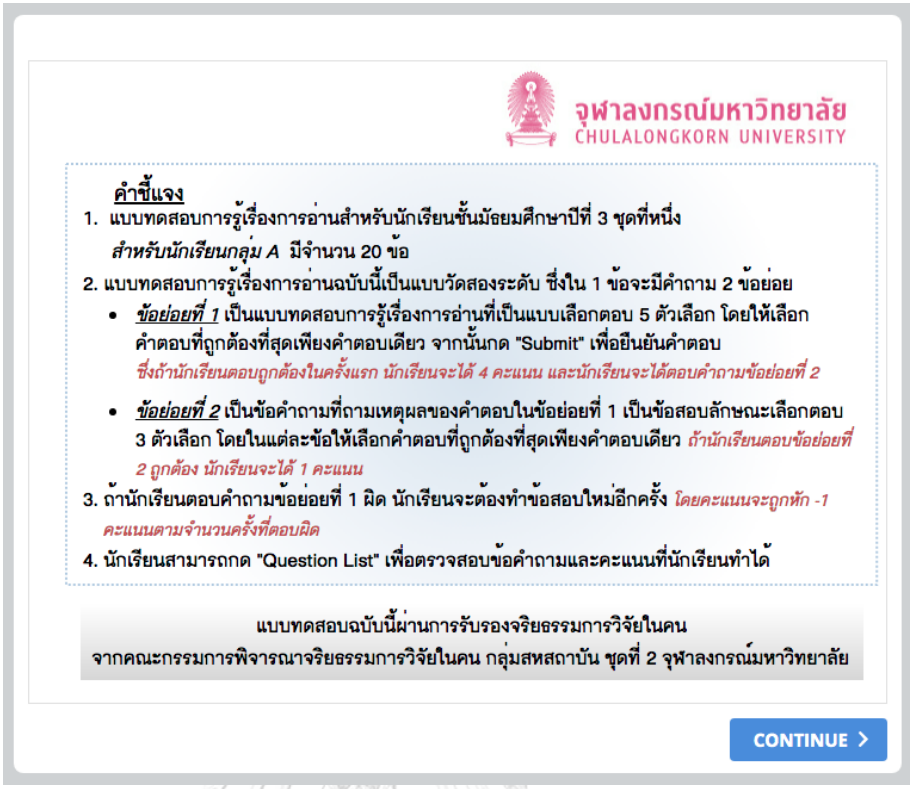


Figure 36 User interface

1.1.1.2 Dynamic assessment interface

The researcher adapted the theoretical basis of the Graduated prompting approach proposed by Campione and Brown (1985). Each reading literacy instrument comprised 20 two-tier multiple choice items. Students were randomly chosen into 4 groups, including the instructional prompt group, the error-explanation prompt group, the mixed prompt group, and the verification prompt group (control condition). Reading passages were presented before students took the testlet, and they could re-read the text. In each item, they were required to answer first-tier items. Students were allowed to answer the second-tier item if they answered the first-tier correctly; otherwise, they were provided prompts, ranging from implicit to explicit as shown in Figures 37-40. The given instructional prompts were depicted in Figure 41. The error-explanation prompts were in Figure 42. Moreover, the mixed prompts were seen in Figure 43.

Figure 37 Example of reading passage

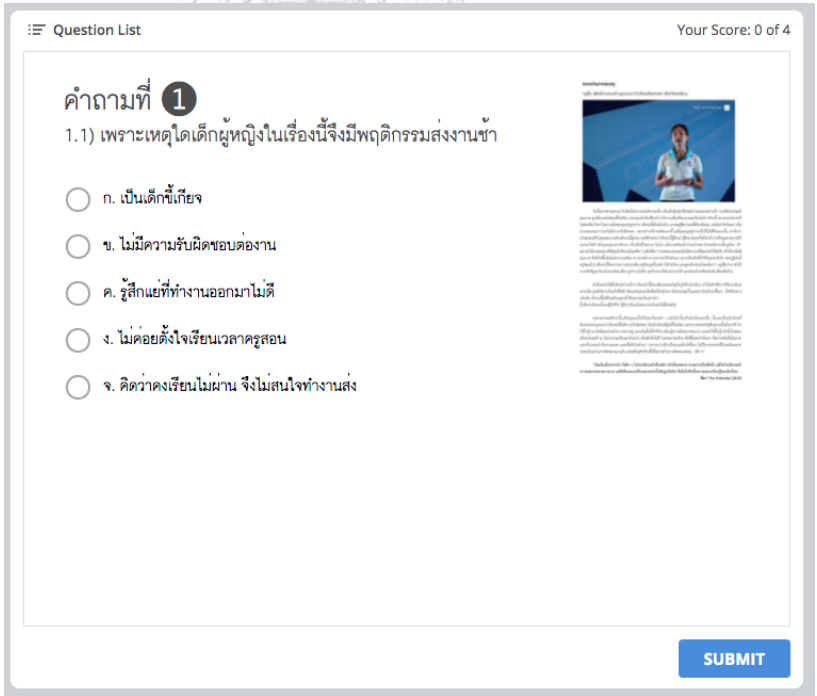


Figure 38 Example of reading literacy item (first-tier item)

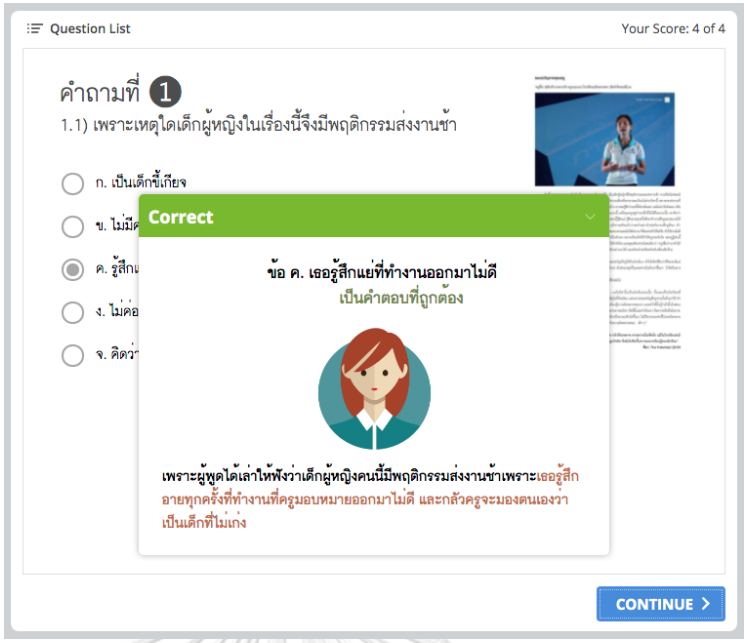


Figure 39 The prompt given when student answered first-tier item correctly

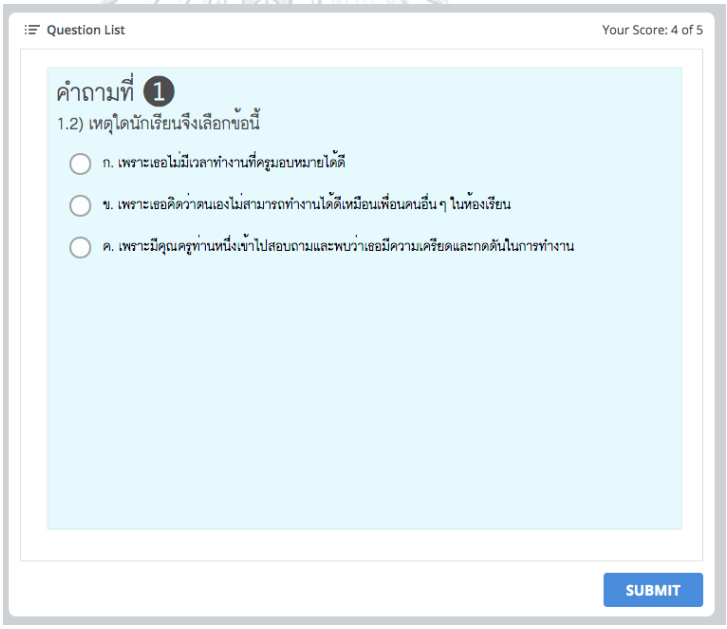


Figure 40 Example of reading literacy item (second-tier item)

1) The interface of the Instructional prompt of computerized dynamic assessment

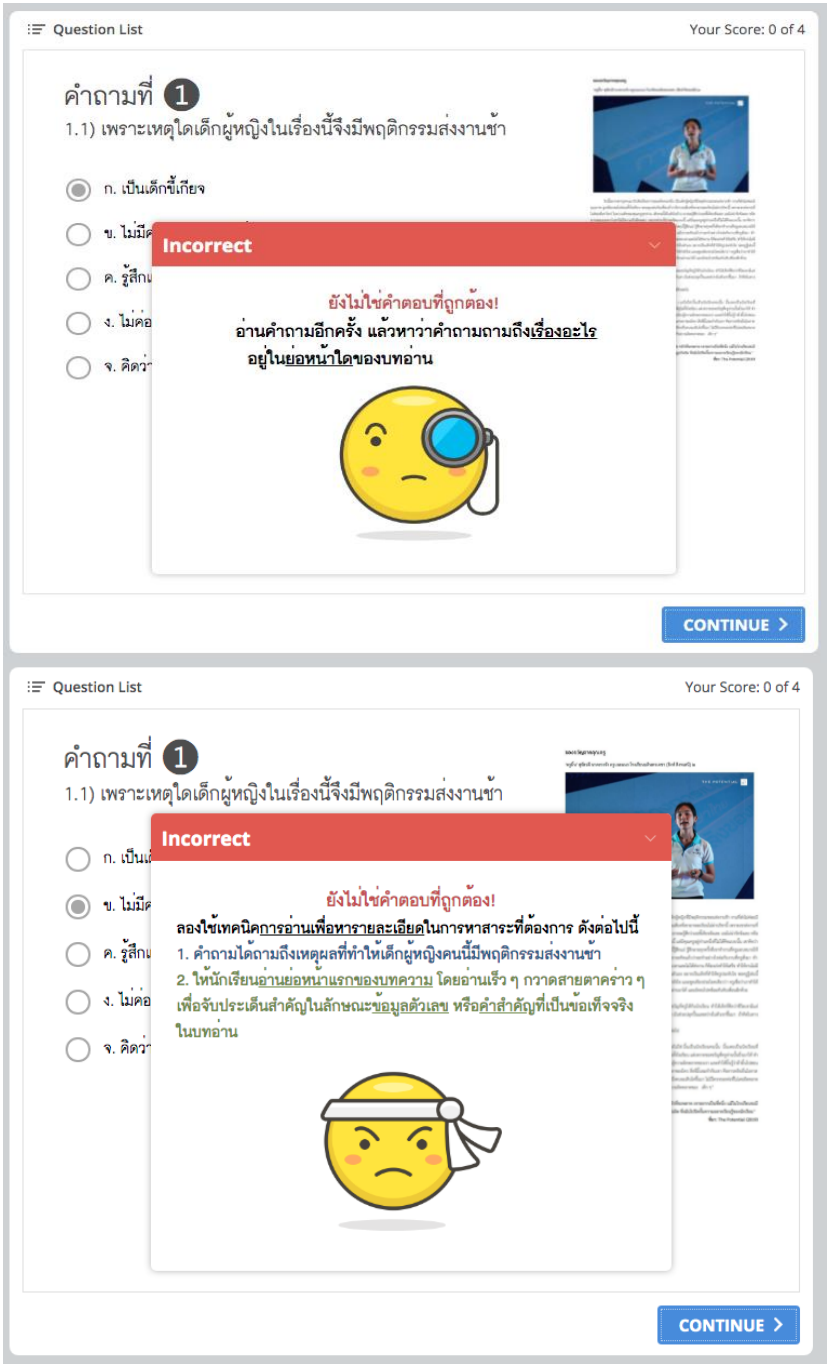


Figure 41 Instructional prompt of computerized dynamic assessment

2) The interface of error-explanation prompt of computerized dynamic assessment

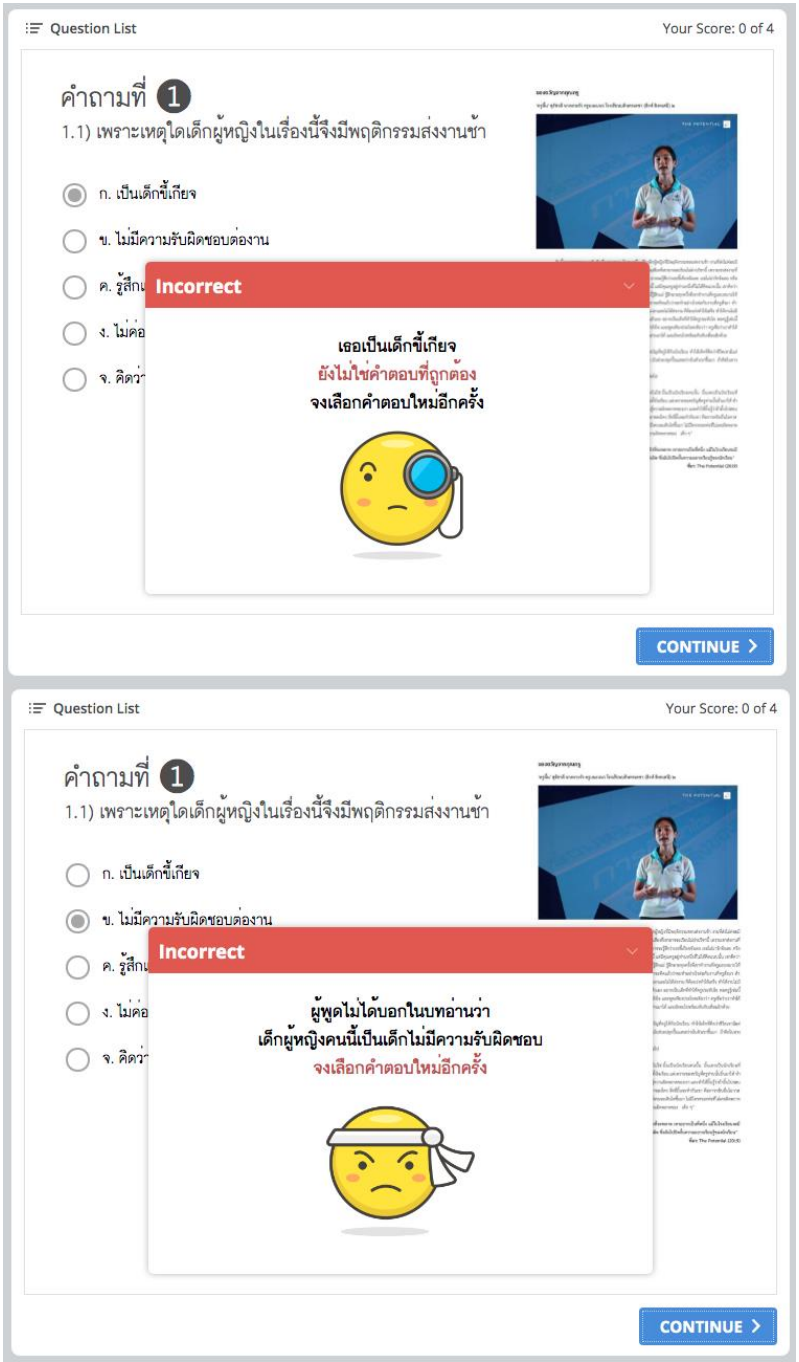


Figure 42 Error-explanation prompt of computerized dynamic assessment

3) The interface of mixed prompt of computerized dynamic assessment

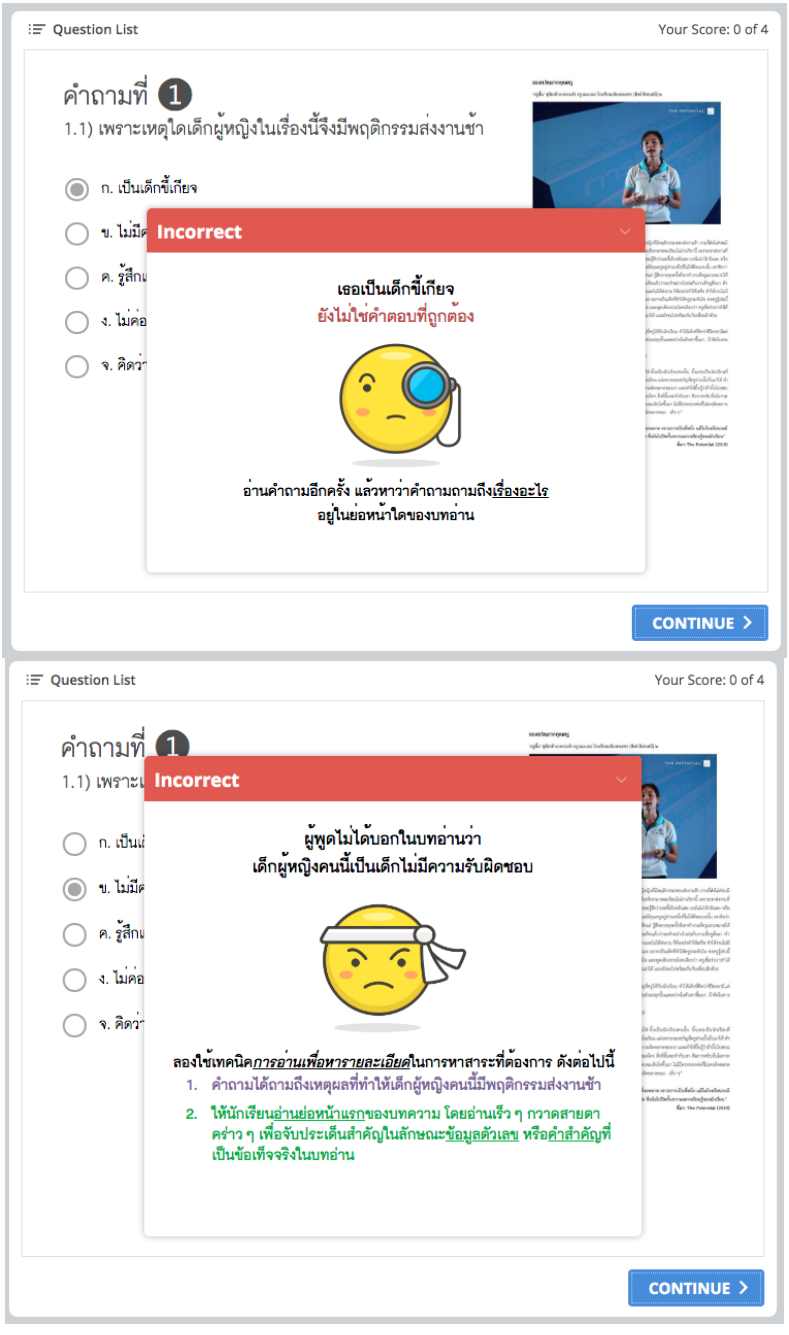
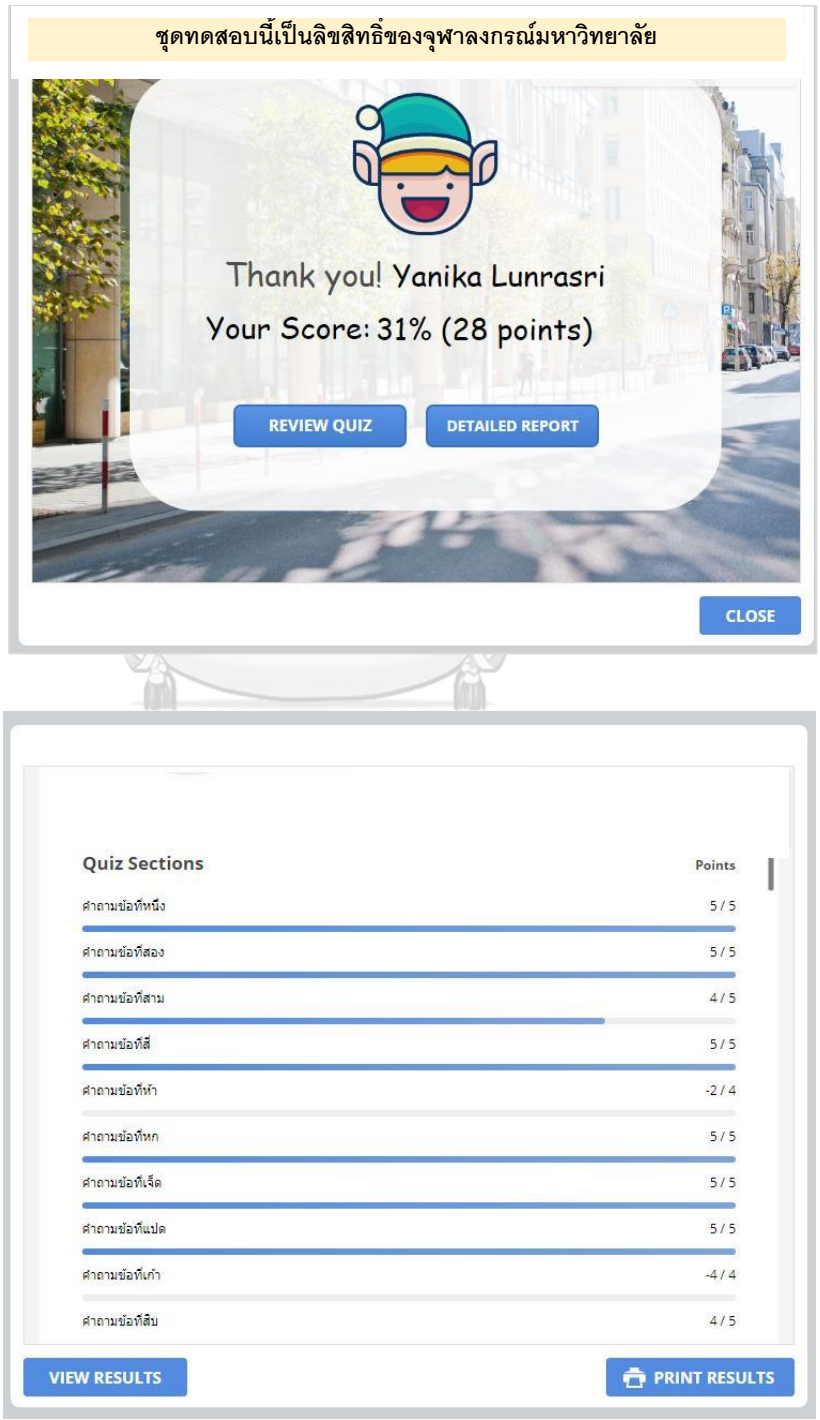


Figure 43 Mixed prompt of computerized dynamic assessment

1.1.1.3 Scoring interface

As depicted in Figure 44, the interface automatically generated the overall scores and each individual item for students. The amount of mediation students used was also counted in each item. Moreover, the scoring report was automatically sent to the email of the students and each individual score was sent to the researcher directly.



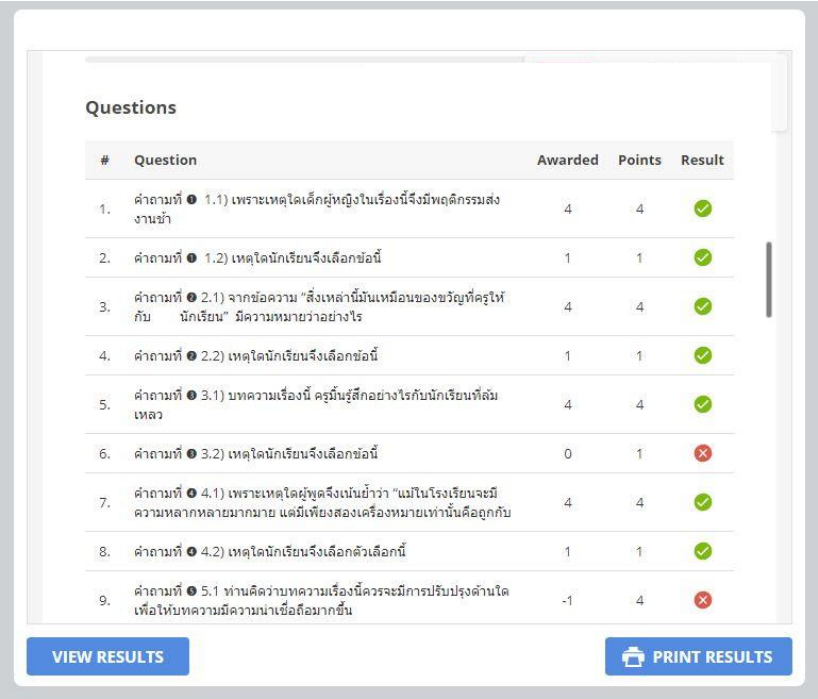


Figure 44 Scoring report of computerized dynamic assessment

1.1.2 The types of prompts

There were many different types of prompts. In this research, three prompting conditions were studied, including instructional prompting, error-explanation prompting, and mixed prompting. Each group of students received different prompting conditions. Each prompt differed in terms of the pattern of prompts and the content of prompts. Examples of questions in locate information, understand, and evaluate and reflect were shown in Figure 45-47.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

คำถามการรู้เรื่องการอ่านข้อที่ 1	
คำถามระดับที่หนึ่ง	**คำถามระดับที่สอง**
1.1 เพราะเหตุใดเด็กผู้หญิงในเรื่องนี้จึงมีพฤติกรรมส่งงานช้า ก. เป็นเด็กขี้เกียจ ข. ไม่มีความรับผิดชอบต่องาน *ค. รู้สึกแย่ที่ทำงานออกมาไม่ดี ง. ไม่ค่อยตั้งใจเรียนเวลาครูสอน จ. คิดว่าคงเรียนไม่ผ่าน จึงไม่สนใจทำงานส่ง	1.2 เหตุใดนักเรียนจึงเลือกข้อนี้ (ค) ก. เพราะเธอไม่มีเวลาทำงานที่ครูมอบหมายได้ดี ข. เพราะเธอคิดว่าตนเองไม่สามารถทำงานได้ดีเหมือนเพื่อนคนอื่น ๆ ในห้องเรียน *ค. เพราะมีคุณครูท่านหนึ่งเข้าไปสอบถามและพบว่าเธอมีความเครียดและกดดันในการทำงาน
สมรรถนะที่วัด: การระบุสารสนเทศในเนื้อเรื่อง (Locate information) **สมรรถนะย่อย**: เข้าถึงและค้นคืนสารสนเทศภายในบทอ่าน (Access and retrieve information within a text)	

Figure 45 Example of reading literacy item in locate information subscale (Item 1)

คำถามการรู้เรื่องการอ่านข้อที่ 2	
คำถามระดับที่หนึ่ง	**คำถามระดับที่สอง**
2.1 จากข้อความ ***"สิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน"*** มีความหมายว่าอย่างไร ก. ครูได้มอบของขวัญให้กับนักเรียน *ข. ครูให้กำลังใจและเชื่อมั่นในตัวของนักเรียน ค. ครูให้โอกาสนักเรียนส่งงานช้ากว่ากำหนดได้ ง. ครูทำให้นักเรียนไม่เครียดเวลาที่สอนในชั้นเรียน จ. ครูเปิดโอกาสให้นักเรียนได้ทำในสิ่งที่ตนเองสนใจ	2.2 เหตุใดนักเรียนจึงเลือกข้อนี้ (ข) ก. เพราะครูทำหน้าที่เป็นที่ปรึกษาให้แก่นักเรียนยามมีปัญหาเดือดร้อน *ข. เพราะครูเป็นผู้ที่ให้โอกาสแก่นักเรียนยามนักเรียนทำผิดพลาด ค. เพราะครูเป็นผู้ที่รักและค่อยดูแลนักเรียนเหมือนสมาชิกภายในครอบครัว
สมรรถนะที่วัด: การมีความเข้าใจในเนื้อเรื่อง (Understand) **สมรรถนะย่อย**: แสดงถึงความเข้าใจในความหมายที่แท้จริงของเนื้อเรื่อง (Represent literal meaning)	

Figure 46 Example of reading literacy item in understand subscale (Item 2)

คำถามการรู้เรื่องการอ่านข้อที่ 4	
คำถามระดับที่หนึ่ง	**คำถามระดับที่สอง**
4.1 เพราะเหตุใดผู้พูดจึงเน้นย้ำว่า ***"แม้ในโรงเรียนจะมีความหลากหลายมากมาย แต่มีเพียงสองเครื่องหมายเท่านั้นคือถูกกับผิด ซึ่งมันไปปิดกั้นความอยากเรียนรู้ของนักเรียน"*** *ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง ข. เพื่อเปรียบเทียบความแตกต่างระหว่างสิ่งที่นักเรียนทำถูกและสิ่งที่นักเรียนทำผิด ค. เพื่อเสนอแนะวิธีการเรียนรู้ที่ถูกต้องให้แก่นักเรียนนำไปปฏิบัติ ง. เพื่อสนับสนุนครูในการส่งเสริมนักเรียนให้เรียนรู้ได้ด้วยตนเอง จ. เพื่อเป็นกำลังใจให้แก่คุณครูในการเรียนรู้จากความผิดพลาดของตนเอง	4.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้ ก. เพราะผู้พูดไม่ต้องการให้นักเรียนเป็นเด็กไม่ดี ข. เพราะผู้พูดต้องการสะท้อนปัญหาของเด็กนักเรียนที่ล้มเหลวในโรงเรียน ค. เพราะผู้พูดต้องการให้โอกาสนักเรียนในการเรียนรู้จากความผิดพลาดของตน
สมรรถนะหลัก: การประเมินและสะท้อนความคิดเห็น (Evaluate and reflect) **สมรรถนะย่อย**: สะท้อนให้เห็นถึงเนื้อหาและรูปแบบ (Reflect on content and form)	

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Figure 47 Example of reading literacy item in evaluate and reflect subscale (Item 4)

1.1.2.1 Instructional prompting method

The prompt or mediation helped a student to better understand the text or the question. The instructional prompt was arranged from most implicit to most explicit prompts. When a student answered incorrectly for the first time, he or she would be provided the most implicit instructional prompt and was allowed to answer the question again. If a student answered incorrectly for the second try, the second instructional prompt was more explicit. If a student still could not answer correctly, the third instructional prompt was provided with very explicit information. If a student could not respond correctly, the fourth instructional prompt was

provided to show the correct answer and explain why that choice was correct. The pattern of prompting was based on the incorrect attempt of students as presented in Figures 48-50.

การออกแบบวิธีการชี้แนะ (promptings)	
1. การชี้แนะโดยการสอน (Instructional prompting)	
ตอบผิด ครั้งที่ 1 (Prompt 1)	**ยังไม่ใช่คำตอบที่ถูกต้อง!** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร อยู่ในย่อหน้าใดของบทอ่าน
ตอบผิด ครั้งที่ 2 (Prompt 2)	**ยังไม่ใช่คำตอบที่ถูกต้อง!** ลองใช้เทคนิคการอ่านเพื่อหารายละเอียดในการหาสาระที่ต้องการ ดังต่อไปนี้ 1. คำถามได้ถามถึงเหตุผลที่ทำให้เด็กผู้หญิงคนนี้มีพฤติกรรมส่งงานช้า 2. ให้นักเรียนอ่านย่อ**หน้าแรก**ของบทความ โดย**อ่านเร็ว ๆ กวาดสายตาคร่าว ๆ** เพื่อจับประเด็นสำคัญในลักษณะข้อมูลตัวเลข หรือคำสำคัญที่เป็น**ข้อเท็จจริง**ในบทอ่าน
ตอบผิด ครั้งที่ 3 (Prompt 3)	**ยังไม่ใช่คำตอบที่ถูกต้อง!** สาระที่ต้องการคือ**เหตุผลที่ทำให้เด็กผู้หญิงมีพฤติกรรมส่งงานช้า** ประเด็นที่เกี่ยวข้องจะอยู่ในย่อหน้าแรกคือตอนที่ผู้พูดได้เล่าถึงครูท่านหนึ่งที่เข้าไปพูดคุยและสอบถามนักเรียนในเรื่องนี้ ตามข้อความข้างล่างนี้ (บรรทัดที่ 4-6) "แต่มีคุณครูอยู่ท่านหนึ่งที่ไม่ได้คิดแบบนั้น เขาคิดว่าน่าจะลองเข้าไปคุยสอบถามกับเด็กคนนี้ดูก่อนจนได้คำตอบว่าเด็กคนนี้รู้สึกแย่ รู้สึกอายทุกครั้งที่เขาทำงานที่ครูมอบหมายให้ออกมาไม่ดี กลัวคุณครูมองว่าตัวเขา เป็นเด็กที่ไม่ฉลาด ไม่เก่ง"
ตอบผิด ครั้งที่ 4 (Prompt 4)	**คำตอบที่ถูกต้องคือ ข้อ ค. เธอรู้สึกแย่ที่ทำงานออกมาไม่ดี** ผู้พูดได้เล่าให้ฟังว่าเด็กผู้หญิงคนนี้มีพฤติกรรมส่งงานช้าเพราะ**เธอรู้สึกอายทุกครั้งที่ทำงานที่ครูมอบหมายออกมาไม่ดี และกลัวครูจะมองตนเองว่าเป็นเด็กที่ไม่เก่ง**

Figure 48 Instructional prompting method in locate information (Item 1)

การออกแบบวิธีการชี้แนะ (promptings)	
1. การชี้แนะโดยการสอน (Instructional prompting)	
ตอบผิด ครั้งที่ 1 (Prompt 1)	**ยังไม่ใช่คำตอบที่ถูกต้อง!** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร และประโยคดังกล่าวอยู่ในตำแหน่งใดของบทอ่าน
ตอบผิด ครั้งที่ 2 (Prompt 2)	**ยังไม่ใช่คำตอบที่ถูกต้อง!** ลองใช้เทคนิคการอ่านแปลความเพื่อหาความหมายของประโยค ดังต่อไปนี้ 1. ผู้พูดได้กล่าวว่า**สิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน** 2. ให้นักเรียนอ่านย่อ**หน้าแรกและย่อหน้าที่สอง**ของบทความ หาตำแหน่งของประโยคดังกล่าว 3. ให้นักเรียนแปลความหมายแฝงของคำว่า "ของขวัญ" และความหมายของประโยคดังกล่าวโดยพิจารณา**บริบทของข้อความ** โดยอ่านประโยคสนับสนุนก่อนหน้า และประโยคสนับสนุนที่ตามหลัง
ตอบผิด ครั้งที่ 3 (Prompt 3)	**ยังไม่ใช่คำตอบที่ถูกต้อง! ผู้พูดได้กล่าวว่าสิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน** ซึ่งประโยคนี้อยู่ใน***ย่อหน้าที่สอง*** ของบทอ่าน ซึ่งการแปลความหมายของประโยคสามารถพิจารณาจากบริบทข้อความจากประโยคก่อนหน้าได้ เนื่องจากผู้พูดกล่าวคำว่า **"สิ่งเหล่านี้"** แทนประโยคสนับสนุนที่ผู้พูดได้เล่าไว้ก่อนหน้าแล้ว ตามข้อความข้างล่างนี้ (บรรทัดที่ 9-11) "ครูคิดแล้วว่าเด็กคนนี้ต้องการความช่วยเหลือ ครูจึงปลุกปั้นพลัง ให้กำลังใจ และพูดเพียงประโยคเดียวว่า ครูเชื่อว่าเราทำได้ จากเด็กที่ดูจะเรียนไม่จบพร้อมเพื่อน ดูทำงานไม่ทันสุดท้ายเขาก็เรียนผ่านมาได้ แถมยังจบไปพร้อมกันกับเพื่อนอีกด้วย"
ตอบผิด ครั้งที่ 4 (Prompt 4)	**คำตอบที่ถูกต้องคือ ข้อ ข. ครูให้กำลังใจและเชื่อมั่นในตัวของนักเรียน** เพราะผู้พูดได้เล่าว่า ครูคิดว่าเด็กต้องการความช่วยเหลือ ครูจึงได้ให้กำลังใจนักเรียนและแสดงความเชื่อมั่นในตัวนักเรียนโดยพูดว่า **"ครูเชื่อว่าเราทำได้"** ซึ่งสิ่งเหล่านี้เป็นเหมือนของขวัญที่ครูให้กับนักเรียน ทำให้เด็กนักเรียนรู้สึกว่าชีวิตตัวเองมีคุณค่าขึ้นมา ทำให้สามารถเรียนผ่านและจบพร้อมกับเพื่อนได้ ซึ่งคำว่า "ของขวัญ" จึงเป็นความหมายแฝงของการที่ครูได้ให้กำลังใจและเชื่อมั่นในตัวของนักเรียน

Figure 49 Instructional prompting method in understand subscale (Item 2)

การออกแบบวิธีการชี้แนะ (promptings)	
1. การชี้แนะโดยการสอน (Instructional prompting)	
ตอบผิด ครั้งที่ 1 (Prompt 1)	**ยังไม่ใช่คำตอบที่ถูกต้อง! ยังไม่ใช่คำตอบที่ถูกต้อง!** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร หาว่าข้อความที่ผู้พูดกล่าวเน้นอยู่ในย่อหน้าใดของบทอ่าน
ตอบผิด ครั้งที่ 2 (Prompt 2)	**ยังไม่ใช่คำตอบที่ถูกต้อง!** ลองใช้เทคนิคการสะท้อนความคิดเห็นในเชิงคุณภาพและรูปแบบการเขียน เพื่อหาสาระที่ต้องการดังต่อไปนี้ 1. คำถามได้ถามเกี่ยวกับจุดมุ่งหมายในการพูด ซึ่งต้องอาศัยความรู้ทั่วไป มุมมองหรือทัศนคติของนักเรียนมาเชื่อมโยงกับสิ่งที่อ่านในเนื้อเรื่อง โดยข้อความที่ผู้พูดกล่าวเน้นอยู่ในย่อหน้าที่สี่ 2. เริ่มจากหามุมมองของผู้พูด โดยหาว่าใครเป็นผู้เล่าเรื่อง ผู้เล่าเรื่องที่อยู่ในเหตุการณ์หรือผู้เล่าเรื่องที่มิได้อยู่ในเหตุการณ์ ผู้เล่าเรื่องมีความสัมพันธ์กับเรื่องในลักษณะใด และผู้เล่าเรื่องกำลังเล่าให้ใครฟัง 3. จากนั้นให้นักเรียนดูข้อความที่ผู้พูดกล่าวเน้นสะท้อนจุดมุ่งหมายในการพูดอย่างไร เช่น เพื่อเล่าเรื่อง เพื่อชี้แจง/อธิบายวิธี เพื่อแสดงความคิดเห็น เพื่อโน้มน้าวใจ เพื่อโต้แย้ง เป็นต้น
ตอบผิด ครั้งที่ 3 (Prompt 3)	**ยังไม่ใช่คำตอบที่ถูกต้อง!** โดยมุมมองของผู้พูดคือเป็นบุคคลที่อยู่ในเหตุการณ์ ผู้พูดเป็นผู้เล่าเรื่องและกำลังเล่าเรื่องให้ครูท่านอื่น ๆ ฟัง เนื้อเรื่องเล่าถึงครูท่านหนึ่งที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อน และตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก ผู้พูดจึงกล่าวเน้นว่ามันเป็นเรื่องน่ากลัวที่เด็กในโรงเรียนจะกลัวผิดพลาด เพราะโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้น ซึ่งปิดกั้นความอยากเรียนรู้ของเด็ก ให้นักเรียนพิจารณาประโยคในคำถามอีกครั้ง และหาว่าผู้พูดมีจุดมุ่งหมายในการพูดอย่างไร
ตอบผิด ครั้งที่ 4 (Prompt 4)	**คำตอบที่ถูกต้องคือ ข้อ ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง** สิ่งที่อ่านในเนื้อเรื่องเป็นเรื่องของครูที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อนและตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก **ดังนั้นเนื้อเรื่องจึงเกี่ยวข้องกับเด็กที่เคยทำผิดพลาด ซึ่งจุดมุ่งหมายของผู้พูดคือเพื่อแสดงความคิดเห็นว่าให้นักเรียนลองผิดลองถูกได้ด้วยตนเอง เนื่องจากโรงเรียนมีเพียงสิ่งที่ถูกและผิดเท่านั้น ซึ่งทำให้เป็นการปิดกั้นเด็กในการเรียนรู้ เพราะเด็กกลัวทำผิด**

Figure 50 Instructional prompting method in evaluate and reflect subscale (Item 4)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.1.2.2 Error-explanation prompting method

The prompt or mediation was based on the choice a student selected and the number of attempts a student answered incorrectly. If a student chooses the incorrect answer choice, he or she would receive the error-explanation prompt provided depending on the selected response, ranging from implicit to explicit. Examples of an error-explanation prompting method were presented in Figures 51-53.

การออกแบบวิธีการชี้แนะ (promptings)				
2. การชี้แนะโดยการอธิบายข้อผิดพลาด (Error-explanation prompting)				
ตอบผิดครั้งที่ 1 Prompt 1	ก	ข	ง	จ
	เธอเป็นเด็กขี้เกียจ **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง	เธอไม่มีความรับผิดชอบต่องาน **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง	เธอไม่ค่อยตั้งใจเรียนเวลาครูสอน **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง	เธอคิดว่าคงเรียนไม่ผ่านวิชานี้อยู่แล้ว จึงไม่สนใจทำงานส่ง **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง
ตอบผิดครั้งที่ 2 Prompt 2	ก	ข	ง	จ
	ผู้พูดไม่ได้บอกในบทอ่านว่าเด็กผู้หญิงคนนี้เป็นเด็กขี้เกียจ จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้บอกในบทอ่านว่าเด็กผู้หญิงคนนี้เป็นเด็กไม่มีความรับผิดชอบ จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้บอกในบทอ่านว่าเด็กผู้หญิงคนนี้เป็นเด็กไม่ตั้งใจเรียน จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้บอกในบทอ่านว่าเด็กผู้หญิงคนนี้ไม่สนใจทำงานส่ง เพราะเธอคิดว่าเธอคงเรียนไม่ผ่านวิชานี้อยู่แล้ว จงเลือกคำตอบใหม่อีกครั้ง
ตอบผิดครั้งที่ 3 Prompt 3	ก	ข	ง	จ
	ผู้พูดไม่ได้บอกว่าเด็กผู้หญิงคนนี้ส่งงานช้าเพราะเป็นเด็กขี้เกียจ เพราะผู้พูดกล่าวว่า แม้ครูส่วนใหญ่คิดว่าเด็กผู้หญิงคนนี้เป็นเด็กขี้เกียจ แต่มีครูท่านหนึ่งที่ไม่ได้คิดแบบนั้น จึงลองเข้าไปพูดคุยและ สอบถามเด็กดู จงเลือกคำตอบใหม่อีกครั้ง	**ผู้พูดไม่ได้บอกว่าเด็กผู้หญิงคนนี้ส่งงานช้าเพราะเป็นเด็กไม่มีความรับผิดชอบ** เพราะผู้พูดกล่าวว่า แม้ครูส่วนใหญ่คิดว่าเด็กผู้หญิงคนนี้เป็นเด็กที่ไม่มีความรับผิดชอบ แต่มีครูท่านหนึ่งที่ไม่ได้คิดแบบนั้น จึงลองเข้าไปพูดคุยและ สอบถามเด็กดู จงเลือกคำตอบใหม่อีกครั้ง	**ผู้พูดไม่ได้บอกว่าเด็กผู้หญิงคนนี้ส่งงานช้าเพราะไม่ค่อยตั้งใจเรียน** เพราะผู้พูดกล่าวว่า แม้ครูส่วนใหญ่คิดว่าเธอไม่น่ารัก และดูเหมือนไม่ค่อยตั้งใจเรียน แต่มีครูท่านหนึ่งที่ไม่ได้คิดแบบนั้น จึงลองเข้าไปพูดคุย และ สอบถามเด็กดู จงเลือกคำตอบใหม่อีกครั้ง	**ผู้พูดไม่ได้บอกว่าเด็กผู้หญิงคนนี้ส่งงานช้าเพราะเธอคิดว่าคงเรียนไม่ผ่านวิชานี้อยู่แล้วจึงไม่สนใจทำงานส่ง** เพราะ ผู้พูดกล่าวว่า มีครูท่านหนึ่งที่เข้าไปสอบถามเด็กดู และพบว่าความจริงแล้วเธอไม่ได้คิดว่าจะไม่ผ่านวิชานี้ จงเลือกคำตอบใหม่อีกครั้ง
ตอบผิดครั้งที่ 4 Prompt 4	ก	ข	ง	จ
	คำตอบที่ถูกต้องคือ ข้อ ค. รู้สึกแย่ที่ทำงานออกมาไม่ดี ผู้พูดได้เล่าให้ฟังว่าเด็กผู้หญิงคนนี้มีพฤติกรรมส่งงานช้าเพราะ**เธอรู้สึกอายทุกครั้งที่ทำงานที่ครูมอบหมายออกมาไม่ดี และกลัวครูจะมองตนเองว่าเป็นเด็กที่ไม่เก่ง**	**คำตอบที่ถูกต้องคือ ข้อ ค. รู้สึกแย่ที่ทำงานออกมาไม่ดี** ผู้พูดได้เล่าให้ฟังว่าเด็กผู้หญิงคนนี้มีพฤติกรรมส่งงานช้าเพราะ**เธอรู้สึกอายทุกครั้งที่ทำงานที่ครูมอบหมายออกมาไม่ดี และกลัวครูจะมองตนเองว่าเป็นเด็กที่ไม่เก่ง**	**คำตอบที่ถูกต้องคือ ข้อ ค. รู้สึกแย่ที่ทำงานออกมาไม่ดี** ผู้พูดได้เล่าให้ฟังว่าเด็กผู้หญิงคนนี้มีพฤติกรรมส่งงานช้าเพราะ**เธอรู้สึกอายทุกครั้งที่ทำงานที่ครูมอบหมายออกมาไม่ดี และกลัวครูจะมองตนเองว่าเป็นเด็กที่ไม่เก่ง**	**คำตอบที่ถูกต้องคือ ข้อ ค. รู้สึกแย่ที่ทำงานออกมาไม่ดี** ผู้พูดได้เล่าให้ฟังว่าเด็กผู้หญิงคนนี้มีพฤติกรรมส่งงานช้าเพราะ**เธอรู้สึกอายทุกครั้งที่ทำงานที่ครูมอบหมายออกมาไม่ดี และกลัวครูจะมองตนเองว่าเป็นเด็กที่ไม่เก่ง**

Figure 51 Error-explanation prompting method in locate information subscale (Item 1)

การออกแบบวิธีการชี้แนะ (promptings)				
2. การชี้แนะโดยการอธิบายข้อผิดพลาด (Error-explanation prompting)				
ตอบผิดครั้งที่ 1 Prompt 1	**ก**	**ค**	**ง**	**จ**
	ครูได้มอบของขวัญให้กับนักเรียน **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง	ครูให้โอกาสนักเรียนส่งงานช้ากว่ากำหนดได้ **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง	ครูทำให้นักเรียนไม่เครียดเวลาที่สอนในชั้นเรียน **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง	ครูเปิดโอกาสให้นักเรียนได้ทำในสิ่งที่ตนเองสนใจ **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง
ตอบผิดครั้งที่ 2 Prompt 2	**ก**	**ค**	**ง**	**จ**
	ผู้พูดไม่ได้พูดถึงครูที่มอบของขวัญให้แก่นักเรียนในบทอ่านเลย จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้พูดถึงครูที่ให้โอกาสนักเรียนส่งงานช้ากว่ากำหนดได้ในบทอ่านเลย จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้พูดถึงครูที่ทำให้นักเรียนไม่เครียดเวลาที่สอนในชั้นเรียนในบทอ่านเลย จงเลือกคำตอบใหม่อีกครั้ง	แม้ผู้พูดจะสนับสนุนความอยากเรียนรู้ของนักเรียน แต่ผู้พูดไม่ได้พูดถึงครูที่เปิดโอกาสให้นักเรียนได้ทำสิ่งที่ตนเองสนใจในบทอ่าน จงเลือกคำตอบใหม่อีกครั้ง
ตอบผิดครั้งที่ 3 Prompt 3	**ก**	**ค**	**ง**	**จ**
	ของขวัญที่ครูให้กับนักเรียนไม่ได้หมายถึงครูมอบกล่องของขวัญให้กับนักเรียน เพราะคำว่า “ของขวัญ” เป็นความหมายแฝงของผู้พูด ไม่ใช่ความหมายโดยตรง ซึ่งผู้พูดเล่าว่าตนเองเคยเป็นนักเรียนที่ล้มเหลวมาก่อน และได้ของขวัญจากครู ทำให้มีวันนี้ได้ โดยผู้พูดไม่ได้กล่าวถึงครูมอบของขวัญให้กับนักเรียนเลย จงเลือกคำตอบใหม่อีกครั้ง	**ของขวัญที่ครูให้กับนักเรียนไม่ได้หมายถึงครูให้โอกาสนักเรียนส่งงานช้ากว่ากำหนด** เพราะคำว่า “ของขวัญ” เป็นความหมายแฝงของผู้พูด ไม่ใช่ความหมายโดยตรง ซึ่งผู้พูดเล่าว่าตนเองเคยเป็นนักเรียนที่ล้มเหลวมาก่อน และได้ของขวัญจากครู ทำให้มีวันนี้ได้ โดยผู้พูดไม่ได้กล่าวถึงครูให้นักเรียนส่งงานช้ากว่ากำหนดเลย จงเลือกคำตอบใหม่อีกครั้ง	**ของขวัญที่ครูให้กับนักเรียนไม่ได้หมายถึงครูทำให้นักเรียนไม่เครียดเวลาที่สอนในชั้นเรียน**เพราะคำว่า “ของขวัญ” เป็นความหมายแฝงของผู้พูด ไม่ใช่ความหมายโดยตรง ซึ่งผู้พูดเล่าว่าตนเองเคยเป็นนักเรียนที่ล้มเหลวมาก่อน และได้ของขวัญจากครู ทำให้มีวันนี้ได้ โดยผู้พูดไม่ได้กล่าวถึงครูที่ทำให้นักเรียนไม่เครียดเวลาที่สอนในชั้นเรียน เลย จงเลือกคำตอบใหม่อีกครั้ง	**ของขวัญที่ครูให้กับนักเรียนไม่ได้หมายถึงครูที่เปิดโอกาสให้นักเรียนได้ทำสิ่งที่ตนเองสนใจ** เพราะคำว่า “ของขวัญ” เป็นความหมายแฝงของผู้พูด ไม่ใช่ความหมายโดยตรง ซึ่งผู้พูดเล่าว่าตนเองเคยเป็นนักเรียนที่ล้มเหลวมาก่อน และได้ของขวัญจากครู ทำให้มีวันนี้ได้ ซึ่งแม้ผู้พูดจะสนับสนุนความอยากเรียนรู้ของนักเรียน แต่ผู้พูดไม่ได้กล่าวถึงครูที่เปิดโอกาสให้นักเรียนได้ทำสิ่งที่ตนเองสนใจ จงเลือกคำตอบใหม่อีกครั้ง
ตอบผิดครั้งที่ 4 Prompt 4	**ก**	**ค**	**ง**	**จ**
	คำตอบที่ถูกต้องคือ ข้อ ข. ครูให้กำลังใจและเชื่อมั่นในตัวของนักเรียน เพราะผู้พูดได้เล่าว่า ครูคิดว่าเด็กต้องการความช่วยเหลือ ครูจึงได้ให้กำลังใจนักเรียนและแสดงความเชื่อมั่นในตัวนักเรียนโดยพูดว่า “ครูเชื่อว่าเราทำได้” ซึ่งสิ่งเหล่านี้เป็นเหมือนของขวัญที่ครูให้กับนักเรียน ทำให้เด็กนักเรียนรู้สึกว่าชีวิตตัวเองมีคุณค่าขึ้นมา ทำให้สามารถเรียนผ่านและจบพร้อมกับเพื่อนได้ ซึ่งคำว่า “ของขวัญ” จึงเป็นความหมายแฝงของการที่ครูได้ให้กำลังใจและเชื่อมั่นในตัวของนักเรียน	**คำตอบที่ถูกต้องคือ ข้อ ข. ครูให้กำลังใจและเชื่อมั่นในตัวของนักเรียน** เพราะผู้พูดได้เล่าว่า ครูคิดว่าเด็กต้องการความช่วยเหลือ ครูจึงได้ให้กำลังใจนักเรียนและแสดงความเชื่อมั่นในตัวนักเรียนโดยพูดว่า “ครูเชื่อว่าเราทำได้” ซึ่งสิ่งเหล่านี้เป็นเหมือนของขวัญที่ครูให้กับนักเรียน ทำให้เด็กนักเรียนรู้สึกว่าชีวิตตัวเองมีคุณค่าขึ้นมา ทำให้สามารถเรียนผ่านและจบพร้อมกับเพื่อนได้ ซึ่งคำว่า “ของขวัญ” จึงเป็นความหมายแฝงของการที่ครูได้ให้กำลังใจและเชื่อมั่นในตัวของนักเรียน	**คำตอบที่ถูกต้องคือ ข้อ ข. ครูให้กำลังใจและเชื่อมั่นในตัวของนักเรียน** เพราะผู้พูดได้เล่าว่า ครูคิดว่าเด็กต้องการความช่วยเหลือ ครูจึงได้ให้กำลังใจนักเรียนและแสดงความเชื่อมั่นในตัวนักเรียนโดยพูดว่า “ครูเชื่อว่าเราทำได้” ซึ่งสิ่งเหล่านี้เป็นเหมือนของขวัญที่ครูให้กับนักเรียน ทำให้เด็กนักเรียนรู้สึกว่าชีวิตตัวเองมีคุณค่าขึ้นมา ทำให้สามารถเรียนผ่านและจบพร้อมกับเพื่อนได้ ซึ่งคำว่า “ของขวัญ” จึงเป็นความหมายแฝงของการที่ครูได้ให้กำลังใจและเชื่อมั่นในตัวของนักเรียน	**คำตอบที่ถูกต้องคือ ข้อ ข. ครูให้กำลังใจและเชื่อมั่นในตัวของนักเรียน** เพราะผู้พูดได้เล่าว่า ครูคิดว่าเด็กต้องการความช่วยเหลือ ครูจึงได้ให้กำลังใจนักเรียนและแสดงความเชื่อมั่นในตัวนักเรียนโดยพูดว่า “ครูเชื่อว่าเราทำได้” ซึ่งสิ่งเหล่านี้เป็นเหมือนของขวัญที่ครูให้กับนักเรียน ทำให้เด็กนักเรียนรู้สึกว่าชีวิตตัวเองมีคุณค่าขึ้นมา ทำให้สามารถเรียนผ่านและจบพร้อมกับเพื่อนได้ ซึ่งคำว่า “ของขวัญ” จึงเป็นความหมายแฝงของการที่ครูได้ให้กำลังใจและเชื่อมั่นในตัวของนักเรียน

Figure 52 Error-explanation prompting method in understand subscale (Item 2)

การออกแบบวิธีการชี้แนะ (promptings)				
2. การชี้แนะโดยการอธิบายข้อผิดพลาด (Error-explanation prompting)				
ตอบผิดครั้งที่ 1 Prompt 1	ข	ค	ง	จ
	เพื่อเปรียบเทียบความแตกต่างระหว่างสิ่งที่นักเรียนทำถูกและสิ่งที่นักเรียนทำผิด **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง	เพื่อเสนอแนะวิธีการเรียนรู้ที่ถูกต้องให้แก่นักเรียนนำไปปฏิบัติ **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง	เพื่อสนับสนุนครูในการส่งเสริมนักเรียนให้เรียนรู้ได้ด้วยตนเอง **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง	เพื่อเป็นกำลังใจให้แก่คุณครูในการเรียนรู้จากความผิดพลาดของตนเอง **ยังไม่ใช่คำตอบที่ถูกต้อง** จงเลือกคำตอบใหม่อีกครั้ง
ตอบผิดครั้งที่ 2 Prompt 2	ข	ค	ง	จ
	แม้ผู้พูดบอกว่ามีแค่สองเครื่องหมายคือถูกกับผิด แต่ผู้พูดไม่ได้กล่าวเปรียบเทียบสิ่งที่นักเรียนทำถูกหรือทำผิด จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้กล่าวถึงวิธีการเรียนรู้ที่ถูกต้องให้แก่นักเรียนนำไปปฏิบัติ จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้กล่าวถึงการส่งเสริมนักเรียนให้เรียนรู้ได้ด้วยตนเอง จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้กล่าวถึงการเป็นกำลังใจให้แก่คุณครูในการเรียนรู้จากความผิดพลาดของตนเอง จงเลือกคำตอบใหม่อีกครั้ง
ตอบผิดครั้งที่ 3 Prompt 3	ข	ค	ง	จ
	ผู้พูดไม่ได้มีจุดมุ่งหมายเพื่อเปรียบสิ่งที่นักเรียนทำถูกหรือทำผิด เพราะผู้พูดกล่าวว่าโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้นทำให้นักเรียนกลัวการทำผิดพลาด ถึงแม้ผู้พูดบอกว่ามีแค่สองเครื่องหมายคือถูกกับผิด แต่ผู้พูดไม่ได้กล่าวเปรียบเทียบสิ่งที่นักเรียนทำถูกหรือทำผิด จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้มีจุดมุ่งหมายเสนอแนะวิธีการเรียนรู้ที่ถูกต้องให้แก่นักเรียน เพราะผู้พูดกล่าวว่าโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้นทำให้นักเรียนกลัวการทำผิดพลาด ซึ่งผู้พูดไม่ได้กล่าวถึงวิธีการเรียนรู้ที่ถูกต้องให้แก่นักเรียนนำไปปฏิบัติ จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้มีจุดมุ่งหมายเพื่อสนับสนุนครูในการเรียนรู้ด้วยตนเองของนักเรียน เพราะผู้พูดกล่าวว่าโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้นทำให้นักเรียนกลัวการทำผิดพลาด ซึ่งผู้พูดไม่ได้กล่าวถึงการส่งเสริมนักเรียนให้เรียนรู้ได้ด้วยตนเอง จงเลือกคำตอบใหม่อีกครั้ง	ผู้พูดไม่ได้มีจุดมุ่งหมายเพื่อให้กำลังใจคุณครู เพราะผู้พูดกล่าวว่าโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้นทำให้นักเรียนกลัวการทำผิดพลาด ซึ่งผู้พูดไม่ได้หมายถึงการให้กำลังใจคุณครู จงเลือกคำตอบใหม่อีกครั้ง
ตอบผิดครั้งที่ 4 Prompt 4	ข	ค	ง	จ
	คำตอบที่ถูกต้องคือ ข้อ ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง สิ่งที่อ่านในเนื้อเรื่องเป็นเรื่องของครูที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อนและตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก **ดังนั้นเนื้อเรื่องจึงเกี่ยวข้องกับเด็กที่เคยทำผิดพลาด ซึ่งจุดมุ่งหมายของผู้พูดคือเพื่อแสดงความคิดเห็นว่าให้นักเรียนลองผิดลองถูกได้ด้วยตนเอง เนื่องจากโรงเรียนมีเพียงสิ่งที่ถูกและผิดเท่านั้น ซึ่งทำให้เป็นการปิดกั้นเด็กในการเรียนรู้ เพราะเด็กกลัวทำผิด**	**คำตอบที่ถูกต้องคือ ข้อ ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง** สิ่งที่อ่านในเนื้อเรื่องเป็นเรื่องของครูที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อนและตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก **ดังนั้นเนื้อเรื่องจึงเกี่ยวข้องกับเด็กที่เคยทำผิดพลาด ซึ่งจุดมุ่งหมายของผู้พูดคือเพื่อแสดงความคิดเห็นว่าให้นักเรียนลองผิดลองถูกได้ด้วยตนเอง เนื่องจากโรงเรียนมีเพียงสิ่งที่ถูกและผิดเท่านั้น ซึ่งทำให้เป็นการปิดกั้นเด็กในการเรียนรู้ เพราะเด็กกลัวทำผิด**	**คำตอบที่ถูกต้องคือ ข้อ ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง** สิ่งที่อ่านในเนื้อเรื่องเป็นเรื่องของครูที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อนและตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก **ดังนั้นเนื้อเรื่องจึงเกี่ยวข้องกับเด็กที่เคยทำผิดพลาด ซึ่งจุดมุ่งหมายของผู้พูดคือเพื่อแสดงความคิดเห็นว่าให้นักเรียนลองผิดลองถูกได้ด้วยตนเอง เนื่องจากโรงเรียนมีเพียงสิ่งที่ถูกและผิดเท่านั้น ซึ่งทำให้เป็นการปิดกั้นเด็กในการเรียนรู้ เพราะเด็กกลัวทำผิด**	**คำตอบที่ถูกต้องคือ ข้อ ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง** สิ่งที่อ่านในเนื้อเรื่องเป็นเรื่องของครูที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อนและตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก **ดังนั้นเนื้อเรื่องจึงเกี่ยวข้องกับเด็กที่เคยทำผิดพลาด ซึ่งจุดมุ่งหมายของผู้พูดคือเพื่อแสดงความคิดเห็นว่าให้นักเรียนลองผิดลองถูกได้ด้วยตนเอง เนื่องจากโรงเรียนมีเพียงสิ่งที่ถูกและผิดเท่านั้น ซึ่งทำให้เป็นการปิดกั้นเด็กในการเรียนรู้ เพราะเด็กกลัวทำผิด**

Figure 53 Error-explanation prompting method in evaluate and reflect subscale (Item 4)

1.1.2.3 Mixed prompting method

Mixed prompting method was the mixture between the instructional prompting and the error-explanation prompting methods. Therefore, prompting was based on the instruction to guide a student to answer the question correctly and the reason why the selected choice was incorrect. If a student answered incorrectly, the prompt would be provided based on the number of attempts and choices a student selected. Examples of mixed prompting method were presented in Figures 54-56.

การออกแบบวิธีการชี้แนะ (promptings)				
3. การชี้แนะแบบผสมผสาน (Mixed prompting)				
ตอบผิดครั้งที่ 1 Prompt 1	**ก**	**ข**	**ง**	**จ**
	เธอเป็นเด็กขี้เกียจ **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร อยู่ในย่อหน้าใดของบทอ่าน (?)	เธอไม่มีความรับผิดชอบต่องาน **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร อยู่ในย่อหน้าใดของบทอ่าน (?)	เธอไม่ค่อยตั้งใจเรียนเวลาครูสอน **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร อยู่ในย่อหน้าใดของบทอ่าน (?)	เธอคิดว่าคงเรียนไม่ผ่านวิชานี้อยู่แล้ว จึงไม่สนใจทำงานส่ง **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร อยู่ในย่อหน้าใดของบทอ่าน (?)
ตอบผิดครั้งที่ 2 Prompt 2	**ก**	**ข**	**ง**	**จ**
	ผู้พูดไม่ได้บอกในบทอ่านว่าเด็กผู้หญิงคนนี้เป็นเด็กขี้เกียจ ลองใช้เทคนิคการอ่านเพื่อหารายละเอียดในการหาสาระที่ต้องการ ดังต่อไปนี้ 1. คำถามได้ถามถึง**เหตุผลที่ทำให้เด็กผู้หญิงคนนี้มีพฤติกรรมส่งงานช้า** 2. ให้นักเรียนอ่านย่อ**หน้าแรก**ของบทความ โดย**อ่านเร็ว ๆ กวาดสายตาคร่าว ๆ** เพื่อจับประเด็นสำคัญในลักษณะ**ข้อมูลตัวเลข** หรือ**คำสำคัญ**ที่เป็น**ข้อเท็จจริง**ในบทอ่าน	ผู้พูดไม่ได้บอกในบทอ่านว่าเด็กผู้หญิงคนนี้เป็นเด็กไม่มีความรับผิดชอบ ลองใช้เทคนิคการอ่านเพื่อหารายละเอียดในการหาสาระที่ต้องการ ดังต่อไปนี้ 1. คำถามได้ถามถึง**เหตุผลที่ทำให้เด็กผู้หญิงคนนี้มีพฤติกรรมส่งงานช้า** 2. ให้นักเรียนอ่านย่อ**หน้าแรก**ของบทความ โดย**อ่านเร็ว ๆ กวาดสายตาคร่าว ๆ** เพื่อจับประเด็นสำคัญในลักษณะ**ข้อมูลตัวเลข** หรือ**คำสำคัญ**ที่เป็น**ข้อเท็จจริง**ในบทอ่าน	ผู้พูดไม่ได้บอกในบทอ่านว่าเด็กผู้หญิงคนนี้เป็นเด็กไม่ตั้งใจเรียน ลองใช้เทคนิคการอ่านเพื่อหารายละเอียดในการหาสาระที่ต้องการ ดังต่อไปนี้ 1. คำถามได้ถามถึง**เหตุผลที่ทำให้เด็กผู้หญิงคนนี้มีพฤติกรรมส่งงานช้า** 2. ให้นักเรียนอ่านย่อ**หน้าแรก**ของบทความ โดย**อ่านเร็ว ๆ กวาดสายตาคร่าว ๆ** เพื่อจับประเด็นสำคัญในลักษณะ**ข้อมูลตัวเลข** หรือ**คำสำคัญ**ที่เป็น**ข้อเท็จจริง**ในบทอ่าน	ผู้พูดไม่ได้บอกในบทอ่านว่าเด็กผู้หญิงคนนี้ไม่สนใจทำงานส่ง เพราะเธอคิดว่าเธอคงเรียนไม่ผ่านวิชานี้อยู่แล้ว ลองใช้เทคนิคการอ่านเพื่อหารายละเอียดในการหาสาระที่ต้องการ ดังต่อไปนี้ 1. คำถามได้ถามถึง**เหตุผลที่ทำให้เด็กผู้หญิงคนนี้มีพฤติกรรมส่งงานช้า** 2. ให้นักเรียนอ่านย่อ**หน้าแรก**ของบทความ โดย**อ่านเร็ว ๆ กวาดสายตาคร่าว ๆ** เพื่อจับประเด็นสำคัญในลักษณะ**ข้อมูลตัวเลข** หรือ**คำสำคัญ**ที่เป็น**ข้อเท็จจริง**ในบทอ่าน
ตอบผิดครั้งที่ 3 Prompt 3	**ก**	**ข**	**ง**	**จ**
	ผู้พูดไม่ได้บอกว่าเด็กผู้หญิงคนนี้ส่งงานช้าเพราะเป็นเด็กขี้เกียจ เพราะผู้พูดกล่าวว่า แม้ครูส่วนใหญ่คิดว่าเด็กผู้หญิงคนนี้เป็นเด็กขี้เกียจ แต่มีครูท่านหนึ่งที่ไม่ได้คิดแบบนั้น จึงลองเข้าไปพูดคุยและ สอบถามเด็กดู	**ผู้พูดไม่ได้บอกว่าเด็กผู้หญิงคนนี้ส่งงานช้าเพราะเป็นเด็กไม่มีความรับผิดชอบ** เพราะผู้พูดกล่าวว่า แม้ครูส่วนใหญ่คิดว่าเด็กผู้หญิงคนนี้เป็นเด็กที่ไม่มีความรับผิดชอบ แต่มีครูท่านหนึ่งที่ไม่ได้คิดแบบนั้น จึงลองเข้าไปพูดคุยและสอบถามเด็กดู	**ผู้พูดไม่ได้บอกว่าเด็กผู้หญิงคนนี้ส่งงานช้าเพราะไม่ค่อยตั้งใจเรียน** เพราะผู้พูดกล่าวว่า แม้ครูส่วนใหญ่คิดว่าเธอไม่น่ารัก และดูเหมือนไม่ค่อยตั้งใจเรียน แต่มีครูท่านหนึ่งที่ไม่ได้คิดแบบนั้น จึงลองเข้าไปพูดคุยและสอบถามเด็กดู	**ผู้พูดไม่ได้บอกว่าเด็กผู้หญิงคนนี้ส่งงานช้าเพราะเธอคิดว่าคงเรียนไม่ผ่านวิชานี้อยู่แล้วจึงไม่สนใจทำงานส่ง** เพราะ ผู้พูดกล่าวว่า มีครูท่านหนึ่งที่เข้าไปสอบถามเด็กดูและพบว่าความจริงแล้วเธอไม่ได้คิดว่าจะไม่ผ่านวิชานี้

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Figure 54 Mixed prompting method in locate information subscale (Item 1)

การออกแบบวิธีการชี้แนะ (prompting)				
3. การชี้แนะแบบผสมผสาน (Mixed prompting)				
ตอบผิดครั้งที่ 1 Prompt 1	**ก**	**ค**	**ง**	**จ**
	ครูได้มอบของขวัญให้กับนักเรียน **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร และประโยคดังกล่าวอยู่ในตำแหน่งใดของบทอ่าน (?)	ครูให้โอกาสนักเรียนส่งงานช้ากว่ากำหนดได้ **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร และประโยคดังกล่าวอยู่ในตำแหน่งใดของบทอ่าน (?)	ครูทำให้นักเรียนไม่เครียดเวลาที่สอนในชั้นเรียน **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร และประโยคดังกล่าวอยู่ในตำแหน่งใดของบทอ่าน (?)	ครูเปิดโอกาสให้นักเรียนได้ทำในสิ่งที่ตนเองสนใจ **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร และประโยคดังกล่าวอยู่ในตำแหน่งใดของบทอ่าน (?)
ตอบผิดครั้งที่ 2 Prompt 2	**ก**	**ค**	**ง**	**จ**
	ผู้พูดไม่ได้พูดถึงครูที่มอบ ของขวัญให้แก่นักเรียน ในบทอ่าน ลองใช้เทคนิคการอ่านแปลความเพื่อหาความหมายของประโยค ดังต่อไปนี้ 1. ผู้พูดได้กล่าวว่า**สิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน** 2. ให้นักเรียนอ่านย่อ**หน้าแรกและย่อหน้าที่สอง**ของบทความ หาตำแหน่งของประโยคดังกล่าว 3. ให้นักเรียนแปลความหมายแฝงของคำว่า “ของขวัญ” และความหมายของประโยคดังกล่าวโดยพิจารณา**บริบทของข้อความ** โดยอ่าน**ประโยคสนับสนุนก่อนหน้า** และ**ประโยคสนับสนุนที่ตามหลัง**	ผู้พูดไม่ได้พูดถึงครูที่ให้โอกาสนักเรียนส่งงานช้ากว่ากำหนดได้ในบทอ่าน ลองใช้เทคนิคการอ่านแปลความเพื่อหาความหมายของประโยค ดังต่อไปนี้ 1. ผู้พูดได้กล่าวว่า**สิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน** 2. ให้นักเรียนอ่านย่อ**หน้าแรกและย่อหน้าที่สอง**ของบทความ หาตำแหน่งของประโยคดังกล่าว 3. ให้นักเรียนแปลความหมายแฝงของคำว่า “ของขวัญ” และความหมายของประโยคดังกล่าวโดยพิจารณา**บริบทของข้อความ** โดยอ่าน**ประโยคสนับสนุนก่อนหน้า** และ**ประโยคสนับสนุนที่ตามหลัง**	ผู้พูดไม่ได้พูดถึงครูที่ทำให้นักเรียนไม่เครียดเวลาที่สอนในชั้นเรียนในบทอ่าน ลองใช้เทคนิคการอ่านแปลความเพื่อหาความหมายของประโยค ดังต่อไปนี้ 1. ผู้พูดได้กล่าวว่า**สิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน** 2. ให้นักเรียนอ่านย่อ**หน้าแรกและย่อหน้าที่สอง**ของบทความ หาตำแหน่งของประโยคดังกล่าว 3. ให้นักเรียนแปลความหมายแฝงของคำว่า “ของขวัญ” และความหมายของประโยคดังกล่าวโดยพิจารณา**บริบทของข้อความ** โดยอ่าน**ประโยคสนับสนุนก่อนหน้า** และ**ประโยคสนับสนุนที่ตามหลัง**	แม้ผู้พูดจะสนับสนุนความอยากเรียนรู้ของนักเรียน แต่ผู้พูดไม่ได้พูดถึงครูที่เปิดโอกาสให้นักเรียนได้ทำสิ่งที่ตนเองสนใจในบทอ่าน ลองใช้เทคนิคการอ่านแปลความเพื่อหาความหมายของประโยค ดังต่อไปนี้ 1. ผู้พูดได้กล่าวว่า**สิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน** 2. ให้นักเรียนอ่านย่อ**หน้าแรกและย่อหน้าที่สอง**ของบทความ หาตำแหน่งของประโยคดังกล่าว 3. ให้นักเรียนแปลความหมายแฝงของคำว่า “ของขวัญ” และความหมายของประโยคดังกล่าวโดยพิจารณา**บริบทของข้อความ** โดยอ่าน**ประโยคสนับสนุนก่อนหน้า** และ**ประโยคสนับสนุนที่ตามหลัง**
ตอบผิดครั้งที่ 3 Prompt 3	**ก**	**ค**	**ง**	**จ**
	ของขวัญที่ครูให้กับนักเรียนไม่ได้หมายถึงครูมอบกล่องของขวัญให้กับนักเรียน เพราะคำว่า “ของขวัญ” เป็นความหมายแฝงของผู้พูด ไม่ใช่ความหมายโดยตรง ซึ่งผู้พูดเล่าว่าตนเองเคยเป็นนักเรียนที่ล้มเหลวมาก่อน และได้ของขวัญจากครู ทำให้มีวันนี้ได้ โดยผู้พูดไม่ได้กล่าวถึงครูมอบของขวัญให้กับนักเรียนเลย **ผู้พูดได้กล่าวว่าสิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน** ซึ่งประโยคนี้อยู่ใน***ย่อหน้าที่สอง***ของบทอ่าน ซึ่งการแปลความหมายของประโยคสามารถพิจารณาจากบริบทข้อความจาก**ประโยคก่อนหน้า**ได้ เนื่องจากผู้พูดกล่าวคำว่า **“สิ่งเหล่านี้”** แทนประโยคสนับสนุนที่ผู้พูดได้เล่าไว้ก่อนหน้าแล้ว ตามข้อความข้างล่างนี้ (บรรทัดที่ 9-11)	**ของขวัญที่ครูให้กับนักเรียนไม่ได้หมายถึงครูให้โอกาสนักเรียนส่งงานช้ากว่ากำหนด** เพราะคำว่า “ของขวัญ” เป็นความหมายแฝงของผู้พูด ไม่ใช่ความหมายโดยตรง ซึ่งผู้พูดเล่าว่าตนเองเคยเป็นนักเรียนที่ล้มเหลวมาก่อน และได้ของขวัญจากครู ทำให้มีวันนี้ได้ โดยผู้พูดไม่ได้กล่าวถึงครูให้นักเรียนส่งงานช้ากว่ากำหนดเลย **ผู้พูดได้กล่าวว่าสิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน** ซึ่งประโยคนี้อยู่ใน***ย่อหน้าที่สอง***ของบทอ่าน ซึ่งการแปลความหมายของประโยคสามารถพิจารณาจากบริบทข้อความจาก**ประโยคก่อนหน้า**ได้ เนื่องจากผู้พูดกล่าวคำว่า **“สิ่งเหล่านี้”** แทนประโยคสนับสนุนที่ผู้พูดได้เล่าไว้ก่อนหน้าแล้ว ตามข้อความข้างล่างนี้ (บรรทัดที่ 9-11)	**ของขวัญที่ครูให้กับนักเรียนไม่ได้หมายถึงครูทำให้นักเรียนไม่เครียดเวลาที่สอนในชั้นเรียน** เพราะคำว่า “ของขวัญ” เป็นความหมายแฝงของผู้พูด ไม่ใช่ความหมายโดยตรง ซึ่งผู้พูดเล่าว่าตนเองเคยเป็นนักเรียนที่ล้มเหลวมาก่อน และได้ของขวัญจากครู ทำให้มีวันนี้ได้ โดยผู้พูดไม่ได้กล่าวถึงครูที่ทำให้นักเรียนไม่เครียดเวลาที่สอนในชั้นเรียนเลย **ผู้พูดได้กล่าวว่าสิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน** ซึ่งประโยคนี้อยู่ใน***ย่อหน้าที่สอง***ของบทอ่าน ซึ่งการแปลความหมายของประโยคสามารถพิจารณาจากบริบทข้อความจาก**ประโยคก่อนหน้า**ได้ เนื่องจากผู้พูดกล่าวคำว่า **“สิ่งเหล่านี้”** แทนประโยคสนับสนุนที่ผู้พูดได้เล่าไว้ก่อนหน้าแล้ว ตามข้อความข้างล่างนี้ (บรรทัดที่ 9-11)	**ของขวัญที่ครูให้กับนักเรียนไม่ได้หมายถึงครูที่เปิดโอกาสให้นักเรียนได้ทำสิ่งที่ตนเองสนใจ** เพราะคำว่า “ของขวัญ” เป็นความหมายแฝงของผู้พูด ไม่ใช่ความหมายโดยตรง ซึ่งผู้พูดเล่าว่าตนเองเคยเป็นนักเรียนที่ล้มเหลวมาก่อน และได้ของขวัญจากครู ทำให้มีวันนี้ได้ ซึ่งแม้ผู้พูดจะสนับสนุนความอยากเรียนรู้ของนักเรียน แต่ผู้พูดไม่ได้กล่าวถึงครูที่เปิดโอกาสให้นักเรียนได้ทำสิ่งที่ตนเองสนใจ **ผู้พูดได้กล่าวว่าสิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน** ซึ่งประโยคนี้อยู่ใน***ย่อหน้าที่สอง***ของบทอ่าน ซึ่งการแปลความหมายของประโยคสามารถพิจารณาจากบริบทข้อความจาก**ประโยคก่อนหน้า**ได้ เนื่องจากผู้พูดกล่าวคำว่า **“สิ่งเหล่านี้”** แทนประโยคสนับสนุนที่ผู้พูดได้เล่าไว้ก่อนหน้าแล้ว ตามข้อความข้างล่างนี้ (บรรทัดที่ 9-11)

Figure 55 Mixed prompting method in understand subscale (Item 2)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

การออกแบบวิธีการชี้แนะ (promptings)				
3. การชี้แนะแบบผสมผสาน (Mixed prompting)				
ตอบผิดครั้งที่ 1 Prompt 1	ข	ค	ง	จ
	เพื่อเปรียบเทียบความแตกต่างระหว่างสิ่งที่นักเรียนทำถูกและสิ่งที่นักเรียนทำผิด **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร หาว่าข้อความที่ผู้พูดกล่าวเน้นอยู่ในย่อหน้าใดของบทอ่าน (?)	เพื่อเสนอแนะวิธีการเรียนรู้ที่ถูกต้องให้แก่นักเรียนนำไปปฏิบัติ **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร หาว่าข้อความที่ผู้พูดกล่าวเน้นอยู่ในย่อหน้าใดของบทอ่าน (?)	เพื่อสนับสนุนครูในการส่งเสริมนักเรียนให้เรียนรู้ได้ด้วยตนเอง **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร หาว่าข้อความที่ผู้พูดกล่าวเน้นอยู่ในย่อหน้าใดของบทอ่าน (?)	เพื่อเป็นกำลังใจให้แก่นักเรียนในการเรียนรู้จากความผิดพลาดของตนเอง **ยังไม่ใช่คำตอบที่ถูกต้อง** อ่านคำถามอีกครั้ง แล้วหาว่าคำถามถามถึงเรื่องอะไร หาว่าข้อความที่ผู้พูดกล่าวเน้นอยู่ในย่อหน้าใดของบทอ่าน (?)
ตอบผิดครั้งที่ 2 Prompt 2	ข	ค	ง	จ
	แม้ผู้พูดบอกว่ามีแค่สองเครื่องหมายคือถูกกับผิด แต่ผู้พูดไม่ได้กล่าวเปรียบเทียบสิ่งที่นักเรียนทำถูกหรือทำผิด ลองใช้เทคนิคการสะท้อนความคิดเห็นในเชิงคุณภาพและรูปแบบการเขียน เพื่อหาสาระที่ต้องการ ดังต่อไปนี้ 1. คำถามได้ถามเกี่ยวกับจุดมุ่งหมายในการพูด ซึ่งต้องอาศัยความรู้ทั่วไป มุมมองหรือทัศนคติของนักเรียนมาเชื่อมโยงกับสิ่งที่อ่านในเนื้อเรื่อง โดยข้อความที่ผู้พูดกล่าวเน้นอยู่ในย่อหน้าที่สี่ 2. เริ่มจากหามุมมองของผู้พูด โดยหาว่าใครเป็นผู้เล่าเรื่อง ผู้เล่าเรื่องที่อยู่ในเหตุการณ์หรือผู้เล่าเรื่องที่มิได้อยู่ในเหตุการณ์ ผู้เล่าเรื่องมีความสัมพันธ์กับเรื่องในลักษณะใด และผู้เล่าเรื่องกำลังเล่าให้ใครฟัง 3. จากนั้นให้นักเรียนดูข้อความที่ผู้พูดกล่าวเน้นสะท้อนจุดมุ่งหมายในการพูดอย่างไร เช่น เพื่อเล่าเรื่อง เพื่อชี้แจง/อธิบายวิธี เพื่อแสดงความคิดเห็น เพื่อโน้มน้าวใจ เพื่อโต้แย้ง เป็นต้น	ผู้พูดไม่ได้กล่าวถึงวิธีการเรียนรู้ที่ถูกต้องให้แก่นักเรียนนำไปปฏิบัติ ลองใช้เทคนิคการสะท้อนความคิดเห็นในเชิงคุณภาพและรูปแบบการเขียน เพื่อหาสาระที่ต้องการ ดังต่อไปนี้ 1. คำถามได้ถามเกี่ยวกับจุดมุ่งหมายในการพูด ซึ่งต้องอาศัยความรู้ทั่วไป มุมมองหรือทัศนคติของนักเรียนมาเชื่อมโยงกับสิ่งที่อ่านในเนื้อเรื่อง โดยข้อความที่ผู้พูดกล่าวเน้นอยู่ในย่อหน้าที่สี่ 2. เริ่มจากหามุมมองของผู้พูด โดยหาว่าใครเป็นผู้เล่าเรื่อง ผู้เล่าเรื่องที่อยู่ในเหตุการณ์หรือผู้เล่าเรื่องที่มิได้อยู่ในเหตุการณ์ ผู้เล่าเรื่องมีความสัมพันธ์กับเรื่องในลักษณะใด และผู้เล่าเรื่องกำลังเล่าให้ใครฟัง 3. จากนั้นให้นักเรียนดูข้อความที่ผู้พูดกล่าวเน้นสะท้อนจุดมุ่งหมายในการพูดอย่างไร เช่น เพื่อเล่าเรื่อง เพื่อชี้แจง/อธิบายวิธี เพื่อแสดงความคิดเห็น เพื่อโน้มน้าวใจ เพื่อโต้แย้ง เป็นต้น	ผู้ไม่ได้กล่าวถึงการส่งเสริมนักเรียนให้เรียนรู้ได้ด้วยตนเอง ลองใช้เทคนิคการสะท้อนความคิดเห็นในเชิงคุณภาพและรูปแบบการเขียน เพื่อหาสาระที่ต้องการ ดังต่อไปนี้ 1. คำถามได้ถามเกี่ยวกับจุดมุ่งหมายในการพูด ซึ่งต้องอาศัยความรู้ทั่วไป มุมมองหรือทัศนคติของนักเรียนมาเชื่อมโยงกับสิ่งที่อ่านในเนื้อเรื่อง โดยข้อความที่ผู้พูดกล่าวเน้นอยู่ในย่อหน้าที่สี่ 2. เริ่มจากหามุมมองของผู้พูด โดยหาว่าใครเป็นผู้เล่าเรื่อง ผู้เล่าเรื่องที่อยู่ในเหตุการณ์หรือผู้เล่าเรื่องที่มิได้อยู่ในเหตุการณ์ ผู้เล่าเรื่องมีความสัมพันธ์กับเรื่องในลักษณะใด และผู้เล่าเรื่องกำลังเล่าให้ใครฟัง 3. จากนั้นให้นักเรียนดูข้อความที่ผู้พูดกล่าวเน้นสะท้อนจุดมุ่งหมายในการพูดอย่างไร เช่น เพื่อเล่าเรื่อง เพื่อชี้แจง/อธิบายวิธี เพื่อแสดงความคิดเห็น เพื่อโน้มน้าวใจ เพื่อโต้แย้ง เป็นต้น	แม้ผู้พูดกล่าวว่าจะเป็นครูที่เป็นกำลังใจให้นักเรียน และให้โอกาสกับความผิดพลาดของเด็กๆในบทอ่าน แต่คำกล่าวนี้ไม่ใช่ข้อความที่ยกมา ในข้อคำถาม ลองใช้เทคนิคการสะท้อนความคิดเห็นในเชิงคุณภาพและรูปแบบการเขียน เพื่อหาสาระที่ต้องการ ดังต่อไปนี้ 1. คำถามได้ถามเกี่ยวกับจุดมุ่งหมายในการพูด ซึ่งต้องอาศัยความรู้ทั่วไป มุมมองหรือทัศนคติของนักเรียนมาเชื่อมโยงกับสิ่งที่อ่านในเนื้อเรื่อง โดยข้อความที่ผู้พูดกล่าวเน้นอยู่ในย่อหน้าที่สี่ 2. เริ่มจากหามุมมองของผู้พูด โดยหาว่าใครเป็นผู้เล่าเรื่อง ผู้เล่าเรื่องที่อยู่ในเหตุการณ์หรือผู้เล่าเรื่องที่มิได้อยู่ในเหตุการณ์ ผู้เล่าเรื่องมีความสัมพันธ์กับเรื่องในลักษณะใด และผู้เล่าเรื่องกำลังเล่าให้ใครฟัง 3. จากนั้นให้นักเรียนดูข้อความที่ผู้พูดกล่าวเน้นสะท้อนจุดมุ่งหมายในการพูดอย่างไร เช่น เพื่อเล่าเรื่อง เพื่อชี้แจง/อธิบายวิธี เพื่อแสดงความคิดเห็น เพื่อโน้มน้าวใจ เพื่อโต้แย้ง เป็นต้น
ตอบผิดครั้งที่ 3 Prompt 3	ข	ค	ง	จ
	ผู้พูดไม่ได้มีจุดมุ่งหมายเพื่อเปรียบสิ่งที่นักเรียนทำถูกหรือทำผิด เพราะผู้พูดกล่าวว่าโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้นทำให้นักเรียนกลัวการทำผิดพลาด ถึงแม้ผู้พูดบอกว่ามีแค่สองเครื่องหมายคือถูกกับผิด แต่ผู้พูดไม่ได้กล่าวเปรียบเทียบสิ่งที่นักเรียนทำถูกหรือทำผิด	ผู้พูดไม่ได้มีจุดมุ่งหมายเสนอแนะวิธีการเรียนรู้ที่ถูกต้องให้แก่นักเรียน เพราะผู้พูดกล่าวว่าโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้นทำให้นักเรียนกลัวการทำผิดพลาด ซึ่งผู้พูดไม่ได้กล่าวถึงวิธีการเรียนรู้ที่ถูกต้องให้แก่นักเรียนนำไปปฏิบัติ	ผู้พูดไม่ได้มีจุดมุ่งหมายเพื่อสนับสนุนครูในการเรียนรู้ด้วยตนเองของนักเรียน เพราะผู้พูดกล่าวว่าโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้นทำให้นักเรียนกลัวการทำผิดพลาด ซึ่งผู้พูดไม่ได้กล่าวถึงการส่งเสริมนักเรียนให้เรียนรู้ได้ด้วยตนเอง	ผู้พูดไม่ได้มีจุดมุ่งหมายเพื่อให้กำลังใจนักเรียน เพราะแม้ผู้พูดกล่าวว่าจะเป็นครูที่เป็นกำลังใจให้นักเรียน และให้โอกาสกับความผิดพลาดของเด็กๆ แต่คำกล่าวนี้ไม่ใช่ข้อความที่ยกมาในข้อคำถาม ซึ่งผู้พูดกล่าวว่าโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้นทำให้นักเรียนกลัวการทำผิดพลาด

การออกแบบวิธีการชี้แนะ (promptings)			
3. การชี้แนะแบบผสมผสาน (Mixed prompting)			
	โดย**มุมมองของผู้พูดคือเป็นบุคคลที่อยู่ในเหตุการณ์** ผู้พูดเป็นผู้เล่าเรื่องและกำลังเล่าเรื่องให้ครูท่านอื่น ๆ ฟัง เนื้อเรื่องเล่าถึงครูท่านหนึ่งที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อน และตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก ผู้พูดจึงกล่าวเน้นว่ามันเป็นเรื่องน่ากลัวที่เด็กในโรงเรียนจะกลัวผิดพลาด เพราะโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้น ซึ่งปิดกั้นความอยากเรียนรู้ของเด็ก ให้นักเรียนพิจารณาประโยคในคำถามอีกครั้ง และหาว่า**ผู้พูดมีจุดมุ่งหมายในการพูดอย่างไร**	โดย**มุมมองของผู้พูดคือเป็นบุคคลที่อยู่ในเหตุการณ์** ผู้พูดเป็นผู้เล่าเรื่องและกำลังเล่าเรื่องให้ครูท่านอื่น ๆ ฟัง เนื้อเรื่องเล่าถึงครูท่านหนึ่งที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อน และตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก ผู้พูดจึงกล่าวเน้นว่ามันเป็นเรื่องน่ากลัวที่เด็กในโรงเรียนจะกลัวผิดพลาด เพราะโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้น ซึ่งปิดกั้นความอยากเรียนรู้ของเด็ก ให้นักเรียนพิจารณาประโยคในคำถามอีกครั้ง และหาว่า**ผู้พูดมีจุดมุ่งหมายในการพูดอย่างไร**	โดย**มุมมองของผู้พูดคือเป็นบุคคลที่อยู่ในเหตุการณ์** ผู้พูดเป็นผู้เล่าเรื่องและกำลังเล่าเรื่องให้ครูท่านอื่น ๆ ฟัง เนื้อเรื่องเล่าถึงครูท่านหนึ่งที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อน และตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก ผู้พูดจึงกล่าวเน้นว่ามันเป็นเรื่องน่ากลัวที่เด็กในโรงเรียนจะกลัวผิดพลาด เพราะโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้น ซึ่งปิดกั้นความอยากเรียนรู้ของเด็ก	โดย**มุมมองของผู้พูดคือเป็นบุคคลที่อยู่ในเหตุการณ์** ผู้พูดเป็นผู้เล่าเรื่องและกำลังเล่าเรื่องให้ครูท่านอื่น ๆ ฟัง เนื้อเรื่องเล่าถึงครูท่านหนึ่งที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อน และตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก ผู้พูดจึงกล่าวเน้นว่ามันเป็นเรื่องน่ากลัวที่เด็กในโรงเรียนจะกลัวผิดพลาด เพราะโรงเรียนมีเพียงเครื่องหมายถูกกับผิดเท่านั้น ซึ่งปิดกั้นความอยากเรียนรู้ของเด็ก ให้นักเรียนพิจารณาประโยคในคำถามอีกครั้ง และหาว่า**ผู้พูดมีจุดมุ่งหมายในการพูดอย่างไร**
ตอบผิดครั้งที่ 4 Prompt 4	**ข**	**ค**	**ง**	**จ**
	คำตอบที่ถูกต้องคือ ข้อ ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง สิ่งที่อ่านในเนื้อเรื่องเป็นเรื่องของครูที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อนและตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก **ดังนั้นเนื้อเรื่องจึงเกี่ยวข้องกับเด็กที่เคยทำผิดพลาด ซึ่งจุดมุ่งหมายของผู้พูดคือเพื่อแสดงความคิดเห็นว่าให้นักเรียนลองผิดลองถูกได้ด้วยตนเอง เนื่องจากโรงเรียนมีเพียงสิ่งที่ถูกและผิดเท่านั้น ซึ่งทำให้เป็นการปิดกั้นเด็กในการเรียนรู้ เพราะเด็กกลัวทำผิด**	**คำตอบที่ถูกต้องคือ ข้อ ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง** สิ่งที่อ่านในเนื้อเรื่องเป็นเรื่องของครูที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อนและตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก **ดังนั้นเนื้อเรื่องจึงเกี่ยวข้องกับเด็กที่เคยทำผิดพลาด ซึ่งจุดมุ่งหมายของผู้พูดคือเพื่อแสดงความคิดเห็นว่าให้นักเรียนลองผิดลองถูกได้ด้วยตนเอง เนื่องจากโรงเรียนมีเพียงสิ่งที่ถูกและผิดเท่านั้น ซึ่งทำให้เป็นการปิดกั้นเด็กในการเรียนรู้ เพราะเด็กกลัวทำผิด**	**คำตอบที่ถูกต้องคือ ข้อ ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง** สิ่งที่อ่านในเนื้อเรื่องเป็นเรื่องของครูที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อนและตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก **ดังนั้นเนื้อเรื่องจึงเกี่ยวข้องกับเด็กที่เคยทำผิดพลาด ซึ่งจุดมุ่งหมายของผู้พูดคือเพื่อแสดงความคิดเห็นว่าให้นักเรียนลองผิดลองถูกได้ด้วยตนเอง เนื่องจากโรงเรียนมีเพียงสิ่งที่ถูกและผิดเท่านั้น ซึ่งทำให้เป็นการปิดกั้นเด็กในการเรียนรู้ เพราะเด็กกลัวทำผิด**	**คำตอบที่ถูกต้องคือ ข้อ ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง** สิ่งที่อ่านในเนื้อเรื่องเป็นเรื่องของครูที่เคยเป็นนักเรียนที่ล้มเหลวมาก่อนและตั้งใจว่าถ้าสอนเด็กคนไหนก็ตาม จะหยิบยื่นโอกาสให้เด็กที่ทำผิดพลาดและสร้างความเชื่อมั่นในตัวเด็ก **ดังนั้นเนื้อเรื่องจึงเกี่ยวข้องกับเด็กที่เคยทำผิดพลาด ซึ่งจุดมุ่งหมายของผู้พูดคือเพื่อแสดงความคิดเห็นว่าให้นักเรียนลองผิดลองถูกได้ด้วยตนเอง เนื่องจากโรงเรียนมีเพียงสิ่งที่ถูกและผิดเท่านั้น ซึ่งทำให้เป็นการปิดกั้นเด็กในการเรียนรู้ เพราะเด็กกลัวทำผิด**

Figure 56 Mixed prompting method in evaluate and reflect subscale (Item 4)

1.2 Results of the validation of the computerized dynamic assessment for reading literacy of Grade 9 students

The data were collected from the large pilot group of Grade 9 students. There were a total of 525 pilot samples administered for the validation of the instruments of the computerized dynamic assessment for reading literacy. The pilot study was carried out in January-March 2020. The results were shown in four main topics as follows: content validation, psychometric validation, and reliability.

1.2.1 Content validation

1.2.1.1 Content validity of the instruments of the computerized dynamic assessment for reading literacy

After the instruments of the computerized dynamic assessment for reading literacy were constructed, the researcher asked four groups of experts in the field of Thai language and in the field of educational measurement and evaluation to validate the reading literacy tests. The three sets of reading literacy instruments were separated into four sets of the judgmental reading literacy tests in accordance with the similarity of tasks (personal, public, occupational, and educational tasks). Each set of the judgmental reading literacy test comprised five experts, four of them were in the field of Thai language and reading literacy and one was in the field of educational measurement and evaluation. There were a total of 20 experts taking part in this judgmental process. A list of experts was presented in Appendix A. They were asked to check the content validity using IOC index (Item-Objective Congruency Index). The IOC index ranged from -1 to 1. The IOC index higher than 0.5 was acceptable (Rovinelli & Hambleton, 1977).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

With respect to the IOC index of reading literacy Form C-DA1, results suggested that all items in the first-tier reading literacy had an IOC index ranging from 0.6 – 1.0, except for items 9, 10, 15 that needed to be revised. Moreover, the same items (i.e., Items 9, 10, 15) in the second-tier reading literacy needed to be changed. Others in both first-tier and second-tier items had been suggested to revise in terms of language use. For reading literacy Form C-DA2, the result suggested that most items in the first-tier

reading literacy had IOC index ranged from 0.6 - 1.0, except for 2 items (i.e., Items 30 and 35) that needed to be changed. Because the second-tier items relied on the first-tier items as the rational-based items, these two second-tier items needed to be revised. Also, some items were suggested to make a minor revision in language use. For reading literacy Form C-DA3, most items had IOC ranged from 0.6 – 1.0, except for two items (items 50 and 51) needed to make a revision. The same items in the second-tier items also needed some revision. Hence, only three items for reading literacy Form C-DA1, two items for reading literacy Form C-DA2, and two items for reading literacy Form C-DA3 had been changed in accordance with experts' judgment. The result of an IOC index of the reading literacy instruments was shown in Appendix C.

1.2.1.2 Content validity of the prompting methods of the computerized dynamic assessment for reading literacy

After the reading literacy tests for the computerized dynamic assessment was constructed, the researcher also developed prompting methods for each item. There were four different prompting methods, including 1) the instructional prompt, 2) the error-explanation prompt, and 3) the mixture of instructional and error-explanation prompt, and 4) the verification prompt (control group). The groups of experts were the same groups taking part in the judgment of reading literacy tests. They were not only asked to validate the reading literacy tests but also to validate the prompting methods in each item, except the verification prompt. They were required to check whether the prompts were correct, clear, and understandable and also whether they clearly matched with the defined prompt guidelines. A list of experts was presented in Appendix A. They were asked to check the content validity of prompting methods using IOC index (Item-Objective Congruency Index). The IOC index ranged from -1 to 1. The IOC index higher than 0.5 was acceptable.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

With respect to the IOC index of prompting methods of reading literacy Form C-DA1, results suggested that all prompts in the reading literacy Form C-DA2 had IOC

index higher than 0.5, except for the error-explanation and mixed prompts of item 9 that needed to be revised. For prompts in the reading literacy Form C-DA2, the result suggested that most prompts in the reading literacy Form C-DA2 had IOC index higher than 0.5, except for three prompts that needed to be changed, which included the instructional prompt of item 31, the mixed prompt of item 33, and the error-explanation prompt of item 34. For prompts in the reading literacy Form C-DA3, all prompts had IOC higher than 0.5, indicating that all of them were acceptable. As a result, several prompts matched with the defined prompt guidelines. Few prompts had to be changed in accordance with experts' judgment to make them more comprehensible. The result of an IOC index of the prompting methods was shown in Appendix D.

1.2.2 Psychometric validation of the computerized dynamic assessment for reading literacy

Reading literacy instruments were administered to 525 ninth-grade students. They were randomly requested to complete two out of three test sets, which eventually included 359 samples for reading Form C-DA1, 347 samples for reading Form C-DA2, and 344 samples for reading Form C-DA3. Research instruments were three sets of reading literacy tests composed of two-tier multiple choice items. Reading literacy items were constructed to measure students' reading literacy ability to answer first-tier questions and their reasoning with second-tier questions. Practically, the scoring methods of two-tier items can be either pair scoring or individual scoring schemes. The pair-scoring method treats both first-tier and second-tier questions as a single item and assigns a score when both of them are correct, whereas the individual-scoring method treats both questions as the individual items and credits each item independently. In this present study, due to the concern of information loss, the researcher treated the reading literacy test items as a single item for the individual-scoring method.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2.2.1 The overall model fit

The model comparison of three reading literacy test sets were evaluated using model-data fit measures. The overall model fit of the three sets of reading literacy tests were shown below.

1) The model fit of reading literacy Form C-DA1

As Table 30 presented for reading literacy Form C-DA1, 2PL Bifactor models was the best fitting model by having smaller information criterion value for the first-tier items of reading literacy Form C-DA1. Moreover, the 2PL Bifactor model represented the smallest RMSEA, G^2 and Log-likelihood estimates. Because of the nested models, a series of χ^2- difference tests were considered to decide which one is the best-fitting model. Given to the reading literacy Form C-DA1, the results of the χ^2- difference tests showed that the bifactor model fitted the data significantly better than the testlet model ($\chi^2(16) = 44.02$, $p < .05$), which in turn fitted the data significantly better than the 2PL model ($\chi^2(20) = 67.19$, $p < .05$). As a result, it was found that the 2PL Bifactor model was the most fitted for the first-tier data.

For the second-tier data of reading literacy Form C-DA1, it was found that the 2PL Bifactor model had the smallest RMSEA values, Log-likelihood and G^2 values. For the results of the χ^2- difference tests, the bifactor model fitted the data significantly better than the testlet model ($\chi^2(16) = 30.42$, $p < .05$) and the 2PL unidimensional model ($\chi^2(20) = 40.59$, $p < .05$). Hence, the 2PL Bifactor model was preferred as shown in Table 30.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 30 The overall model fit of reading literacy Form C-DA1

Reading literacy Form C-DA1	Model	Log-likelihood	Akaike's Information Criterion (AIC)	Bayesian Information Criterion (BIC)	RMSEA	Deviance Statistic (G^2)	Likelihood ratio Chi-square Statistic		
							Pair	$\chi^2(df)$	*p*
First-tier	1PL	-4393.11	8828.22	8909.77	0.06	4650.41			
	2PL	-4316.93	8713.87	8869.20	0.04	4498.06	2PL vs Testlet	23.17*(4)	0.00
	M2PL	-4335.40	8762.81	8941.44	0.04	4535.00			
	Testlet	-4305.35	8698.70	8869.56	0.04	4474.89	Testlet vs Bifactor	44.02*(16)	0.00
	2PL Bifactor	-4283.33	8686.67	8919.67	0.03	4430.86	Bifactor vs 2PL	67.19*(20)	0.00
Second-tier	1PL	-4718.10	9478.20	9559.75	0.05	5297.08			
	2PL	-4667.52	9415.05	9570.38	0.04	5195.93	2PL vs Testlet	10.16*(4)	0.00
	M2PL	-4695.48	9482.96	9661.59	0.05	5251.83			
	Testlet	-4662.44	9412.88	9583.75	0.04	5185.76	Testlet vs Bifactor	30.42*(16)	0.01
	2PL Bifactor	-4647.23	9414.46	9647.46	0.04	5155.33	Bifactor vs 2PL	40.59*(20)	0.00

*p< .05

2) The model fit of reading literacy Form C-DA2

As seen in Table 31 for reading literacy Form C-DA2, the 2PL Bifactor model was the best fitting model by having smaller information criterion values for the first-tier items of reading literacy test Form C-DA2. The smallest RMSEA, Log-likelihood, and G^2 of first-tier items obtained by the 2PL Bifactor model, representing a good fit to the data. Thus, the 2PL Bifactor model was chosen as the most fitted for the first-tier data. When considering the results of the χ^2- difference tests of the parallel Form C-DA2, bifactor model significantly fitted the data compared to the testlet model ($\chi^2(16) = 33.48$, $p < .05$) and 2PL unidimensional model ($\chi^2(20) = 65.74$, $p < .05$).

Similarly, it was found that the 2PL Bifactor model was the best fitting model for the second-tier items by having a smaller model information criterion. Also, the smallest Log-likelihood, RMSEA, and G^2 values were also obtained by the 2PL Bifactor

model. For the results of the χ^2- difference tests, the bifactor model fitted the data significantly better than the testlet model ($\chi^2(16) = 35.19$, $p < .05$) and the 2PL unidimensional model ($\chi^2(20) = 46.92$, $p < .05$).

Table 31 The overall model fit of reading literacy Form C-DA2

Reading literacy Form C-DA2	Model	Log-likelihood	Akaike's Information Criterion (AIC)	Bayesian Information Criterion (BIC)	RMSEA	Deviance Statistic (G^2)	Likelihood ration Chi-square Statistic		
							Pair	$\chi^2(df)$	*p*
First-tier	1PL	-4262.34	8566.67	8646.83	0.09	4724.95			
	2PL	-4112.73	8305.47	8458.15	0.04	4425.75	2PL vs Testlet	32.26*(4)	0.00
	M2PL	-4161.56	8415.13	8590.72	0.09	4523.41			
	Testlet	-4096.60	8281.21	8449.16	0.05	4393.48	Testlet vs Bifactor	33.48*(16)	0.00
	2PL Bifactor	-4079.86	8279.72	8508.75	0.04	4360.00	Bifactor vs 2PL	65.74*(20)	0.00
Second-tier	1PL	-4578.71	9199.43	9280.27	0.05	5168.87			
	2PL	-4513.79	9107.59	9261.56	0.03	5039.03	2PL vs Testlet	11.79*(4)	0.01
	M2PL	-4530.64	9153.29	9330.36	0.06	5072.73			
	Testlet	-4507.89	9103.79	9273.16	0.03	5027.23	Testlet vs Bifactor	35.19*(16)	0.00
	2PL Bifactor	-4490.30	9100.60	9331.56	0.03	4992.04	Bifactor vs 2PL	46.92*(20)	0.00

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3) The model fit of reading literacy Form C-DA3

As Table 32 presented for reading literacy Form C-DA3, 2PL Bifactor models was the best fitting model by having smaller information criterion values for the first-tier items of reading literacy Form C-DA3. The smallest estimates of Log-likelihood, RMSEA, and G^2 obtained from the 2PL Bifactor model. Thus, 2PL Bifactor model was chosen as the most fitted for the first-tier data. Moreover, taking consideration of parallel Form C-DA3, bifactor model was statistically significant for the data than testlet model ($\chi^2(16) = 44.87$, $p < .05$) and 2PL unidimensional model ($\chi^2(20) = 47.69$, $p < .05$).

Similarly, it was found that the 2PL Bifactor models were the best fitting model for the second-tier items because of its smallest Log-likelihood, RMSEA,

and G^2 values. For the results of the χ^2- difference tests, the bifactor model fitted the data significantly better than the testlet model ($\chi^2(16) = 36.71$, $p < .05$) and the 2PL unidimensional model ($\chi^2(20) = 39.59$, $p < .05$).

Table 32 The overall model fit of reading literacy Form C-DA3

Reading literacy Form C-DA3	Model	Log-likelihood	Akaike's Information Criterion (AIC)	Bayesian Information Criterion (BIC)	RMSEA	Deviance Statistic (G^2)	Likelihood ratio Chi-square Statistic		
							Pair	$\chi^2(df)$	*p*
First-tier	1PL	-4223.06	8488.12	8568.77	0.07	4577.67			
	2PL	-4124.14	8328.29	8481.91	0.04	4379.84	2PL vs Testlet	2.81(4)	0.58
	M2PL	-4152.60	8397.20	8573.87	0.08	4436.76			
	Testlet	-4122.73	8333.47	8502.45	0.05	4377.02	Testlet vs. Bifactor	44.87* (16)	0.00
	2PL Bifactor	-4100.29	8320.59	8551.03	0.04	4332.15	Bifactor vs. 2PL	47.69* (20)	0.00
Second-tier	1PL	-4498.41	9038.83	9119.48	0.05	5060.80			
	2PL	-4428.14	8936.29	9089.92	0.03	4920.26	2PL vs Testlet	2.87(4)	0.57
	M2PL	-4497.65	9087.31	9263.98	0.07	5059.27			
	Testlet	-4426.71	8941.42	9110.41	0.03	4917.38	Testlet vs. Bifactor	36.71* (16)	0.00
	2PL Bifactor	-4408.35	8936.70	9167.14	0.03	4880.66	Bifactor vs. 2PL	39.59* (20)	0.00

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2.2.2 Item parameter estimates

The instruments were made up of 20 dichotomous items for students at the ninth-grade level in Thailand. Because of the combined four passages in reading literacy, the 2PL Bifactor model was selected as the most fitted model and the most appropriate item response model. Even though the reading tests were two-tier multiple choice tests, the score in each tier was counted as an individual-scoring method.

1) Item parameter estimates of reading literacy Form C-DA1

Estimated values of the discrimination values, item difficulty values, along with the intercept values of reading literacy Form C-DA1 were presented in Table 33. The values in the first column (a1) reflected the item slopes for general

factor, (a2) for specific factors. The discrimination of items was characterized by their slopes. The positive slopes showed that the probability of a correct response of a high-achieving student was higher than a low-achieving student, whereas the negative slopes depicted the opposite trend. MDISC represented the item discrimination, while MDIFF represented the item difficulty. The results found the larger magnitude of discrimination on the general dimension than on the group dimension, except four first-tier items (i.e., Items 1, 7, 13, and 18). However, they were not in the same testlets. Thus, none of them needed to be revised.

For the multidimensional discrimination estimates, a majority of the first-tier items had good MDISC. It was found that a majority of first-tier items were at moderate level, ranging from 0.55 – 1.65, and six items were highly MDISC items (i.e., Items 1, 2, 3, 7, 13, and 18). Only item 10 needed to be improved due to very low discriminating value. Regarding the multidimensional difficulty estimates, all first-tier items were produced in the acceptable range from -0.67 – 1.98, except for the item 10 which had a highly MDIFF value. For the second-tier items, more than half of them were good discrimination due to their ranked values at medium level and below. Only three items (i.e., Items 11, 14, and 16) needed to be improved in terms of relatively low discriminating values. Meanwhile, in terms of MDIFF, all items were estimated within the acceptable range, except for item 5 and item 16 which had highly difficult values.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 33 Item parameter estimates for reading literacy Form C-DA1

Passage	Item	Tier	Trait-a_1	Passage-a_2	$MDISC_i$	Discrimination level	$MDIFF_i$	Difficulty level
Passage 1	Item 1	Tier 1	4.08	4.56	6.12	Very high	-0.37	Medium
		Tier 2	0.32	0.28	0.42	Low	1.84	Difficult
	Item 2	Tier 1	1.88	1.01	2.14	Very high	-0.46	Medium
		Tier 2	-0.01	4.30	4.30	Very high	0.35	Medium
	Item 3	Tier 1	1.79	0.54	1.86	Very high	-0.67	Easy
		Tier 2	0.88	0.44	0.98	Moderate	-0.92	Easy
	Item 4	Tier 1	0.53	-0.14	0.55	Low	0.11	Medium
		Tier 2	1.97	-0.09	1.97	Very high	-0.59	Easy
	Item 5	Tier 1	0.36	-0.50	0.62	Low	1.98	Difficult
		Tier 2	0.21	-0.32	0.37	Low	2.34	Very difficult

Passage	Item	Tier	Trait-a_1	Passage-a_2	$MDISC_i$	Discrimination level	$MDIFF_i$	Difficulty level
Passage 2	Item 6	Tier 1	1.49	0.19	1.50	High	-0.45	Medium
		Tier 2	0.78	0.03	0.78	Moderate	0.74	Difficult
	Item 7	Tier 1	1.58	4.30	4.58	Very high	0.09	Medium
		Tier 2	0.67	0.27	0.72	Moderate	0.26	Medium
	Item 8	Tier 1	0.58	-0.12	0.59	Low	0.57	Difficult
		Tier 2	1.29	5.03	5.19	Very high	0.29	Medium
	Item 9	Tier 1	1.14	0.39	1.20	Moderate	-0.03	Medium
		Tier 2	0.70	0.58	0.91	Moderate	0.00	Medium
	Item 10	Tier 1	0.25	0.17	*0.30*	Very low	2.27	Very difficult
		Tier 2	0.82	0.12	0.83	Moderate	0.13	Medium
Passage 3	Item 11	Tier 1	1.34	-0.30	1.38	High	-0.14	Medium
		Tier 2	0.23	0.13	*0.26*	Very low	0.58	Difficult
	Item 12	Tier 1	1.15	0.23	1.17	Moderate	1.04	Difficult
		Tier 2	0.60	0.69	0.92	Moderate	0.03	Medium
	Item 13	Tier 1	1.25	3.37	3.60	Very high	0.56	Difficult
		Tier 2	0.07	0.34	0.35	Low	1.04	Difficult
	Item 14	Tier 1	0.73	0.32	0.80	Moderate	0.79	Difficult
		Tier 2	0.24	-0.15	*0.28*	Very low	0.58	Difficult
	Item 15	Tier 1	1.65	-0.01	1.65	High	-0.04	Medium
		Tier 2	-0.24	-1.03	1.06	Moderate	1.16	Difficult
Passage 4	Item 16	Tier 1	0.71	0.42	0.83	Moderate	1.06	Difficult
		Tier 2	0.05	-0.19	*0.19*	Very low	3.25	Very difficult
	Item 17	Tier 1	1.02	0.14	1.03	Moderate	0.97	Difficult
		Tier 2	0.87	0.29	0.92	Moderate	0.13	Medium
	Item 18	Tier 1	4.19	10.09	10.93	Very high	0.39	Medium
		Tier 2	0.74	-0.07	0.74	Moderate	0.63	Difficult
	Item 19	Tier 1	0.64	-0.50	0.81	Moderate	1.08	Difficult
		Tier 2	0.78	4.40	4.47	Very high	0.18	Medium
	Item 20	Tier 1	0.58	0.32	0.66	Moderate	1.16	Difficult
		Tier 2	0.43	0.16	0.45	Low	0.67	Difficult

Note. $MDISC_i$ = multidimensional discrimination estimate; $MDIFF_i$ = multidimensional difficulty

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2) Item parameter estimates of reading literacy Form C-DA2

As Table 34 shown for reading literacy Form C-DA2, a 2PL Bifactor model accounting for one general reading literacy factor and four passage factors was applied by using MHRM estimation method. The results found the larger magnitude of discrimination on the general dimension than on the group dimension, except five first-tier items (i.e., Items 23, 37, 38, 39, and 40). It was found that four

out of five items (i.e., Items 37, 38, 39, and 40) in passage eight had higher magnitude of discrimination on the group dimension than the general dimension. Thus, revision was made for some low discrimination items in passage eight (i.e., Items 39 and 40).

For the multidimensional discrimination estimates, all of the first-tier items had good discrimination. It was found that most items were at moderate level, and 6 items were found to have highly MDISC values (i.e., Items 21, 23, 28, 29, 32 and 38). Regarding the multidimensional difficulty estimates, all first-tier items were estimated in the acceptable range, except for item 40 which was considered a highly difficult item. For the second-tier items, all of them had good discrimination estimates. Meanwhile, in terms of the MDIFF, all items were estimated within the acceptable range. Most of them were medium MDIFF. Thus, none of the second-tier items needed to be improved in terms of item discrimination and item difficulty estimates.

Table 34 Item parameter estimates for reading literacy Form C-DA2

Passage	Item	Tier	Trait-a_1	Passage-a_2	$MDISC_i$	Discrimination level	$MDIFF_i$	Difficulty level
Passage 5	Item 21	Tier 1	1.83	0.55	1.91	Very high	-0.47	Medium
		Tier 2	0.54	0.34	0.64	Low	0.99	Difficult
	Item 22	Tier 1	1.08	0.17	1.10	Moderate	0.06	Medium
		Tier 2	0.81	0.30	0.87	Moderate	-0.58	Easy
	Item 23	Tier 1	1.74	3.66	4.05	Very high	0.00	Medium
		Tier 2	-0.19	-3.78	3.79	Very high	0.35	Medium
	Item 24	Tier 1	1.46	0.23	1.48	High	-0.00	Medium
		Tier 2	1.16	0.26	1.19	Moderate	0.49	Medium
	Item 25	Tier 1	0.33	-0.38	0.51	Low	1.24	Difficult
		Tier 2	0.24	-0.34	0.42	Low	0.39	Medium
Passage 6	Item 26	Tier 1	1.49	0.75	1.67	High	-0.12	Medium
		Tier 2	0.92	0.32	0.98	Moderate	-0.43	Medium
	Item 27	Tier 1	0.90	0.13	0.91	Moderate	-0.47	Medium
		Tier 2	1.68	1.86	2.50	Very high	-0.42	Medium
	Item 28	Tier 1	1.94	1.43	2.42	Very high	-0.09	Medium
		Tier 2	0.54	0.53	0.75	Moderate	0.17	Medium
	Item 29	Tier 1	2.48	1.39	2.84	Very high	-0.51	Easy
		Tier 2	1.11	0.05	1.11	Moderate	-0.74	Easy
	Item 30	Tier 1	-0.28	-0.39	0.48	Low	1.63	Difficult
		Tier 2	0.35	0.44	0.56	Low	0.57	Difficult

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Passage	Item	Tier	Trait-a_1	Passage-a_2	$MDISC_i$	Discrimination level	$MDIFF_i$	Difficulty level
Passage 7	Item 31	Tier 1	0.78	0.28	0.83	Moderate	-0.07	Medium
		Tier 2	0.58	0.18	0.60	Low	0.36	Medium
	Item 32	Tier 1	7.22	5.71	9.21	Very high	-0.22	Medium
		Tier 2	1.25	-0.22	1.27	Moderate	-0.46	Medium
	Item 33	Tier 1	0.60	-0.27	0.66	Moderate	1.16	Difficult
		Tier 2	0.57	0.60	0.83	Moderate	0.36	Medium
	Item 34	Tier 1	0.90	0.39	0.98	Moderate	0.19	Medium
		Tier 2	-2.49	6.51	6.97	Very high	0.52	Difficult
	Item 35	Tier 1	1.30	0.02	1.30	Moderate	0.01	Medium
		Tier 2	0.57	0.16	0.59	Low	0.80	Difficult
Passage 8	Item 36	Tier 1	1.49	0.76	1.68	High	0.14	Medium
		Tier 2	0.56	0.03	0.56	Low	-0.50	Medium
	Item 37	Tier 1	0.31	0.72	0.79	Moderate	0.09	Medium
		Tier 2	2.23	4.26	4.81	Very high	0.31	Medium
	Item 38	Tier 1	0.72	1.56	1.71	Very high	0.74	Difficult
		Tier 2	0.55	-0.24	0.60	Low	0.65	Difficult
	Item 39	Tier 1	0.12	0.45	0.47	Low	1.63	Difficult
		Tier 2	0.51	0.35	0.62	Low	0.41	Medium
	Item 40	Tier 1	-0.09	0.39	0.40	Low	*2.07*	Very difficult
		Tier 2	0.40	0.18	0.44	Low	0.51	Difficult

3) Item parameter estimates of reading literacy Form C-DA3

As shown in Table 35 for reading literacy Form C-DA3, the results found the larger magnitude of discrimination on the general dimension than on the group dimension, except eight first-tier items (i.e., Items 42, 43, 44, 46, 51, 58, 59, and 60). There were three items in passages nine (i.e., Items 42, 43, and 44) and three items in passage twelve (i.e., Items 58, 59, and 60) received higher magnitude of discrimination on the group dimension than the general dimension. Thus, items with low discrimination estimates in passage nine and passage twelve were revised (i.e., Items 43, 44, 59, and 60).

For the multidimensional discrimination estimates, more than half of the first-tier items had good discrimination. It was found that most items ranged from 0.38 to 1.43, except seven highly discriminating items (i.e., items 42, 46, 51, 52, 54, 56 and 58). Only two items (57 and 59) needed to be revised because of their very low

discrimination values. Regarding the multidimensional difficulty estimates, most of the first-tier items were at the moderate item difficulty and all of them were within the acceptable range, except for four items that had highly MDIFF values (Items 44, 57, 59, and 60). It is interesting that some items (i.e., Items 57 and 59) were the most difficult items and did not discriminate between low and high ability students. For the second-tier items, all of the second-tier items were good discrimination and good difficulty estimate. None needed to be improved.

Table 35 Item parameter estimates for reading literacy Form C-DA3

Passage	Item	Tier	a_1	a_2	$MDISC_i$	Discrimination level	$MDIFF_i$	Difficulty level
Passage 9	Item 41	Tier 1	1.38	0.22	1.40	High	-0.05	Medium
		Tier 2	1.89	2.52	3.15	Very high	-0.48	Medium
	Item 42	Tier 1	1.20	4.01	4.18	Very high	-0.12	Medium
		Tier 2	0.71	-0.07	0.72	Moderate	-0.22	Medium
	Item 43	Tier 1	0.27	0.49	0.57	Low	0.59	Difficult
		Tier 2	0.47	0.32	0.57	Low	0.40	Medium
	Item 44	Tier 1	-0.09	0.37	0.38	Low	3.21	Very difficult
		Tier 2	2.57	-0.19	2.58	Very high	-0.63	Easy
	Item 45	Tier 1	0.75	-0.13	0.77	Moderate	1.44	Difficult
		Tier 2	-0.35	0.04	0.35	Low	1.84	Difficult
Passage 10	Item 46	Tier 1	1.28	2.58	2.89	Very high	0.29	Medium
		Tier 2	0.54	0.39	0.66	Moderate	0.05	Medium
	Item 47	Tier 1	1.17	0.25	1.20	Moderate	-0.31	Medium
		Tier 2	1.90	6.54	6.80	Very high	-0.08	Medium
	Item 48	Tier 1	0.56	-0.12	0.57	Low	0.25	Medium
		Tier 2	-0.12	0.47	0.49	Low	1.54	Difficult
	Item 49	Tier 1	1.20	0.02	1.20	Moderate	-0.09	Medium
		Tier 2	0.61	-0.31	0.68	Moderate	-0.25	Medium
	Item 50	Tier 1	0.54	0.39	0.67	Moderate	1.74	Difficult
		Tier 2	0.64	0.18	0.67	Moderate	0.72	Difficult
Passage 11	Item 51	Tier 1	1.11	1.84	2.15	Very high	0.00	Medium
		Tier 2	0.60	-0.06	0.60	Low	-0.28	Medium
	Item 52	Tier 1	1.84	0.68	1.96	Very high	-0.37	Medium
		Tier 2	0.91	0.35	0.98	Moderate	-0.40	Medium
	Item 53	Tier 1	1.24	-0.71	1.43	High	0.01	Medium
		Tier 2	0.52	0.19	0.55	Low	0.02	Medium
	Item 54	Tier 1	1.94	0.68	2.05	Very high	-0.14	Medium
		Tier 2	2.92	3.08	4.24	Very high	-0.44	Medium
	Item 55	Tier 1	1.00	0.25	1.03	Moderate	0.71	Difficult

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Passage	Item	Tier	a_1	a_2	$MDISC_i$	Discrimination level	$MDIFF_i$	Difficulty level
		Tier 2	1.09	-0.34	1.15	Moderate	0.14	Medium
Passage 12	Item 56	Tier 1	1.77	-0.01	1.77	Very high	-0.37	Medium
		Tier 2	0.70	0.46	0.83	Moderate	0.14	Medium
	Item 57	Tier 1	0.29	0.06	0.30	Very low	4.34	Very difficult
		Tier 2	0.86	0.19	0.88	Moderate	0.15	Medium
	Item 58	Tier 1	0.27	2.25	2.26	Very high	0.87	Difficult
		Tier 2	0.50	-0.02	0.50	Low	0.83	Difficult
	Item 59	Tier 1	0.03	0.26	0.26	Very low	3.63	Very difficult
		Tier 2	0.36	0.21	0.42	Low	1.96	Difficult
	Item 60	Tier 1	-0.11	0.39	0.41	Low	3.49	Very difficult
		Tier 2	0.66	3.07	3.14	Very high	0.44	Medium

1.2.3 Confirmatory factor analysis of the computerized dynamic assessment for reading literacy

1.2.3.1. Confirmatory factor analysis of reading literacy Form C-DA1

The analyses were conducted with CFA to determine the model that best described the structure of reading literacy instrument (C-DA1). According to the analyses of item parameter estimates from bifactor model, one item had been removed because of low discriminating value (Item 10). The results of the CFA are reported in Table 36. The second-order confirmatory factor analysis was examined whether items fitted the reading literacy subscales and all factors fitted the general reading literacy. The results showed that the second-order factors of the reading literacy Form C-DA1 exhibited good model fit (CFI = .983, TLI = .980, RMSEA = .023, and SRMR =.049). The result also presented that chi-square (χ^2)= 170.698, df = 143, and p = .057, which was not statistically significant. The chi-square χ^2/df = 1.193, which indicated a good model fit, as it was below two.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 36 Confirmatory factor analysis of reading literacy Form C-DA1 (n = 359)

Model	χ^2	df	p	RMSEA	CFI	TLI	SRMR
C-DA1	170.698	143	.057	.023	.983	.980	.049

From the findings, the CFA showed that factor loadings of locate information subscale ranged from 0.245 - 0.573 and were statistically significant. Item 6 was found to have the greatest importance, followed by Item 11. For understand subscale, the factor loadings ranged from 0.275 – 0.535 and were statistically significant. The factor loading of Item 3 was the greatest. For evaluate and reflect subscale, factor loadings ranged from 0.123 – 0.649 and were statistically significant, except Item 5. Item 15 was found to have the highest importance as shown in Figure 57.

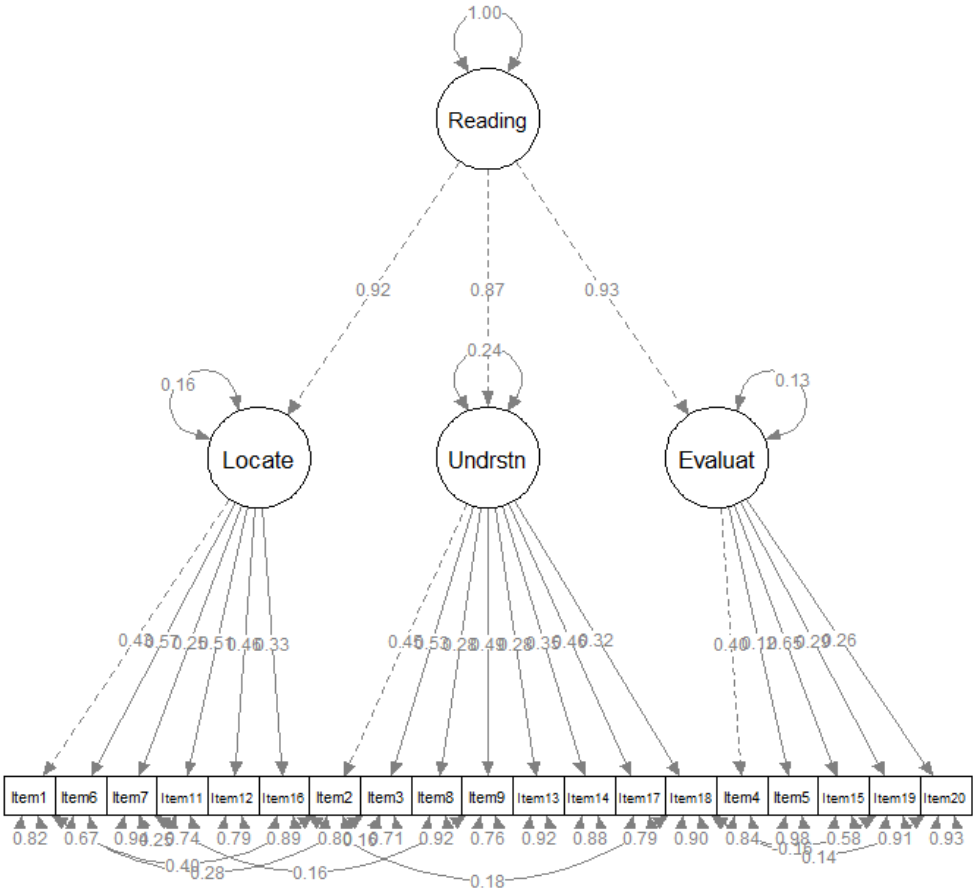
For the second-order factor analysis, the standardized values ranged between 0.871 – 0.934. The highest importance was evaluate and reflect, followed by locate information and understand subscales. Table 37 shows the values of estimate, standard errors, and p-values of the second-order factor model.

Table 37 The CFA estimates, standard errors, and p-values for the second-order factor analysis model of reading literacy Form C-DA1 (n = 359)

Estimate	Item	Beta	S.E.	*p*	R^2
Locate Information	Item1	0.425	-	-	0.181
	Item6	0.573	0.165	0.000	0.328
	Item7	0.245	0.163	0.000	0.060
	Item11	0.512	0.164	0.000	0.262
	Item12	0.459	0.159	0.000	0.211
	Item16	0.327	0.167	0.000	0.107
Understand	Item2	0.448	-	-	0.201
	Item3	0.535	0.138	0.000	0.286
	Item8	0.281	0.148	0.000	0.079
	Item9	0.489	0.168	0.000	0.239
	Item13	0.275	0.146	0.000	0.076
	Item14	0.347	0.150	0.000	0.120
	Item17	0.462	0.160	0.000	0.214
	Item18	0.315	0.151	0.000	0.099
Evaluate and reflect	Item4	0.403	-	-	0.162
	Item5	0.123	0.144	0.069	0.015
	Item15	0.649	0.198	0.000	0.421
	Item19	0.294	0.166	0.000	0.086
	Item20	0.257	0.159	0.000	0.066

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Reading	Locate information	0.916	-	-	0.839
	Understand	0.871	-	-	0.758
	Evaluate and reflect	0.934	-	-	0.873



CHULALONGKORN UNIVERSITY

Figure 57 Second-order factor analysis model of reading literacy Form C-DA1

1.2.3.2. Confirmatory factor analysis of reading literacy Form C-DA2

The analyses were conducted with CFA to determine the model that best described the structure of reading literacy instrument (C-DA2). According to the analyses of item parameter estimates from bifactor model, two item had been removed because of high difficulty and low discriminating values. The results of the CFA are reported in Table 38. The second-order confirmatory factor analysis was examined whether items fitted the reading literacy subscales and all factors fitted the general reading literacy. The results showed that the second-order factors of the

reading literacy Form C-DA2 exhibited good model fit for C-DA2 (CFI = .987, TLI = .985, RMSEA = .024, and SRMR =.051). The result also presented that chi-square (χ^2)= 155.798, df = 130, and p = .061, which was not statistically significant. The chi-square χ^2/df = 1.198, which indicated a good model fit, as it was below 2.

Table 38 Confirmatory factor analysis of reading literacy Form C-DA2 (n = 336)

Model	χ^2	df	p	RMSEA	CFI	TLI	SRMR
C-DA2	155.798	130	0.061	.024	.987	.985	.051

From the findings shown in Figure 58, the CFA showed that factor loadings of locate information subscale ranged from 0.366 – 0.643 and were statistically significant. Item 32 was found to have the greatest importance, followed by Item 26. For understand subscale, the factor loadings ranged from 0.158 – 0.595 and were statistically significant. The factor loading of Item 29 was the greatest. For evaluate and reflect subscale, factor loadings ranged from -0.171 – 0.506 and were statistically significant. Item 24 was found to have the highest importanc and Item 30 had negative value that needed revision.

The second-order analysis showed that the factor loadings were between 0.987- 1.000. Table 39 shows the values of estimate, standard errors, and p-values of the second-order factor model.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 39 The CFA estimates, standard errors, and p-values for the second-order factor analysis model of reading literacy Form C-DA2 (n = 336)

Estimate	Item	Beta	S.E.	p	R^2
Locate Information	Item21	0.518	-	-	0.268
	Item26	0.565	0.110	0.000	0.320
	Item27	0.394	0.117	0.000	0.155
	Item31	0.366	0.118	0.000	0.134
	Item32	0.643	0.108	0.000	0.413
	Item36	0.501	0.108	0.000	0.251

Estimate	Item	Beta	S.E.	*p*	R^2
Understand	Item22	0.500	-	-	0.250
	Item23	0.312	0.114	0.000	0.097
	Item28	0.563	0.113	0.000	0.317
	Item29	0.595	0.105	0.000	0.354
	Item33	0.231	0.112	0.000	0.053
	Item34	0.398	0.119	0.000	0.158
	Item37	0.158	0.122	0.000	0.025
	Item38	0.232	0.111	0.000	0.054
Evaluate and reflect	Item24	0.506	-	-	0.256
	Item25	0.144	0.118	0.022	0.021
	Item30	-0.171	0.112	0.005	0.029
	Item35	0.492	0.114	0.000	0.242
Reading	Locate information	1.000	-	-	1.000
	Understand	1.000	-	-	1.000
	Evaluate and reflect	0.987	-	-	0.974

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Figure 58 Second-order factor analysis model of reading literacy Form C-DA2

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2.3.3. Confirmatory factor analysis of reading literacy Form C-DA3

The analyses were conducted with CFA to determine the model that best described the structure of reading literacy instrument (C-DA3). According to the analyses of item parameter estimates from bifactor model, two item had been removed because of high difficulty and low discriminating values. The results of the CFA are reported in Table 40. The second-order confirmatory factor analysis was examined whether items fitted the reading literacy subscales and all factors fitted the general reading literacy. The results showed that the second-order factors of the reading literacy instruments exhibited good model fit for C-DA3 (CFI = .982, TLI =

.977, RMSEA = .027, and SRMR =.049). The result also presented that chi-square (χ^2)= 102.955, df = 82, and p = .059, which was not statistically significant. The chi-square χ^2/df = 1.255, which indicated a good model fit, as it was below 2.

Table 40 Confirmatory factor analysis of reading literacy Form C-DA3 (n = 344)

Model	χ^2	df	*p*	RMSEA	CFI	TLI	SRMR
C-DA3	102.955	82	.059	.027	.982	.977	.049

From the findings shown in Figure 59, the CFA showed that factor loadings of locate information subscale ranged from 0.335 – 0.590 and were statistically significant. Item 52 was found to have the greatest importance, followed by Item 56. For understand subscale, the factor loadings ranged from 0.159 – 0.581 and were statistically significant. The factor loading of Item 54 was the greatest. For evaluate and reflect subscale, factor loadings ranged from 0.267 – 0.493 and were statistically significant. Item 55 was found to have the highest importance.

The second-order analysis showed that the factor loadings were between 0.909-1.000. Table 41 shows the values of estimate, standard errors, and p-values of the second-order factor model.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Table 41 The CFA estimates, standard errors, and p-values for the second-order factor analysis model of reading literacy Form C-DA3 (n = 344)

Estimate	Item	Beta	S.E.	*P*	R^2
Locate	Item41	0.346	-	-	0.120
Information	Item46	0.335	0.204	0.000	0.112
	Item47	0.467	0.213	0.000	0.218
	Item51	0.420	0.198	0.000	0.177
	Item52	0.590	0.219	0.000	0.348
	Item56	0.582	0.221	0.000	0.338
Understand	Item42	0.316	-	-	0.100
	Item43	0.159	0.200	0.013	0.025
	Item48	0.194	0.207	0.003	0.038

Estimate	Item	Beta	S.E.	*P*	R^2
	Item49	0.488	0.229	0.000	0.238
	Item53	0.461	0.229	0.000	0.212
	Item54	0.581	0.256	0.000	0.337
Evaluate and	Item45	0.362	-	-	0.131
reflect	Item50	0.267	0.204	0.000	0.071
	Item55	0.493	0.230	0.000	0.243
Reading	Locate information	0.909	-	-	0.826
	Understand	1.000	-	-	1.000
	Evaluate and reflect	0.979	-	-	0.958

Figure 59 Second-order factor analysis model of reading literacy Form C-DA3

1.2.4 Reliability of the computerized dynamic assessment for reading literacy

Table 42 provided the reliability estimates using Cronbach's alpha coefficient and omega coefficient. For reading literacy test set 1, the Omega coefficient was 0.83 and the Cronbach's alpha coefficient was 0.81. Regarding reading literacy test set 2, the Omega coefficient was 0.86 and the Cronbach's alpha coefficient was 0.84. With regard to reading literacy test set 3, the Omega coefficient was 0.86 and the Cronbach's alpha coefficient was 0.83. This indicated that the reading literacy test sets had highly acceptable reliability values.

Table 42 Reliability of the instruments for computerized dynamic assessment for reading literacy

Composite	Reading literacy test set 1 (n = 181)		Reading literacy test set 2 (n = 178)		Reading literacy test set 3 (n =166)	
	Cronbach's alpha **(α)**	Omega coefficient**(ω)**	Cronbach's alpha **(α)**	Omega coefficient**(ω)**	Cronbach's alpha **(α)**	Omega coefficient**(ω)**
Full scale	0.81	0.83	0.84	0.86	0.83	0.86

1.2.5 Results of the parallelism of the computerized dynamic assessment for reading literacy of Grade 9 students

Assembling parallel forms were investigated under the unidimensional 2PL model. Parallel test forms were created to be equivalent in terms of statistical and content properties. The item information function (IIF) and test information function (TIF) were used as a statistical target to make sure that the test forms are equal in terms of precision. This study focused on computing the root mean square deviation (RMSD) statistic with respect to item closeness as well as test closeness between the assembled tests. The test parallelism was investigated as follows:

1.2.5.1 Parallel reading literacy forms of Test set 1 (Form C-DA1 and Form C-DA2)

1) Content-related parallelism of the reading literacy forms of Test set 1

After the reading literacy tests were constructed, the researcher asked four groups of experts to review the parallelism of the reading literacy tests. Each group of experts comprised four experts in the field of Thai language as well as another one in the field of educational measurement and evaluation. The list of experts was presented in Appendix A. They were asked to depict their agreement on each item pair of the equivalent tests in terms of the measured constructs and language use using IOC index (Item-Objective Congruency Index). The IOC index ranged from -1 to 1. The IOC index higher than 0.5 was acceptable, representing that the items were parallel forms and were appropriate to use in the main study. The results suggested that all item pairs had IOC index higher than 0.5, except for the two pairs (items 3 and 15) that needed to be revised. The result of an IOC index of the parallelism of reading literacy forms between Form C-DA1 and Form C-DA2 was shown in Appendix E.

2) Statistical-related parallelism of the reading literacy forms of Test set 1 (Form C-DA1 and Form C-DA2)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

According to the conformity of Test set 1, both test forms were applied to information functions and characteristic curves along with seven ability scales (i.e. -3.0, -2.0, -1.0, 0, 1.0, 2.0, 3.0). It can be seen that $RMSD_{IIF}$ estimates ranged from 0.01 to 0.14 and $RMSD_{ICC}$ ranged from 0.01-0.11, showing satisfactory closeness between the two items, with respect to the statistical index lower than 0.5. Moreover, there was no difference between Form C-DA1 and Form C-DA2 in terms of TIF and TCC. The TIF and TCC values were 0.40 and 0.17, respectively. Hence, the two forms showed satisfactory parallelism as presented in Table 43 and Figure 60.

Table 43 Parallel forms of reading literacy of Test Set 1(Form C-DA1 and Form C-DA2)

Item-by-item parallelism					Test parallelism			Evaluate
Pair	$RMSD_{IIIF}$	$RMSD_{ICC}$	Pair	$RMSD_{IIIF}$	$RMSD_{ICC}$	$RMSD_{TIF}$	$RMSD_{TCC}$	
1	0.03	0.02	11	0.05	0.03	0.40	0.17	Parallel
2	0.08	0.04	12	0.11	0.08			
3	0.11	0.07	13	0.00	0.02			
4	0.08	0.07	14	0.04	0.05			
5	0.00	0.02	15	0.02	0.02			
6	0.08	0.04	16	0.02	0.01			
7	0.04	0.03	17	0.00	0.07			
8	0.12	0.11	18	0.01	0.01			
9	0.14	0.06	19	0.01	0.06			
10	0.01	0.07	20	0.02	0.07			

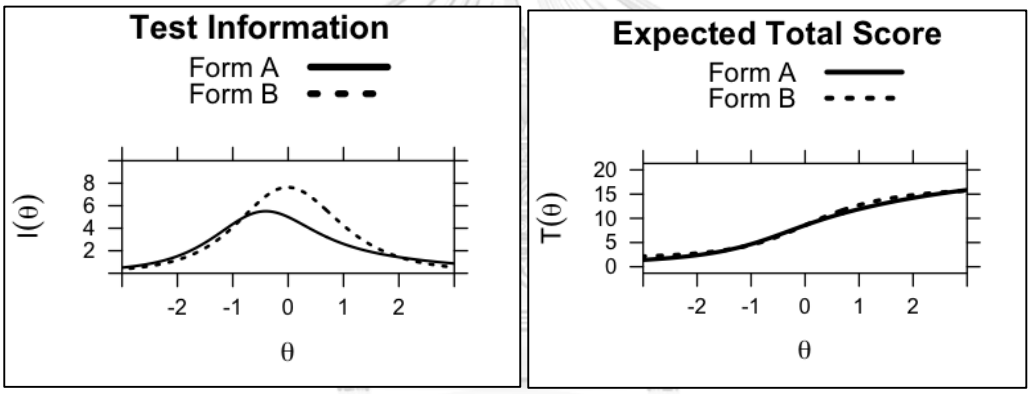


Figure 60 Test information and test characteristic curve of Test set 1

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2.5.2 Parallel reading literacy forms of Test set 2 (Form C-DA1 and Form C-DA3)

1) Content-related parallelism of the reading literacy forms of Test set 2 (Form C-DA1 and Form C-DA3)

After the reading literacy tests were constructed, the researcher asked four groups of experts to review the parallelism of the reading literacy tests. Each group of experts comprised four experts in the field of Thai language as well as another one in the field of educational measurement and evaluation. The list of experts was presented in Appendix A. They were asked to depict their agreement on

each item pair of the equivalent tests in terms of the measured constructs and language use using the content validity using IOC index (Item-Objective Congruency Index). The IOC index ranged from -1 to 1. The IOC index higher than 0.5 was acceptable, representing that the items were parallel forms and were appropriate to use in the main study.

The results suggested that all item pairs had IOC index higher than 0.5, except for the three pairs (items 11, 12 and 15) that needed to be revised. Thus, with respect to content-related parallelism, the researcher revised only the pairs of item 11, item 12, and item 15 to be more equivalent. The result of an IOC index of the parallelism of reading literacy forms between Form C-DA1 and Form C-DA3 is shown in Appendix E.

2) Statistical-related parallelism of the reading literacy forms of Test set 2 (Form C-DA1 and Form C-DA3)

Parallelism of Test set 2 were applied to information functions and characteristic curves along with seven ability scales (i.e. -3.0, -2.0, -1.0, 0, 1.0, 2.0, 3.0). In Table 44 and Figure 61, in accordance with item parallelism, it can be seen that all items showed satisfactory closeness between the two assembled forms because and $RMSD_{IIF}$ estimates ranged from 0.01 to 0.20 and $RMSD_{ICC}$ ranged from 0.02 to 0.13. Moreover, it can be seen that the differences between the test forms were very small ($RMSD_{TIF}$ = 0.17, $RMSD_{TCC}$ = 0.34). Thus, the amount of test information and test characteristic curve of Form C-DA1 and Form C-DA3 were relatively close.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 44 Parallel forms of reading literacy of Test set 2 (Form C-DA1 and Form C-DA3)

Item-by-item parallelism						Test parallelism		Evaluate
Pair	$RMSD_{IIF}$	$RMSD_{ICC}$	Pair	$RMSD_{IIF}$	$RMSD_{ICC}$	$RMSD_{TIF}$	$RMSD_{TCC}$	
1	0.04	0.03	11	0.03	0.02	0.17	0.34	Parallel
2	0.11	0.12	12	0.11	0.08			
3	0.12	0.12	13	0.04	0.05			
4	0.01	0.13	14	0.20	0.07			
5	0.04	0.09	15	0.08	0.05			
6	0.03	0.04	16	0.11	0.10			
7	0.13	0.08	17	0.11	0.10			
8	0.02	0.02	18	0.02	0.06			
9	0.06	0.03	19	0.02	0.05			
10	0.03	0.07	20	0.02	0.09			

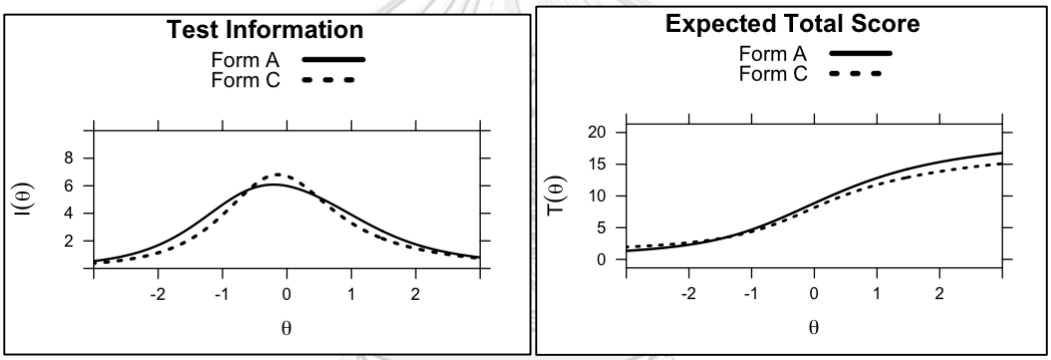


Figure 61 Test information and test characteristic curve of Test set 2

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2.5.3 Parallel reading literacy of Test set 3 (Form C-DA2 and Form C-DA3)

1) Content-related parallelism of reading literacy of Test set 3 (Form C-DA2 and Form C-DA3)

After reading literacy tests were constructed, the researcher asked four groups of experts to review the parallelism of the reading literacy tests. Each group of experts comprised four experts in the field of Thai language as well as another one in the field of educational measurement and evaluation. The list of experts was presented in Appendix A. They were asked to depict their agreement on each item pair of the equivalent tests in terms of the measured constructs and

language use using the content validity using IOC index (Item-Objective Congruency Index). The IOC index ranged from -1 to 1. The IOC index higher than 0.5 was acceptable, representing that the items were parallel forms and were appropriate to use in the main study.

The results suggested that all item pairs had IOC index higher than 0.5, except for the three pairs (items 3, 11 and 12) that needed to be revised. Thus, with respect to content-related parallelism, the researcher revised only the pairs of item 3, item 11 and item 12 to be more equivalent. The result of an IOC index of the parallelism of reading literacy forms between Form C-DA2 and Form C-DA3 is shown in Appendix E.

2) Statistical-related parallelism of the reading literacy of Test set 3 (Form C-DA2 and Form C-DA3)

As shown in Table 45 and Figure 62, parallelism of reading literacy of Test set 3 was applied to information function and characteristic curve. It can be seen that $RMSD_{IIF}$ estimates ranged from 0.00 to 0.15 and $RMSD_{ICC}$ from 0.01-0.14, showing satisfactory closeness between the two forms, with respect to the statistical index lower than 0.5. In contrast, it can be seen that differences are found between the test information functions and test characteristic curves of the assembled forms. The $RMSD_{TIF}$ and $RMSD_{TCC}$ were 0.61 and 0.71, respectively. Thus, Test set 3 did not represent the satisfactory closeness between the two forms. Thus, the researchers deleted four items that had unusually high values of item difficulty estimates (i.e., Items 44, 50, 58, and 59). As a result, the parallelism of Form C-DA2 and Form C-DA3 were met. The $RMSD_{TIF}$ and $RMSD_{TCC}$ were 0.50 and 0.45, respectively.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 45 Parallel forms of reading literacy of Test set 3

Item-by-item parallelism						Test parallelism (2^{nd} analysis)		Evaluate
Pair	$RMSD_{IIF}$	$RMSD_{ICC}$	Pair	$RMSD_{IIF}$	$RMSD_{ICC}$	$RMSD_{TIF}$	$RMSD_{TCC}$	
1	0.10	0.03	11	0.03	0.02	0.50	0.45	Parallel
2	0.03	0.03	12	0.10	0.01			
3	0.01	0.02	13	0.01	0.02			
4	0.08	0.14	14	0.12	0.05			
5	0.02	0.03	15	0.08	0.04			
6	0.06	0.07	16	0.03	0.02			
7	0.01	0.02	17	0.07	0.13			
8	0.12	0.09	18	0.05	0.10			
9	0.15	0.06	19	0.01	0.05			
10	0.00	0.07	20	0.00	0.06			

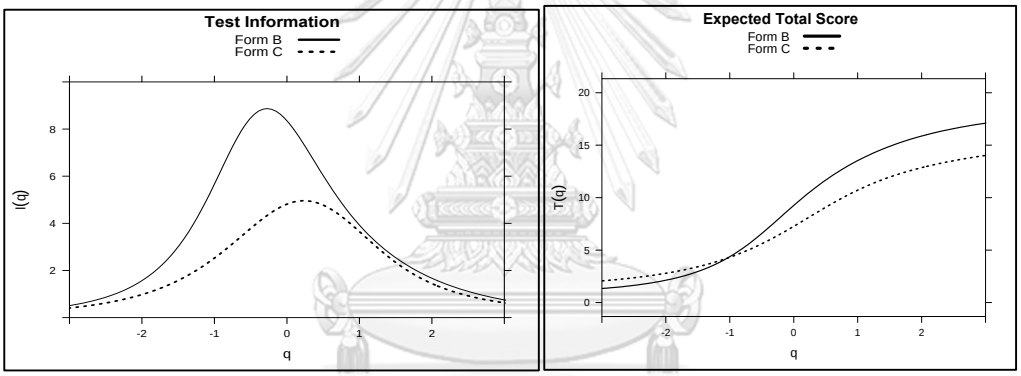


Figure 62 Test information between reading literacy of Test set 3

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Part 2: Results of the Effects of Different Promptings of the Computerized Dynamic Assessment on Reading Literacy Ability, Growth of Learning, and Learning Potential of Grade 9 Students

2.1 Results of the effects of different promptings of the computerized dynamic assessment on reading literacy ability of Grade 9 students

2.1.1 ANCOVA results of different prompting conditions on the overall reading literacy ability

The descriptive statistics of students' pretest and posttest scores on reading literacy are summarized in Table 46 and Figure 63. All prompting conditions received relatively similar scores on the pretest. Group B received the highest score on the pretest (*M*=8.06, *SD*=3.51). Moreover, all prompting conditions had a higher posttest score than the pretest score. The mean posttest score was greater in Group A (*M*=14.99, *SD*=3.33), followed by Group C (*M*=14.82, *SD*=3.09). On the other hand, Group D received the lowest posttest score (*M*=11.16, *SD*=3.80).

Table 46 Descriptive statistics of students' pretest and posttest scores on reading literacy

Group	Pretest score		Posttest score	
	M	*SD*	*M*	*SD*
A	7.93	3.31	14.99	3.33
B	8.06	3.51	14.51	3.57
C	7.47	3.13	14.82	3.09
D	7.86	2.98	11.16	3.80

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

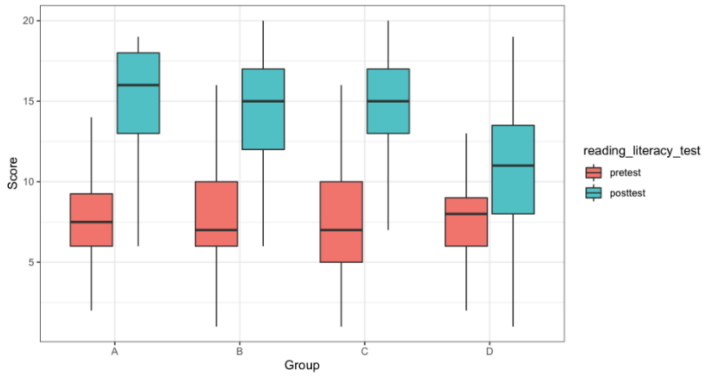


Figure 63 Students' reading literacy pretest and posttest scores

A one-way ANCOVA was performed to investigate the effect of the computerized dynamic assessment (C-DA) on students' reading literacy posttest score. The reading literacy score was measured pre- and post-training of computerized dynamic assessment (C-DA). The preliminary analyses investigating the homogeneity of slope, linearity, and homogeneity of variances assumptions were satisfactory. The results revealed that when the pretest score was considered as the covariate, there was a statistically significant difference in posttest score between the groups ($F(3, 536)=35.63$, $p < .05$). The summary of ANCOVA results is shown in Table 47 and Figure 64. The results of follow-up pairwise comparison revealed that Group D scored significantly lower when compared with students in any other groups as presented in Table 48.

Table 47 Analysis of Covariance for student's reading literacy posttest score with pretest score as covariate

Source	Univariate ANCOVA				
	SS	df	*MS*	*F*	*p*
Intercept	13494.61	1	13494.61	1142.19	.000
Pretest (Covariate)	59.44	1	59.44	5.03*	.025
Group	1262.70	3	420.90	35.63*	.000
Residuals	6332.63	536	11.81		

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

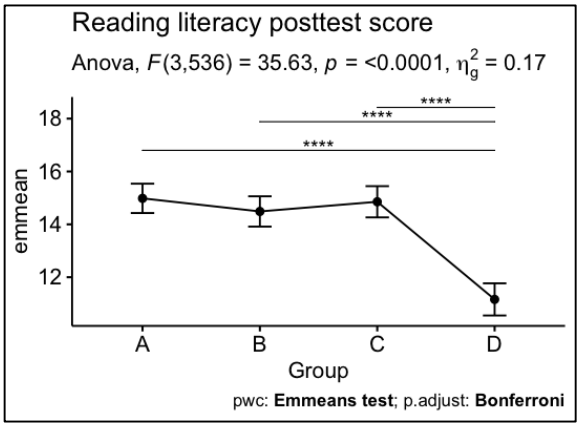


Figure 64 Line plot of the adjusted posttest score

Table 48 Results of pairwise comparisons

Pair comparison		Mean Difference (I - J)	Std. Error	*p*
Group (I)	Group (J)			
A	B	.495	.406	1.000
A	C	.129	.413	1.000
A	D	3.823*	.419	.000
B	C	-.366	.419	1.000
B	D	3.328*	.426	.000
C	D	3.694*	.432	.000

*$p < .05$

2.1.2 ANCOVA results of different prompting conditions on the reading literacy subscales

The descriptive statistics of students' pretest and posttest scores on reading literacy subscale are summarized in Table 49 and Figure 65. For pretest score, locate information subscale was received as the highest score for all prompting conditions, followed by understand and evaluate and reflect subscales. For posttest score, however, understand subscale was obtained the highest subscale score for all prompting conditions. The mean posttest score of locate information subscale was greater in Group A and Group B, respectively (*M*=5.30, *SD*=1.12; *M*=5.15, *SD*=1.11). Group A also received the highest score for the understand subscale (*M*=6.39, *SD*=1.52), followed by Group C (*M*=6.26, *SD*=1.54). Moreover, Group B had the highest evaluate and reflect subscale (*M*=3.45, *SD*=1.31). On the other hand, Group D received the lowest scores for all three subscales.

Table 49 Descriptive statistics of students' pretest and posttest scores on reading literacy subscales

Group	Dependent variable	Pretest score		Posttest score	
		M	*SD*	*M*	*SD*
Locate information	Group A	3.14	1.44	5.30	1.12
	Group B	3.29	1.50	5.15	1.11
	Group C	2.90	1.34	5.11	1.08
	Group D	2.98	1.31	4.37	1.42
Understand	Group A	2.94	1.54	6.39	1.52
	Group B	2.81	1.59	5.90	1.89
	Group C	2.64	1.53	6.26	1.54
	Group D	2.88	1.37	4.54	1.92
Evaluate and reflect	Group A	1.86	1.32	3.30	1.43
	Group B	1.95	1.30	3.45	1.31
	Group C	1.93	1.26	3.44	1.34
	Group D	2.01	1.24	2.25	1.38

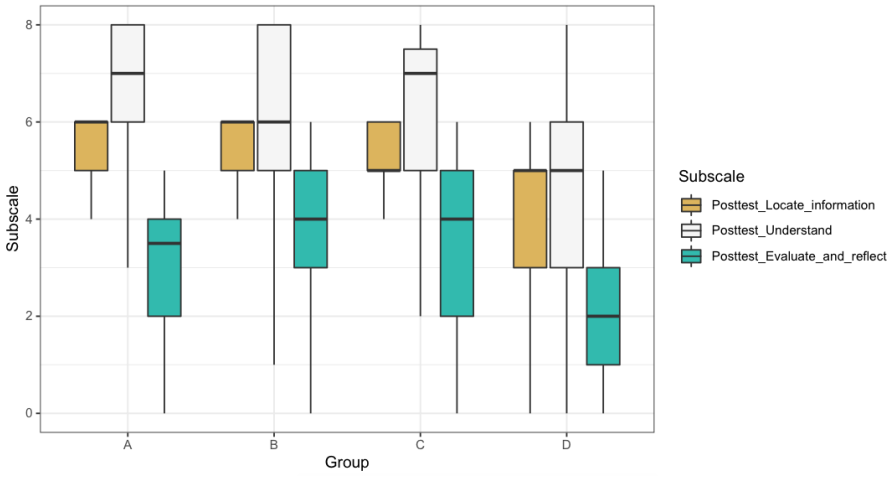


Figure 65 Students' posttest subscale scores

2.1.2.1 ANCOVA results of different prompting conditions on locate information subscale

A one-way ANCOVA was performed to investigate the effect of the computerized dynamic assessment (C-DA) on students' locate information posttest subscale score. As presented in Table 50, the results revealed that there was a statistically significant difference in locate information subscale score among the groups ($F(3, 536)=15.63$, $p < .05$). The results of follow-up pairwise comparison revealed that Group D significantly received lower locate information subscale when compared with students in any other groups as shown in Table 51 and Figure 66.

Table 50 Analysis of Covariance for student's locate information subscale score with pretest score as covariate

Source	Univariate ANCOVA				
	SS	df	*MS*	*F*	*p*
Intercept	2228.49	1	2228.49	1577.43*	.000*
Pretest (Covariate)	.69	1	.69	.49	.484
Group	66.27	3	22.08	15.63*	.000*
Residuals	757.22	536	1.41		

*$p < .05$

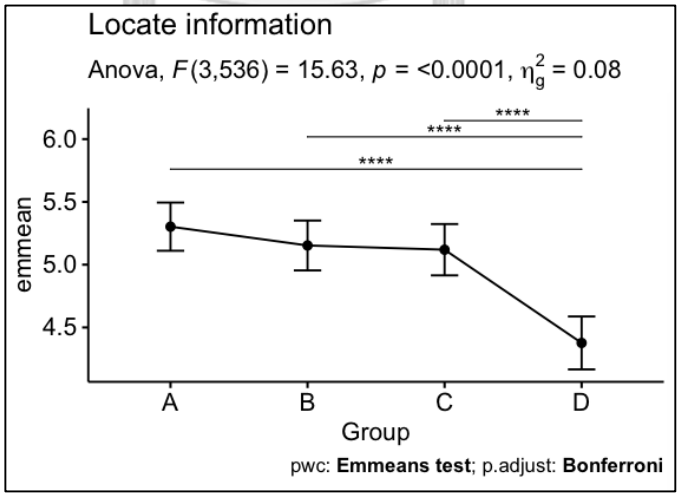


Figure 66 Line plot of the adjusted posttest score for locate information subscale

Table 51 Results of pairwise comparisons of locate information subscale

Pair comparison		Mean Difference	Std. Error	*P*
Group 1 (I)	Group 2 (J)	(I - J)		
A	B	.150	.141	1.000
A	C	.184	.143	1.000
A	D	.926*	.145	.000
B	C	.034	.145	1.000
B	D	.776*	.148	.000
C	D	.742*	.149	.000

*$p < .05$

2.1.2.2 ANCOVA results of different prompting conditions on understand subscale

A one-way ANCOVA was performed to investigate the effect of the computerized dynamic assessment (C-DA) on students' understand subscale score. As depicted in Table 52, the results revealed that significant differences were found in understand subscale score ($F(3, 536)=31.23$, $p < .05$). The results of follow-up pairwise comparison revealed that Group D was significantly different from any other groups as shown in Table 53 and Figure 67.

Table 52 Analysis of Covariance for student's understand subscale score with pretest score as covariate

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Source	Univariate ANCOVA				
	SS	df	*MS*	*F*	*p*
Intercept	3851.92	1	3851.92	1298.85*	.000*
Pretest (Covariate)	2.04	1	2.04	.68	.407
Group	277.85	3	92.61	31.23*	.000*
Residuals	1589.57	536	2.96		

*$p < .05$

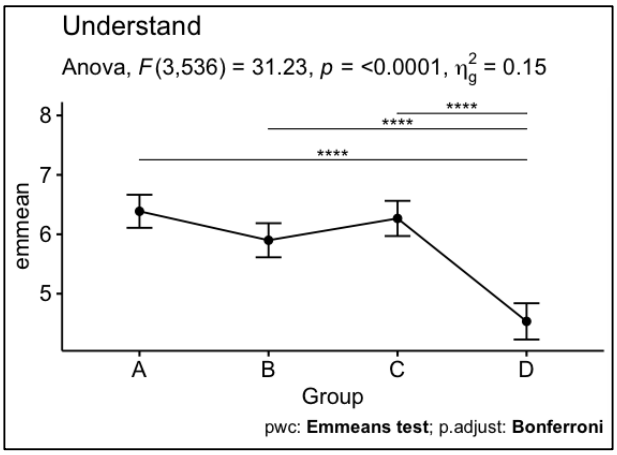


Figure 67 Line plot of the adjusted posttest score for understand subscale

Table 53 Results of pairwise comparisons of understand subscale

Pair comparison		Mean Difference	Std. Error	*P*
Group 1 (I)	Group 2 (J)	(I - J)		
A	B	.487	.203	.102
A	C	.120	.207	1.000
A	D	1.853*	.210	.000
B	C	-.367	.210	.484
B	D	1.365*	.213	.000
C	D	1.733*	.217	.000

*p< .05

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.1.2.3 ANCOVA results of different prompting conditions on evaluate and reflect subscale

A one-way ANCOVA was performed to investigate the effect of the computerized dynamic assessment (C-DA) on students' evaluate and reflect subscale score. As presented in Table 54, the results revealed that there was a statistically significant difference in evaluate and reflect subscale score (F(3, 536)=23.22, p < .05). The results of follow-up pairwise comparison revealed that Group D received lower evaluate and reflect subscale significantly when compared with students inany other groups as presented in Table 55 and Figure 68.

Table 54 Analysis of Covariance for student's evaluate and reflect subscale score with pretest score as covariate

Source	Univariate ANCOVA				
	SS	df	*MS*	*F*	*p*
Intercept	1350.81	1	1350.81	732.14*	.000
Pretest (Covariate)	13.94	1	13.94	7.55*	.006
Group	128.53	3	42.84	23.22*	.000
Residuals	988.93	536	1.84		

*p< .05

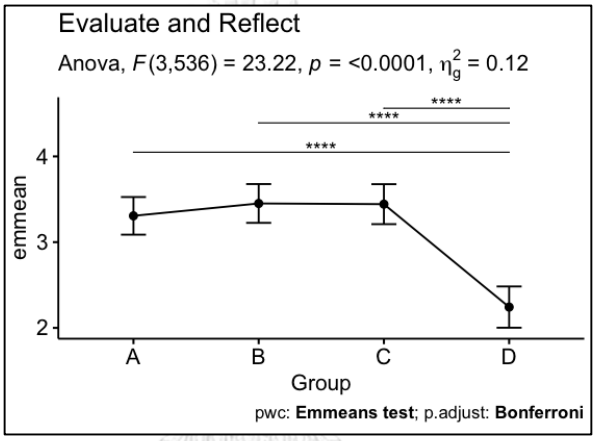


Figure 68 Line plot of the adjusted posttest score for evaluate and reflect subscale

Table 55 Results of pairwise comparisons of evaluate and reflect subscale

Pair comparison		Mean Difference	Std. Error	*p*
Group 1 (I)	Group 2 (J)	(I - J)		
A	B	-.144	.160	1.000
A	C	-.136	.163	1.000
A	D	1.064*	.166	.000
B	C	.008	.165	1.000
B	D	1.209*	.168	.000
C	D	1.200*	.171	.000

*p < .05

2.1.3 Results of different prompting conditions of the computerized dynamic assessment (C-DA) and students' demographic data on reading literacy posttest score

2.1.3.1 Two-way ANCOVA results of different prompting conditions and gender on students' reading literacy posttest score

Table 56 presents the mean scores and standard deviations of prompting groups and gender on reading literacy pretest and posttest scores. All groups obtained higher posttest scores than the pretest score. The highest score on pretest obtained by female students in all prompting groups. Those female students in Group A scored higher than any other groups on the pretest (*M*=8.49, *SD*=3.12). With respect to the posttest score, male students received higher posttest score than female students, except Group D. The highest posttest score of male students was Group C (*M*=15.11, *SD*=3.42), followed by Group A (*M*=15.08, *SD*=3.52). For the group of female students, Group A had the highest score (*M*=14.93, *SD*=3.20) as shown in Figure 69.

Table 56 Descriptive statistics of group and gender on reading literacy pretest and posttest scores

Group	Gender	Pretest score		Posttest score	
		M	SD	*M*	SD
Group A	Female	8.49	3.12	14.93	3.20
	Male	7.16	3.45	15.08	3.52
Group B	Female	8.19	3.76	14.45	3.67
	Male	7.91	3.23	14.58	3.48
Group C	Female	7.57	3.09	14.61	2.84
	Male	7.35	3.22	15.11	3.42
Group D	Female	8.06	2.47	11.44	3.77
	Male	7.59	3.60	10.76	3.84

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

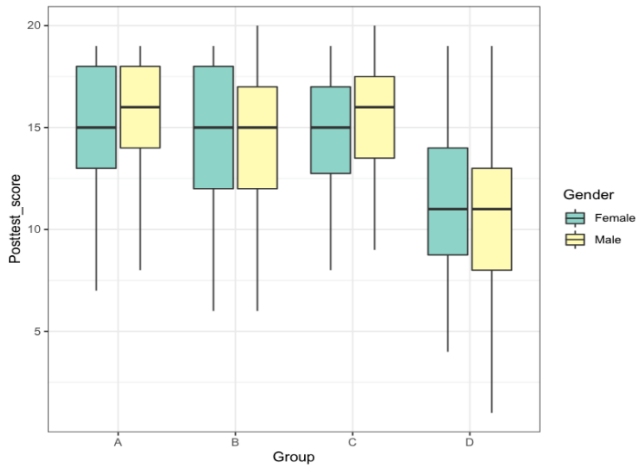


Figure 69 Students' posttest scores with respect to different prompting groups and gender

The two-way ANCOVA was conducted that examined the effect of prompting conditions and gender on students' reading literacy posttest score. The preliminary analyses investigating the homogeneity of slope, linearity, homogeneity of variances assumptions were satisfactory. From the results, an interaction between group and gender could not be demonstrated, $F(3, 532)=.65$, $p = .582$. In addition, there were no significant differences in the impact of gender on posttest score, $F(1, 532)=.08$, $p = .770$. However, there was significant difference in the prompting groups on reading literacy posttest score, $F(3, 532)=36.03$, $p < .05$, as shown in Table 57 and Figure 70. The pairwise comparison showed that Group D (control group) significantly received lower scores when compared with other prompting conditions as summarized in Table 58.

Table 57 A Two-way ANCOVA of prompting groups and gender on student's reading literacy posttest score with pretest score as covariate

Source	Univariate ANCOVA				
	SS	df	*MS*	*F*	*p*
Pretest	61	1	61.00	5.13*	.024
Group	1281.70	3	427.23	36.03*	.000
Gender	1.01	1	1.01	.08	.770
Group*Gender	23.20	3	7.73	.65	.582
Residuals	6308	532	11.85		

*$p < .05$

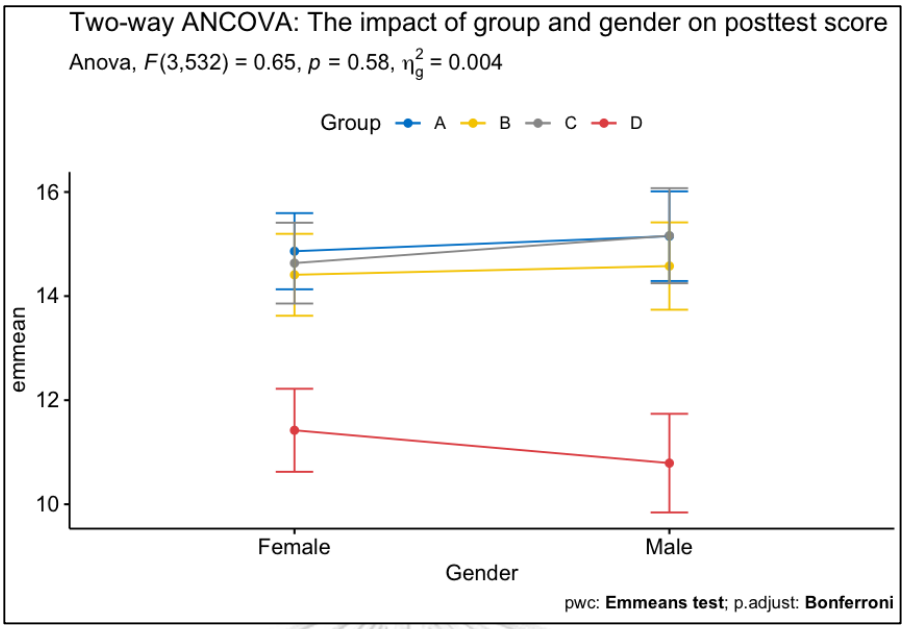


Figure 70 Line plot of prompting groups and gender on student's reading literacy posttest score

Table 58 Results of pairwise comparisons of group difference

Pair comparison		Mean Difference (I - J)	Std. Error	*p*
Group (I)	Group (J)			
A	B	.513	.410	1.000
A	C	.110	.419	1.000
A	D	3.901*	.426	.000
B	C	-.404	.423	1.000
B	D	3.387*	.430	.000
C	D	3.791*	.439	.000

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.3.2 Two-way ANCOVA results of different prompting conditions and levels of learning achievement on students' reading literacy posttest score

Table 59 presents the mean scores and standard deviations of group and learning achievement on the pretest and posttest reading literacy. All groups obtained higher posttest scores than the pretest score. The highest score on the pretest was obtained by very high achievers in all prompting groups except Group C. Those students in Group B scored higher than any other groups on the pretest (*M*=9.27, *SD*=3.80). With respect to the posttest score, the highest posttest score of very high achievers was Group C (*M*=15.10, *SD=2.93*), whereas Group A had the highest mean posttest scores for both high (*M*=14.93, *SD*=3.76) and mid achievers (*M*=15.17, *SD*=3.43). For the group of low achievers, Group B had the highest score (*M*=15.82, *SD*=3.24) as shown in Figure 71.

Table 59 Descriptive statistics of group and learning achievement on reading literacy posttest score

Group	Learning achievement	Pretest score		Posttest score	
		Mean	SD	Mean	SD
Group A	Very high	8.77	3.52	14.85	2.99
	High	8.45	3.01	14.93	3.76
	Mid	7.43	3.04	15.17	3.43
	Low	5.14	1.75	15.32	3.41
Group B	Very high	9.27	3.80	14.86	3.35
	High	7.95	3.16	13.41	3.84
	Mid	6.48	2.15	14.56	3.57
	Low	6.82	3.80	15.82	3.24
Group C	Very high	7.98	3.56	15.10	2.93
	High	8.02	2.71	14.54	2.88
	Mid	6.81	3.43	15.00	2.53
	Low	5.90	1.79	14.52	4.31
Group D	Very high	9.09	2.68	12.40	4.15
	High	8.57	3.09	10.14	3.46
	Mid	6.35	2.01	10.83	3.30
	Low	5.28	2.16	10.61	3.60

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

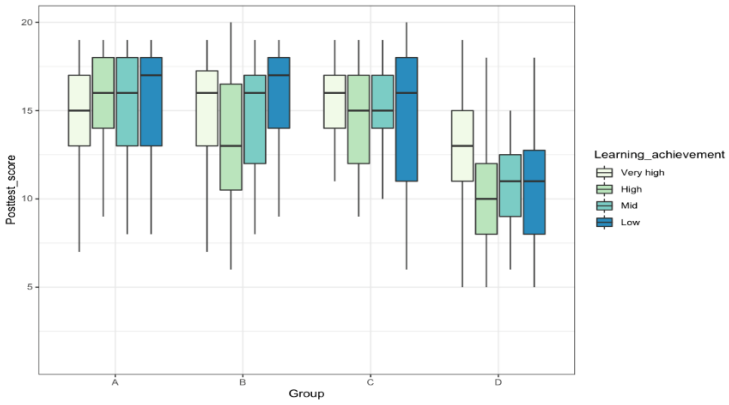


Figure 71 Students' posttest scores with respect to different prompting groups and levels of learning achievement

A two-way ANCOVA was conducted that examined the effect of prompting conditions and levels of learning achievement on students' reading literacy posttest score. The preliminary analyses investigating the homogeneity of slope, linearity, homogeneity of variances assumptions were satisfactory. As summarized in Table 60 and Figure 72, an interaction between group and learning achievement could not be demonstrated, $F(9, 524) = 1.08$, $p = .373$. In addition, there were no significant differences in the effect of levels of learning achievement on posttest score, $F(3, 524) = 0.37$, $p = .773$. However, significant differences in the impact of different prompting groups on posttest score were found, $F(3, 524) = 35.11$, $p < .05$. The pairwise comparison showed that Group D (control group) significantly received lower scores when compared with other prompting conditions as presented in Table 61.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 60 Two-way ANCOVA student's reading literacy posttest score with pretest score as covariate

Source	Univariate ANCOVA				
	SS	df	*MS*	*F*	*p*
Pretest	56.23	1	56.23	4.813*	.029
Group	1230	3	410.33	35.11*	.000
Learning achievement	13	3	32.10	0.37	.773
Group*Learning achievement	114	9	12.66	1.08	.373
Residuals	6122	524	11.68		

*$p < .05$

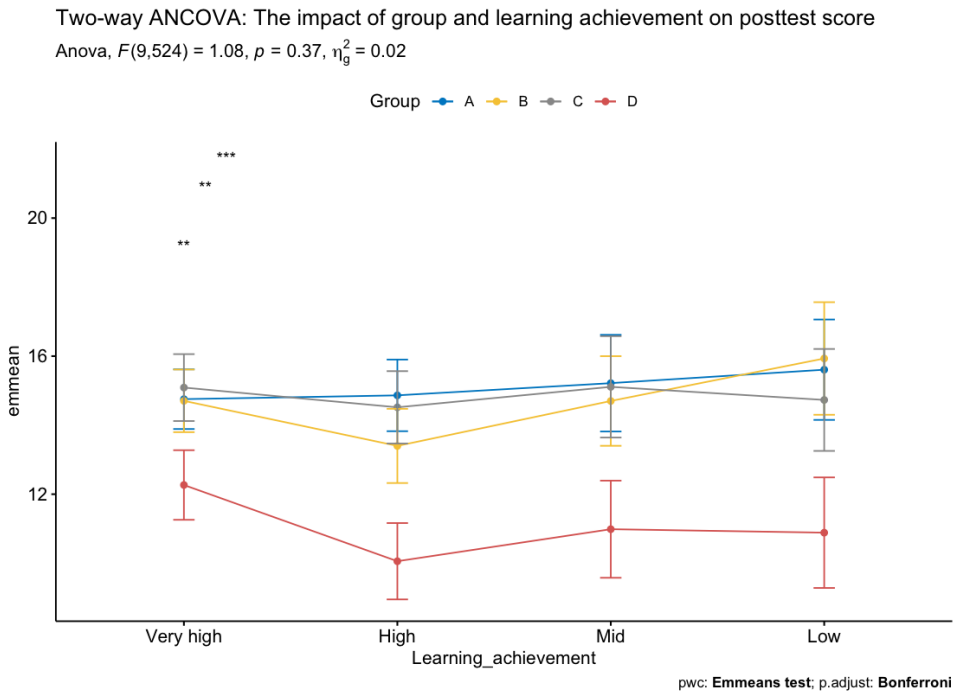


Figure 72 Line plot of prompting groups and levels of learning achievement on student's reading literacy posttest score

Table 61 Results of pairwise comparisons of group difference

Pair comparison		Mean Difference (I - J)	Std. Error	*p*
Group (I)	Group (J)			
A	B	.426	.443	1.000
A	C	.248	.444	1.000
A	D	4.062*	.450	.000
B	C	-.178	.453	1.000
B	D	3.635*	.459	.000
C	D	3.813*	.459	.000

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.2 Results of the effects of different promptings of the computerized dynamic assessment on growth of learning of Grade 9 students

2.2.1 Descriptive statistics of changes in reading literacy scores

The means and standard deviations are presented in Table 62. When considering the overall reading literacy scores, students obtained higher reading literacy scores across three time points as presented in Figure 73. All students scored 6.84 (*SD* = 4.06) in the first assessment, 8.86 (*SD* = 4.23) in the second assessment, and 9.89 (*SD* = 4.47) in the third assessment. Group A obtained the highest scores in understand subscale for the first assessment (*M* = 3.10, *SD* = 1.90) and second assessment (*M* = 3.58, *SD* = 1.90). They also received the highest score for the third time in locate information subscale (*M* = 3.82, *SD* = 1.66). Group B obtained the highest score in understand subscale (*M* = 3.25, *SD* = 1.98; *M* = 3.42, *SD* = 1.96; *M* = 3.95, *SD* = 1.91, respectively) and the lowest score in evaluate and reflect subscale for the three assessment periods (*M* = 1.45, *SD* = 1.27; *M* = 2.24, *SD* = 1.60; *M* = 2.37, *SD* = 1.62, respectively). Group C received the highest score in understand subscale for the first assessment time point (*M* = 3.01, *SD* = 2.00) and the third time point (*M* = 4.23, *SD* = 1.91), whereas they obtained the lowest score in evaluate and reflect subscale across three time points (*M* = 1.66, *SD* = 1.23; *M* = 2.22, *SD* = 1.51; *M* = 2.59, *SD* = 1.45, respectively). Group D obtained understand subscale as the highest score for the first assessment (*M* = 3.13, *SD* = 1.90) and they had the highest score for the second and third time points in locate information subscales (*M* = 3.25, *SD* = 1.84; *M* = 3.34, *SD* = 1.78, respectively).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 62 Descriptive statistics for growth of learning

Variable	Time	Sample									
		Group A (n = 148)		Group B (n = 139)		Group C (n = 131)		Group D (n = 123)		Total (*N* = 541)	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Locate information	1	2.20	1.68	2.21	1.67	2.29	1.71	1.82	1.66	2.14	1.69
	2	3.38	1.61	3.27	1.58	3.55	1.43	3.25	1.84	3.36	1.62
	3	3.82	1.66	3.78	1.69	4.01	1.53	3.34	1.78	3.74	1.68
Understand	1	3.10	1.90	3.25	1.98	3.01	2.00	3.13	1.90	3.12	1.94
	2	3.58	1.90	3.42	1.96	3.48	1.80	3.11	2.00	3.41	1.94
	3	3.78	1.90	3.95	1.91	4.23	1.91	3.22	2.08	3.80	1.97
Evaluate and reflect	1	1.62	1.17	1.45	1.27	1.66	1.23	1.59	1.35	1.58	1.25
	2	2.03	1.50	2.24	1.60	2.22	1.51	1.84	1.46	2.09	1.52
	3	2.35	1.52	2.37	1.62	2.59	1.45	2.02	1.61	2.34	1.56
Overall reading literacy	1	6.92	3.94	6.91	4.00	6.95	4.15	6.54	4.21	6.84	4.06
	2	8.99	4.21	8.94	4.35	9.25	3.72	8.20	4.59	8.86	4.23
	3	9.95	4.37	10.1	4.53	10.8	4.16	8.58	4.58	9.89	4.47

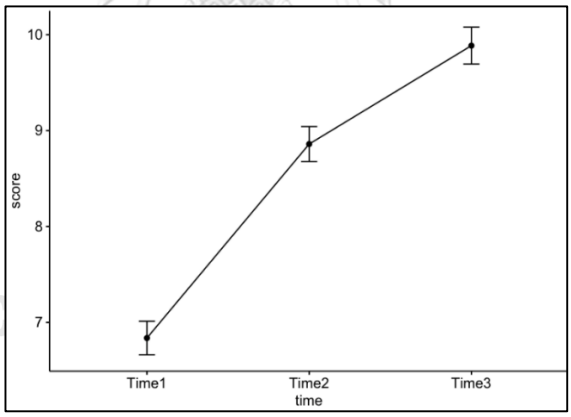


Figure 73 The trajectory plot of reading literacy for all students

2.2.2 Results of univariate latent growth curve models of changes in overall reading literacy and reading literacy subscales

Table 63 presents the model fit statistics. In the initial step in this analysis, unconditional growth models for each variable were examined to identify the statistical model that best describes the growth. Both the no-growth model and linear model were tested in each variable. The results showed that the model fit of reading literacy and each

subscale were satisfactory. The results found significant differences between the two models in overall reading literacy, $\Delta\chi^2$ =207.87, Δdf =2, $p < .05$. Moreover, there were significant differences between the no-growth and linear growth models in locate information, $\Delta\chi^2$ =313.21, Δdf =2, $p < .05$, understand, $\Delta\chi^2$ =37.93, Δdf =2, $p < .05$, and evaluate and reflect, $\Delta\chi^2$ =115.28, Δdf =2, $p < .05$. As a result, the linear model was used for the subsequent analyses.

Table 63 Fit indices of univariate growth curves for variables

Model	χ^2	*df*	TLI	CFI	SRMR	RMSEA	Model comparison	$\Delta\chi^2(\Delta df)$	*p*
Locate information (intercept only)	322.680	*4*	-0.300	0.000	0.304	0.384	Intercept only vs. linear	313.21* (2)	.000
Locate information (linear)	9.465	2	0.939	0.959	0.038	0.083			
Understand (intercept only)	47.427	4	0.787	0.716	0.086	0.142	Intercept only vs. linear	37.93*(2)	.000
Understand (linear)	9.494	2	0.926	0.951	0.043	0.083			
Evaluate and reflect (Intercept only)	116.134	4	-0.008	0.000	0.149	0.228	Intercept only vs. linear	115.28* (2)	.000
Evaluate and reflect (linear)	0.851	2	1.000	1.000	0.012	0.000			
Overall reading literacy (Intercept)	218.610	4	0.397	0.197	0.202	0.315	Intercept only vs. linear	207.87*(2)	.000
Overall reading literacy (linear)	10.744	2	0.951	0.967	0.049	0.090			

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

The researchers conducted the univariate latent growth curve model for all students and the multiple group latent growth curve model to assess prompting-based differences. The mean and variance scores of the intercepts and slopes were considered. Table 64 and Figures 74-75 present the results of trajectory growth of reading literacy and reading literacy subscales.

For the analysis of the unconstrained univariate model (Model 1), the univariate linear growth models showed that reading literacy had an initial score of

6.837 (p < .05), with an increase of 1.015 (p < .05). Regarding reading literacy subscales, locate information had an initial score of 2.158 (p < .05), with an increase of 0.551 (p < .05). Moreover, understand had an initial score of 3.110 (p < .05), with an increase of 0.207 (p < .05). Evaluate and reflect had an initial score of 1.577 (p < .05), with an increase of 0.253 (p < .05).

For the multiple group latent growth curve analysis (Model 2), as shown in Figure 76, the intercept mean and slope mean of the overall reading literacy were significant for all four groups. Group C obtained the highest initial reading literacy score (intercept mean = 6.947, p < .05) and there was higher increase in overall reading literacy scores over time (slope mean = 1.253, p < .05). The intercept variance and the slope variance reached significance for all groups, indicating that there were individual differences in the initial values and their rate of change in reading literacy score. There was negative and significant covariance between the intercept and slope in overall reading literacy. It can be concluded that individuals with higher starting scores did not perform better across time when compared with those with lower starting scores.

The intercept mean and slope mean of locate information subscales were significant for all four groups. It indicates that there was a positive change in locate information scores over time. Group C presented the highest starting locate information score (intercept mean = 2.309, p < .05) and there was higher increase in scores over time (intercept slope = 0.583, p < .05). The intercept variance and the slope variance were significant for all groups, indicating that there were individual differences in the initial values and their rate of change in locate information score. The covariance between the intercept and slope of locate information was significant and negative (β= .578 - .737), suggesting that individuals with lower starting scores exhibited linear growth than did those with higher starting scores.

Furthermore, with respect to understand subscale, the intercept mean was significant for all groups and there was positive change in understand score over

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

time. The highest starting score in understand subscale obtained by Group B (intercept mean = 3.228, $p < .05$), representing their highest starting score. However, differences were identified among prompting groups in the slope mean. Group B received a significant increase in understand subscale score (slope mean = 0.205, $p < .05$); on the other hand, the trend of their scores were lower than other groups. Group C showed the highest increase in understand subscale score (slope mean = 0.364, $p < .05$). Group D received the lowest increased score and there was no significant slope mean (slope mean = 0.021), indicating that there was no change in reading literacy over time. The intercept variance was significant for all groups, suggesting that there were individual differences in the initial score of understand subscale. However, Group A received no significant difference in slope variance (slope variance = 0.091), suggesting that individual differences were not identified in the rate of change in Group A. The covariance between understand intercept and understand slope was significant and negative. This indicates that students with lower starting values had steeper trajectories than their higher performing peers.

Concerning evaluate and reflect subscale, the intercept mean and slope mean were significant for all four groups and there was positive change in evaluate and reflect score over three time points The highest starting score in evaluate and reflect subscale obtained by Group C (intercept mean = 1.649, $p < .05$). However, Group B revealed the highest increased score in evaluate and reflect subscale over time (slope mean = 0.324, $p < .05$), followed by Group C (slope mean = 0.306, $p < .05$). The intercept variance was significant for all groups, suggesting that there were individual differences in the initial score of evaluate and reflect subscale. However, the slope variance was significant only for Group D (slope variance = 0.124, $p < .05$), suggesting that unlike other groups, Group D had individual differences in the rate of change in evaluate and reflect subscales. No significant covariance was identified between the intercept and slope in evaluate and reflect subscale in all groups, except Group C (β = -.330). This suggests that individuals with higher starting

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

scores in Group C did not perform linear growth better than those with lower starting scores.

Table 64 Univariate model of growth in reading literacy and reading literacy subscales

Variable	Model	Intercept		Slope		Covariance	Correlation (Coefficient **β**)
		Mean	Variance	Mean	Variance		
Locate information	**Model 1**	2.158*	1.781*	0.551*	0.165*	-0.338*	-.624
	Model 2						
	Group A	2.210*	1.696*	0.550*	0.097*	-0.244*	-.601
	Group B	2.211*	1.625*	0.524*	0.220*	-0.346*	-.578
	Group C	2.309*	1.829*	0.583*	0.201*	-0.447*	-.737
	Group D	1.867*	1.851*	0.555*	0.132*	-0.301*	-.609
Understand	**Model 1**	3.110*	2.702*	0.207*	0.214*	-0.562*	-.738
	Model 2						
	Group A	3.104*	2.542*	0.229*	0.091	-0.434*	-.903
	Group B	3.228*	2.799*	0.205*	0.288*	-0.675*	-.753
	Group C	2.977*	2.877*	0.364*	0.273*	-0.726*	-.820
	Group D	3.124*	2.564*	0.021	0.155*	-0.379*	-.600
Evaluate and reflect	**Model 1**	1.577*	0.547*	0.253*	0.086*	-0.066	-.302
	Model 2						
	Group A	1.610*	0.385*	0.236*	0.073	-0.023	-.135
	Group B	1.469*	0.619*	0.324*	0.063	-0.031	-.156
	Group C	1.649*	0.432*	0.306*	0.065	-0.055*	-.330
	Group D	1.580*	0.762*	0.140*	0.124*	-0.160	-.520
Overall reading literacy	**Model 1**	6.837*	15.353*	1.015*	1.471*	-3.115*	-.656
	Model 2						
	Group A	6.920*	14.350*	1.016*	0.846*	-2.324*	-.667
	Group B	6.905*	14.731*	1.051*	1.658*	-3.182*	-.644
	Group C	6.947*	15.877*	1.253*	1.927*	-4.105*	-.742
	Group D	6.545*	16.577*	0.723*	1.370*	-3.016*	-.633

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

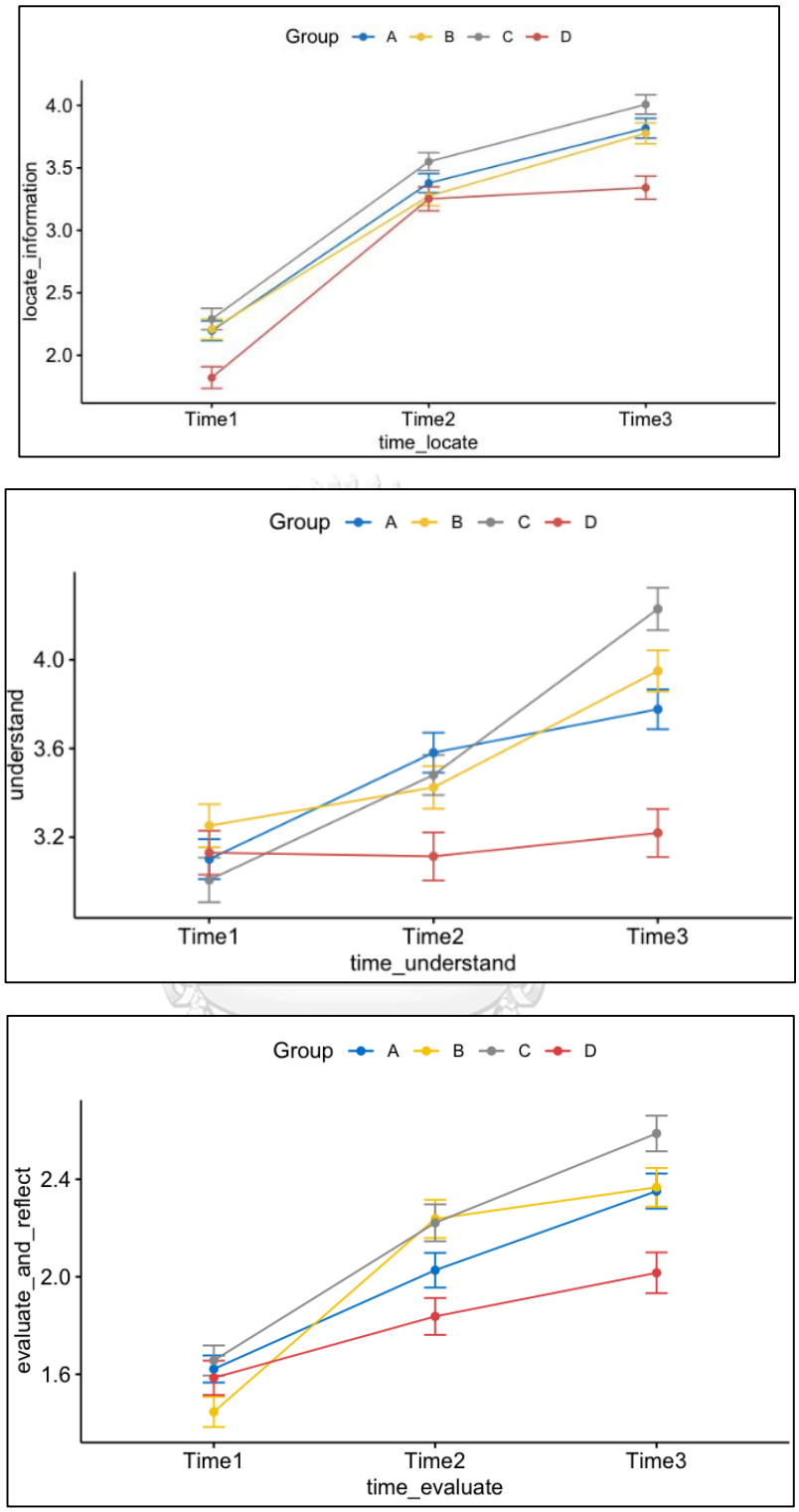


Figure 74 The trajectory plots of reading literacy subscales for different prompting groups

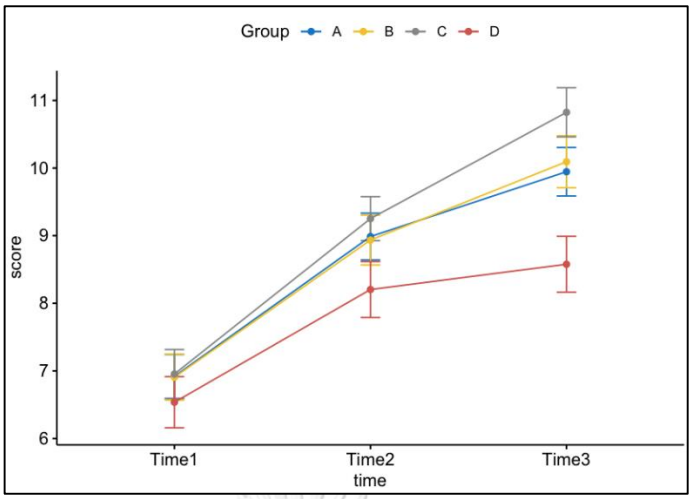
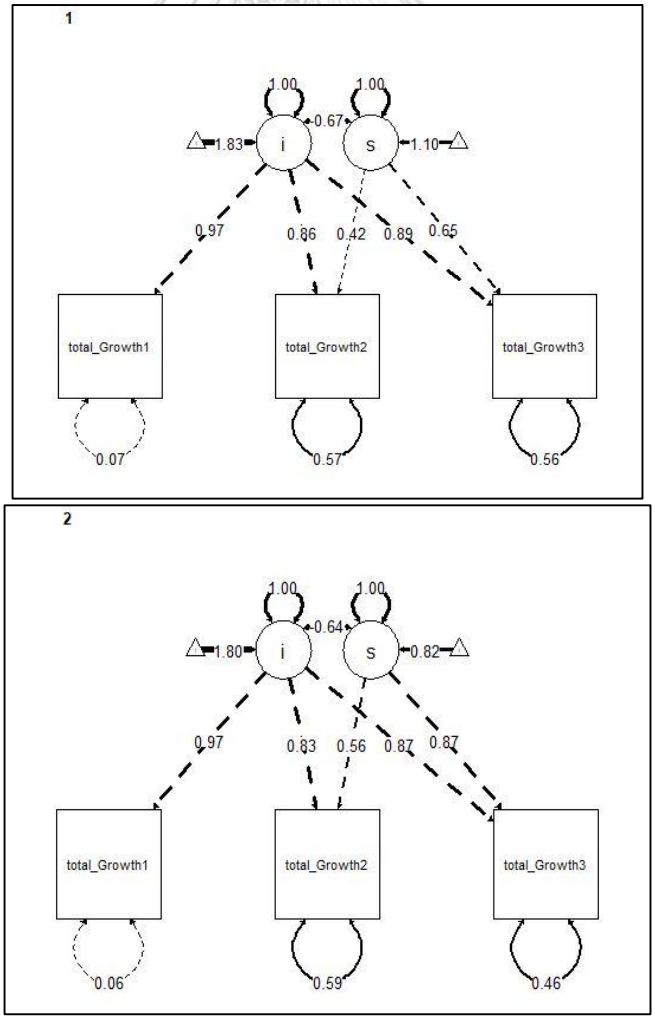


Figure 75 The trajectory plots of overall reading literacy for different prompting groups



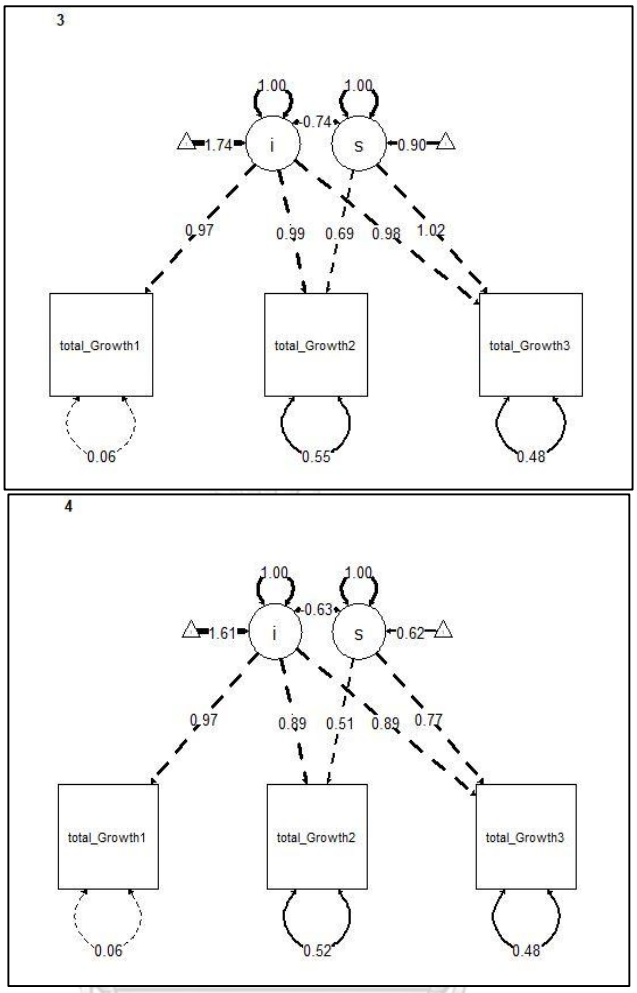


Figure 76 The univariate latent growth curves of overall reading literacy for different prompting groups

Note. All loadings shown in Figure 76 were standardized factor loadings.

2.2.3 Results of multivariate latent growth curve model of the associations of changes in reading literacy subscales

As shown in Table 65, the researchers examined the associations between changes over time in reading literacy subscales. The multivariate model was established to investigate the associations between the intercepts of locate information, understand, and evaluate and reflect and the associations between the slopes of locate information, understand, and evaluate and reflect. The model fit indices of the multivariate latent growth curve model presented good fit to the data, χ^2 = 52.427, df = 15, TLI = 0.948, CFI = 0.978, and RMSEA = 0.068.

Table 65 Fit indices of multivariate growth curves for variables

Model	χ^2	*df*	TLI	CFI	SRMR	RMSEA
Multivariate latent growth model	52.427	*15*	0.948	0.978	0.036	0.068

As depicted in Table 66 and Figure 77, the model showed positive and bidirectional associations among the three subscale variables. The results showed that locate intercept had statistically significant effects on understand intercept (β = .628, p = .000) and evaluate intercept (β = .721, p = .000), indicating an association between the better locate information score and higher understand and greater evaluate and reflect scores at the initial assessment. Moreover, the association between understand intercept and evaluate intercept was significantly positive (β = .539, p = .000). Thus, the initial increased score of each subscale was associated with the higher scores of other two subscales.

Regarding the slope, locate slope had no statistically significant effect on understand slope (β = .228, p = .063). The change in locate information score was not related to changes in understand subscale. The non-significant effect was also found in the understand slope on evaluate slope (β = .307, p = .056). Higher understand score did not correlate with an increase in evaluate and reflect subscale across time points. However, the positive association between locate slope and

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

evaluate and reflect slope was significant (β = .365, p = .016). The higher locate information score was correlated with the increase in evaluate and reflect subscale.

Locate intercept had statistically significant effect on understand slope (β = -.207, p = .018), indicating that students with higher initial locate information score were related to lower understand score over time. On the contrary, the non-significant effect between locate information intercept and evaluate and reflect slope was found (β = -.048, p = .648). Moreover, there was no significant association between understand intercept and evaluate and reflect slope (β = .140, p = .126). Thus, neither locate information nor understand subscales were related to the growth in evaluate and reflect subscale.

In summary, the relationship in growth trajectories among locate information, understand, evaluate and reflect subscales are likely to be strong and positive at the first time point, but the trend of each subscale did not related to other subscales, except the relationship between locate slope and evaluate and reflect slope and association between locate intercept and understand slope.

Table 66 Coefficient estimates and standard error of the multivariate latent growth model for reading literacy variables of all students (N =541)

Bidirectional relationship	Covariance (β)	SE	*p*
Locate intercept <-> Understand intercept	.628*	.123	*.000*
Locate intercept <-> Evaluate intercept	.721*	.078	*.000*
Understand intercept <-> Evaluate intercept	.539*	.087	*.000*
Locate slope <-> Understand slope	.228	.020	*.063*
Locate slope <-> Evaluate slope	.365	.016	.016
Understand slope <-> Evaluate slope	.307	.019	.056
Locate intercept <-> Understand slope	-.207	.047	.018
Locate intercept <-> Evaluate slope	-.048	.036	.648
Understand intercept <->Evaluate slope	.140	.040	.126

*p < .05

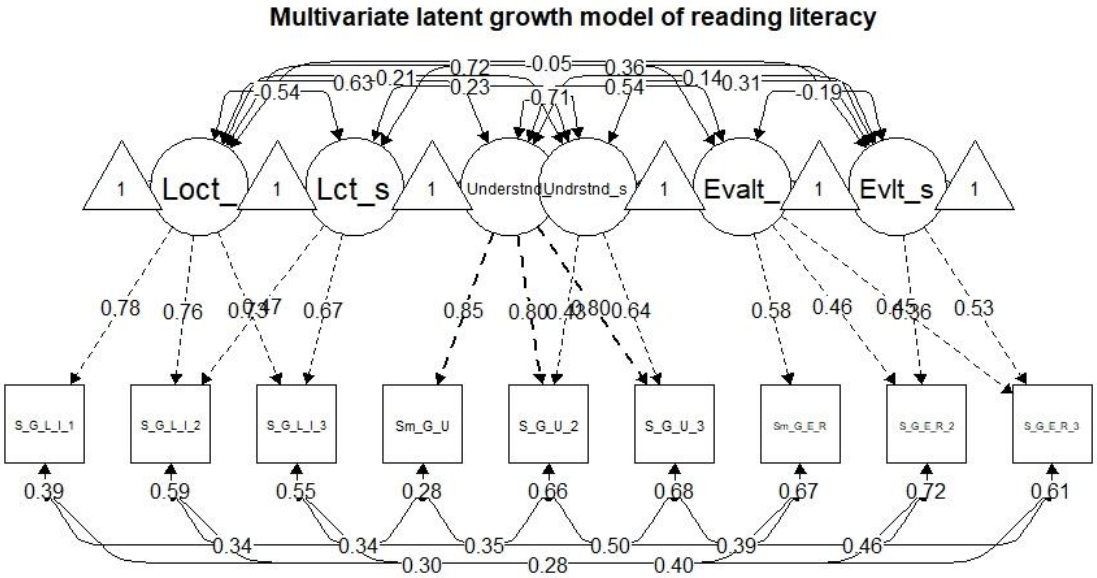


Figure 77 The multivariate latent growth curves of overall reading literacy

Note. All loadings shown in Figure 77 were standardized factor loadings.

Group A revealed that intercept on locate information had statistically significant effects on understand intercept and evaluate and reflect intercept. The result found the highly positive associations between the locate intercept and understand intercept (β = 0.782, p = .000) and between the locate intercept and the evaluate intercept (β = 0.770, p = .000). Accordingly, increased locate information score was highly associated with higher understand score and higher evaluate and reflect subscale score at the initial assessment. Regarding the slope, locate slope had no significant effects of slopes on understand subscale (β = 0.040, p = .92) and evaluate and reflect subscale (β = 0.264, p = .488). Furthermore, understand slope had no significant effects on evaluate and reflect slope (β = -0.691, p = .116). The increase in one subscale was not related to the increase in other subscale scores over three time points. The locate intercept had no associations with understand subscale and evaluate and reflect subscales. However, there was significant association between the understand intercept and evaluate and reflect slope (β =0.405, p = .035). Thus, higher understand score in the initial assessment was related to the greater evaluate and reflect scores over three time points as shown in Table 67.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Group B showed positively bidirectional associations among the three subscale variables. For the intercept, there were positive association between the locate intercept and understand intercept (β = 0.489, p = .000). Moreover, there was association between the locate intercept and evaluate and reflect intercept (β = 0.791, p = .000). This suggests that the increased locate information score was related to the increase in understand subscale and evaluate and reflect subscale at the first assessment. Concerning the associated slopes, there were non-significant associations between all three subscale slopes, suggesting that the increase in one subscale did not exhibit the increase in other subscales across three time points as depicted in Table 67.

Group C obtained the positive associations among bidirectional variables. There were positive association between the locate and understand intercept (β = 0.387, p = .000) and between the locate and evaluate intercept (β = 0.481, p = .000). Moreover, it was found the positive relationship between the understand intercept and evaluate intercept (β = 0.385, p =.010). Thus, increased locate information score was correlated with increased understand and evaluate and reflect subscale scores at the initial assessment. Concerning the associated slopes, the significant and positive associations of slope were also found between understand subscale and evaluate and reflect subscale scores (β = 0.653, p = .025). It indicates that increased understand score was correlated with an increase in evaluate and reflect subscale over three time points. However, non-significant correlations were found between slope in locate information and slopes in understand subscale (β = 0.253, p = .164) and in evaluate and reflect subscales (β = 0.363, p = .174). Thus, increased locate information score did not relate to increased understand and evaluate and reflect scores over time as seen in Table 67.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Group D indicated a positive association among bidirectional variables. There were positive association between the locate and understand intercept (β = 0.927, p = .000) and between the locate and evaluate intercept (β = 0.818, p = .000), as well as the association between the understand and evaluate intercepts (β = 0.655, p = .000). This suggested that an increased locate information score was correlated with increased

understand and evaluate and reflect subscale scores at the first time point. Concerning the slopes, there were non-significant associations between all three subscale slopes. This exhibited that the increase in one subscale did not exhibit the increase in other subscales across three time points. However, there was significant association between the locate intercept and understand slope(β =-0.584, p = .010). Thus, higher locate information score at the first time point was related to lower understand score over three time points as presented in Table 67.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 67 Coefficient estimates and standard error of the multivariate latent growth model for reading literacy subscale variables

Path	Group A			Group B			Group C			Group D		
	Path coefficient (β)	SE	*p*	Path coefficient (β)	SE	*p*	Path coefficient (β)	SE	*p*	Path coefficient (β)	SE	*p*
Locate intercept <-> Understand intercept	0.782*	0.24	*.000*	0.489*	0.24	*.000*	0.387*	0.21	*.000*	0.927*	0.29	*.000*
Locate intercept <-> Evaluate intercept	0.770*	0.13	*.000*	0.791*	0.16	*.000*	0.481*	0.13	*.000*	0.818*	0.18	*.000*
Understand intercept <-> Evaluate intercept	0.649*	0.15	*.000*	0.473*	0.17	*.000*	0.385*	0.15	*.0100*	0.655*	0.21	*.000*
Locate slope <-> Understand slope	0.040	0.03	*.920*	0.224	0.04	*.210*	0.253	0.03	*.164*	0.455	0.05	*.263*
Locate slope <-> Evaluate slope	0.264	0.02	.488	0.500	0.03	.100	0.363	0.03	.174	0.352	0.03	.288
Understand slope <-> evaluate slope	-0.691	0.03	.116	0.400	0.04	.220	0.653*	0.03	.025	0.403	0.03	.236
Locate intercept <-> Understand slope	-0.428	0.08	0.070	-0.122	0.09	.430	-0.081	0.08	.568	-0.584*	0.10	.010
Locate intercept <-> Evaluate slope	0.234	0.06	0.270	-0.389	0.07	.100	0.013	0.06	.946	-0.144	0.08	.470
Understand intercept <-> Evaluate slope	0.405*	0.07	0.030	0.059	0.07	.760	0.049	0.07	.777	0.098	0.09	.583

*$p < .05$

2.2.4 Informative data of parameter estimation for individual growth trajectory

The researchers estimated not only the growth trajectory of group difference but also the individual difference over three time points. The example of individual growth trajectory is randomly shown in Figure 78.

Concerning the performance of examples in Group A, student no. 135 had little change in reading literacy over time. His/her ability rose slowly at the second time point and little decreased at the third time point. Whereas, performance of student no. 420 grew gradually over time. For Group B, student no. 528 remained constant with no growth. In contrast, student no. 309 and student no. 384 increased sharply during the first two time points, but slowed down after completing the third assessment. With respect to growth of examples in Group C, student no. 72 had little change. His/her performance dropped during the second time point, but slightly went up at the third time point. In contrast, students no. 217 and no. 490 gradually increased over time. Regarding examples of Group D, students no. 166 and no.398 had little change in their growth trajectories. Student no. 291 steadily decreased, whereas student no. 224 steadily rose over time.

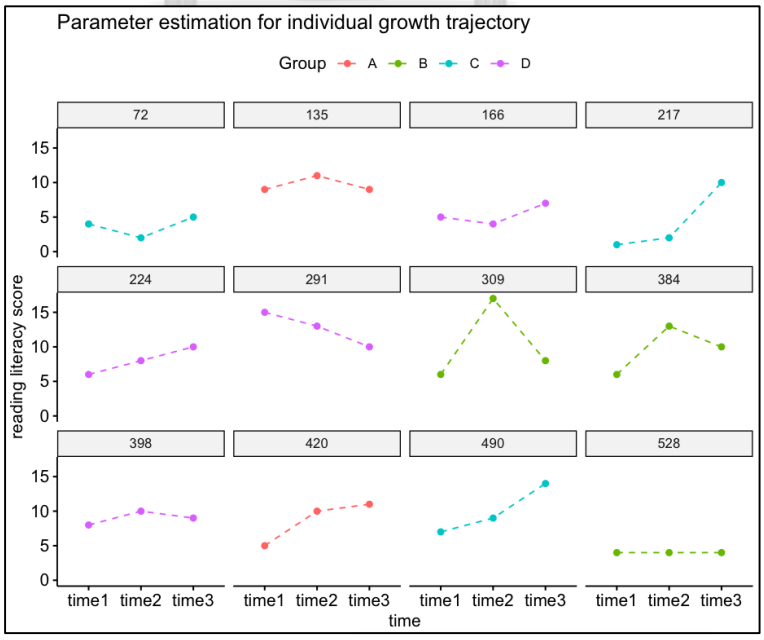


Figure 78 Examples of Individual growth trajectories over three time points

2.2.5 Informative data of growth patterns of reading literacy

The researcher further investigated the patterns of growth trajectories of different prompting methods in accordance with initial reading literacy scores. The initial reading literacy score measured at first time points of students in different prompting conditions were analyzed and compared. This study categorized initial reading literacy scores in accordance with the Quartile ranking. The first group was students who scored 25% of the lowest or those who scored below four points at the first time point. They were called low initial reading literacy. For the second group, they were called medium-low initial reading literacy. Those were students who scored between 50-75% or those who scored between four to six points at the first time point. The third group was medium-high initial reading literacy. There were groups of students who scored between 25-50% or those who scored between six to ten points at the first time point. For the fourth group, they were high initial reading literacy. They were those who scored top 25% or those who scored between ten to twenty points at the first time point. Means and standard deviations of reading literacy scores across groups with differently initial reading literacy were shown in Table 68.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 68 Means and standard deviations of reading literacy scores across groups with differently initial reading literacy over three time points

Initial reading literacy ability	Group	Time 1		Time 2		Time 3	
		M	*SD*	*M*	*SD*	*M*	*SD*
Low	Group A	2.57	1.35	6.80	3.60	7.43	3.12
	Group B	2.77	1.11	7.40	4.84	8.09	4.30
	Group C	3.00	1.03	7.96	4.31	9.40	4.29
	Group D	2.59	1.22	5.65	4.21	6.59	4.53
Medium-low	Group A	5.59	0.50	7.59	3.58	8.50	4.64
	Group B	5.62	0.49	9.33	4.21	12.20	3.86
	Group C	5.39	0.49	9.00	2.52	10.00	4.17
	Group D	5.65	0.49	7.29	2.89	9	3.89

Initial reading literacy ability	Group	Time 1		Time 2		Time 3	
		M	*SD*	*M*	*SD*	*M*	*SD*
Medium-high	Group A	8.45	1.19	9.89	4.01	11.40	4.00
	Group B	8.26	1.14	9.06	3.78	9.70	4.18
	Group C	8.17	1.31	9.03	3.03	11.60	3.93
	Group D	8.50	1.19	9.94	4.35	9.62	4.28
High	Group A	12.60	1.69	12.20	3.50	12.90	3.89
	Group B	13.30	2.44	11.20	3.44	12.80	4.21
	Group C	13.10	1.94	11.60	2.98	12.90	3.22
	Group D	13.20	1.91	12.10	2.77	11.20	3.82

1) Low initial reading literacy score

The growth pattern of the low initial reading score showed that they received the low scores at the first time but increased gradually over time. Mixed prompting group obtained the highest score over time, followed by Error-explanation prompting group, whereas Verification prompting group obtained the lowest reading literacy score over three time points as shown in Figure 79.

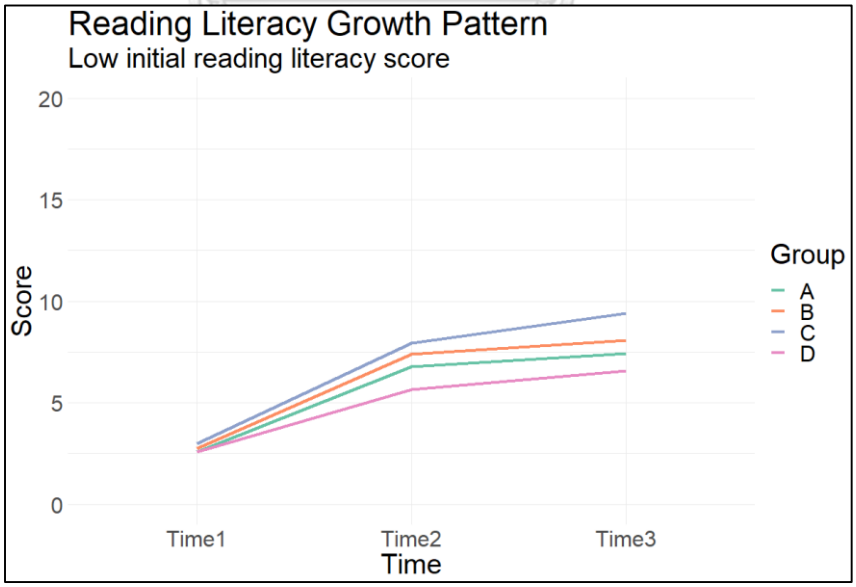


Figure 79 Growth trajectories of reading literacy of low initial reading literacy score

2) Medium-low initial reading literacy score

The growth pattern of the medium-low initial reading literacy groups showed that they received similar scores at the first time point. Error-explanation prompting group increased gradually and obtained the highest growth trajectory over time, followed by Mixed prompting group. Instructional prompting group received the lowest at the third time point as shown in Figure 80.

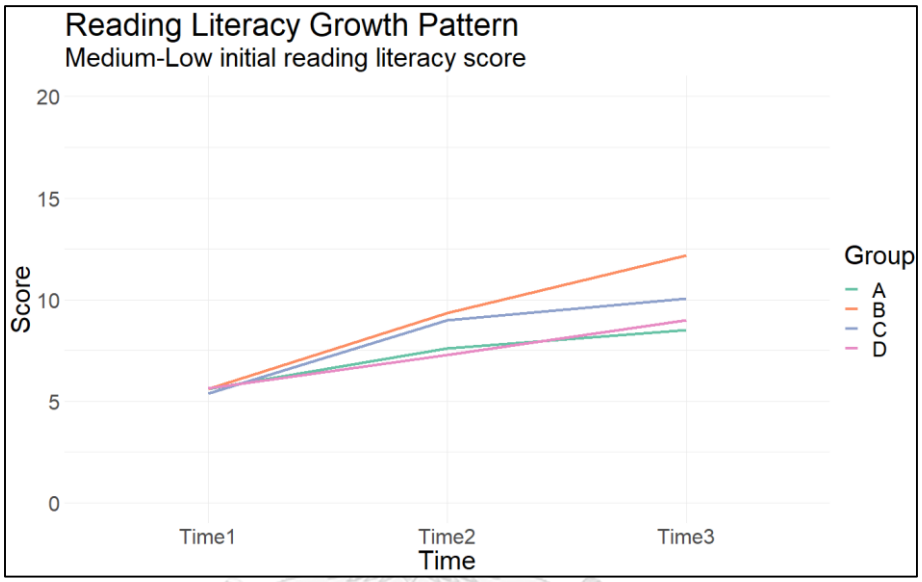


Figure 80 Growth trajectories of reading literacy of medium-low initial reading literacy score

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3) Medium-high initial reading literacy score

The growth pattern of the medium-high initial reading literacy groups showed that they received similar scores at the first time point. Instructional prompting group increased steadily over three time points. Mixed prompting group increased sharply during the second assessment to the third assessment. Error-explanation prompting group grew slightly over time as depicted in Figure 81.

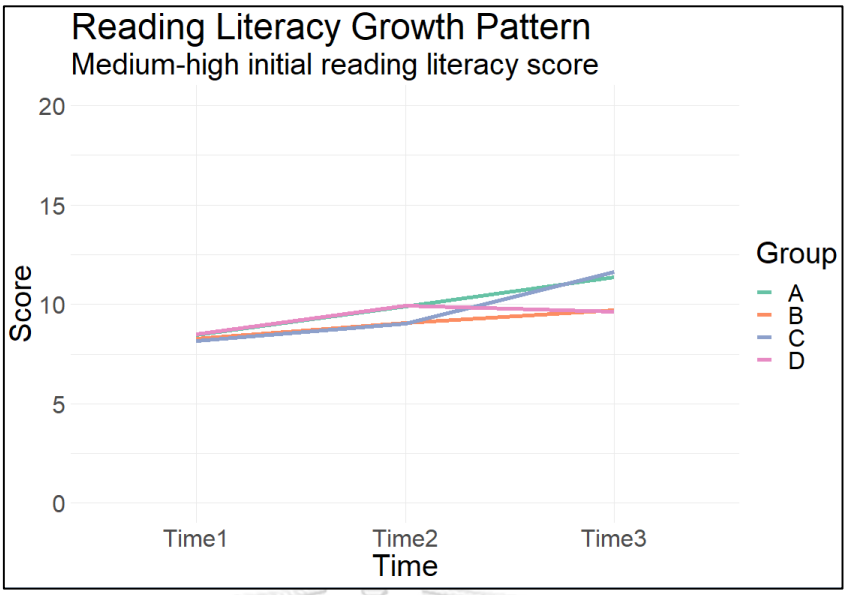


Figure 81 Growth trajectories of reading literacy of medium-hiigh initial reading literacy score

4) High initial reading literacy score

The growth pattern of high initial reading literacy groups showed that the reading literacy of Verification prompting group decreased gradually over time. Error-explanation prompting group also dropped during the second assessment, but rose at the third assessment. Instructioanl prompting group had slightly changes over time as presented in Figure 82.

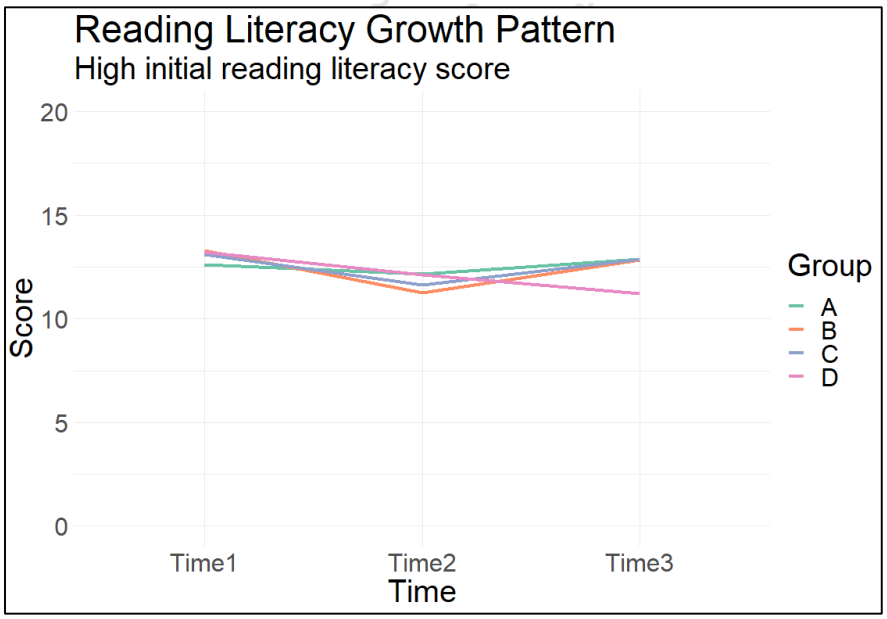


Figure 82 Growth trajectories of reading literacy of high initial reading literacy score

2.3 Results of the effects of different promptings of the computerized dynamic assessment on learning potential of Grade 9 students

2.3.1 Estimation of student's learning availability from the growth curve trajectory

2.3.1.1 Estimation of student's learning availability using the Michaelis-Mentens curve

The researcher analyzed the availability score using the Michaelis-Mentens curve. Availability score was defined as the difference between an individual's current ability and estimated capacity, which is related to the Vygotsky's conceptual zone of proximal development (ZPD). It is the absolute distance between the current ability and estimated capacity. Michaelis-Menten equation was chosen to estimate the capacity parameter.

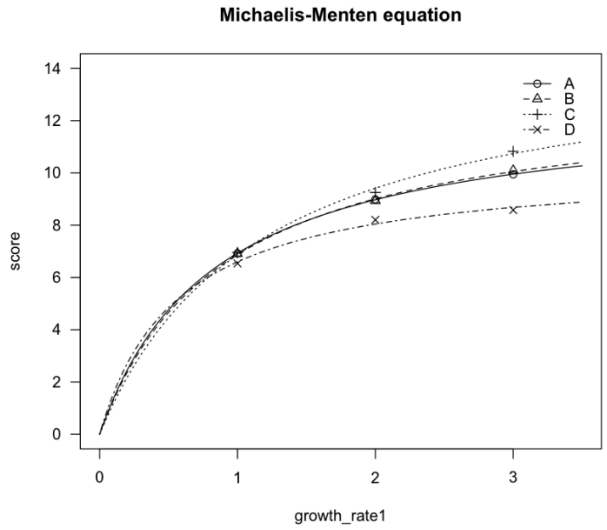


Figure 83 The Michaelis-Mentens curve of reading literacy across three time points

The researchers conducted computerized dynamic assessment for reading literacy over three time points to examine the availability distance between the estimated capacity and current reading literacy score of each individual. The mean and standard deviations of capacity and ability scores of each time point were shown in Table 69. Group C had higher mean reading literacy score than other groups at the first time point (*M* = 6.95, *SD* = 3.02), followed by Group A (*M* = 6.92, *SD* = 2.45) and Group

B (*M* = 6.91, *SD* = 2.88). For the measurement at the second time point, the two highest mean ability scores obtained by Group C (*M* = 9.25, *SD* = 2.39) and Group A (*M* = 8.99, *SD* = 2.64). For the third time point, Group A received the lower ability score than Group B (*M* = 9.95, *SD* = 2.71; *M* = 10.1, *SD* = 2.81, respectively), whereas Group D obtained the lowest mean ability score (*M* = 8.58, *SD* = 2.91). Their mean capacity scores over three time points were 10.1 (*SD* = 2.61) in Group A, 10.1 (*SD* = 2.58) in Group B, 10.4 (*SD* = 2.37) in Group C, and 9.38 (*SD* = 2.80) in Group D. The Michaelis-Mentens curve of reading literacy across three time points is depicted in Figure 83.

Table 69 Means and standard deviations of capacity and ability scores over three time points

Group	Capacity		Ability (Time 1)		Ability (Time 2)		Ability (Time 3)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Group A	10.1	2.61	6.92	2.45	8.99	2.64	9.95	2.71
Group B	10.1	2.58	6.91	2.88	8.94	2.81	10.1	2.87
Group C	10.4	2.37	6.95	3.02	9.25	2.39	10.8	2.68
Group D	9.38	2.80	6.54	2.71	8.20	2.68	8.58	2.91

2.3.1.2 One-Way MANOVA results of availability scores across three time points of different prompting groups

As shown in Table 70, Group C obtained the highest availability score during the initial assessment (*M* = 3.42, *SD* = 3.02) and reached the estimated capacity at the third time point (*M* = -0.44, *SD* = 2.68), followed by Group B with the availability score was 3.17 (*SD* = 2.88) for the first assessment, 1.14 (*SD* = 2.81) for the second assessment and -.01 (*SD* = 2.87) for the third assessment. However, Group D had the lowest initial availability score (*M* = 2.84, *SD* = 2.71) and their obtained availability score was .80 (*SD* = 2.91), which doesn't reach their estimated capacity. Trajectory plots of reading literacy ability and estimated capacity over three time points are presented in Figure 84.

Table 70 Means and standard deviations of availability scores over three time points

Group	Availability score 1		Availability score 2		Availability score 3	
	M	*SD*	*M*	*SD*	*M*	*SD*
Group A	3.14	2.45	1.07	2.64	.10	2.71
Group B	3.17	2.88	1.14	2.81	-.01	2.87
Group C	3.42	3.02	1.12	2.39	-.44	2.68
Group D	2.84	2.71	1.18	2.68	.80	2.91

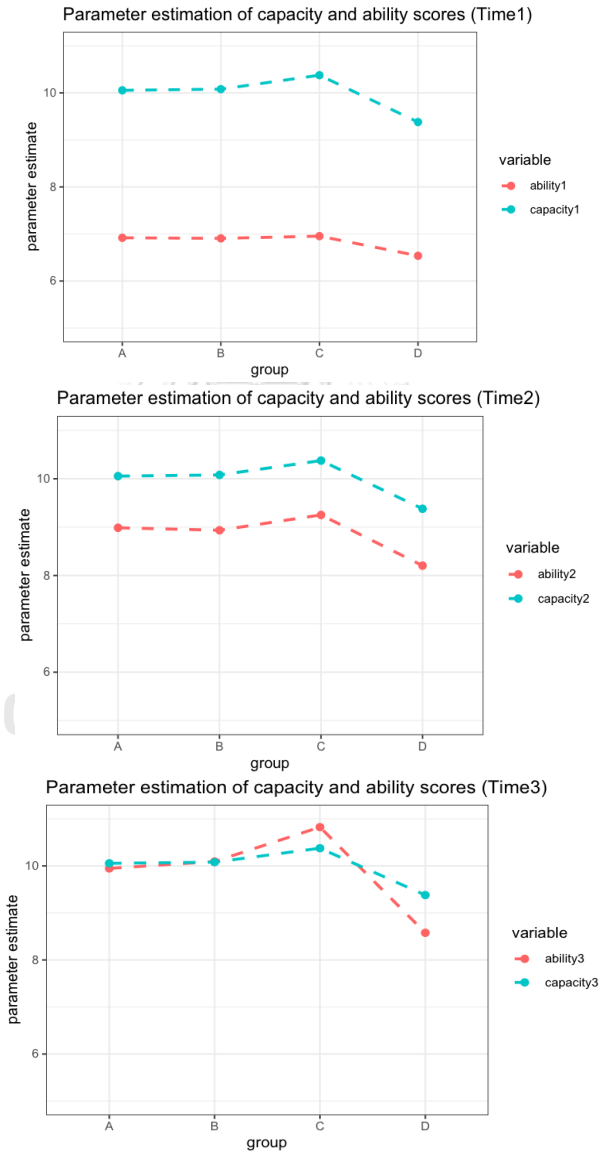


Figure 84 Trajectory plots of reading literacy ability and estimated capacity over three time points

One-way MANOVA was performed to investigate availability scores over three time points (availability score measured at the first time point, availability score measured at the second time point, and availability score measured at the third time point). The independent variable was groups received different types of promptings (instructional prompting (Group A), error-explanation prompting (Group B), mixed prompting (Group C), verification prompting (Group D). The assumption of homogeneity of variance-covariance matrices was satisfactory. Thus, Wilks' statistic was preferred. The results revealed that there were significant group effects of availability scores over three time points, Wilks' lambda = .967, $F(9, 1302) = 1.98$, $p < .05$, shown in Table 71.

The results of univariate tests were further conducted for availability scores measured at the different time points in Table 72. It indicated that there were significant differences among different prompting groups in terms of availability score measured at the third time point, $F(3,537) = 4.360$, $p < .05$. Because the homogeneity of variance assumption was met, follow-up pairwise comparisons using the Bonferroni post-hoc test were further analyzed. The result revealed that Group D had significantly higher availability score when compared with Group C as presented in Table 73. The availability scores over three time points of different prompting groups were depicted in Figure 85.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 71 One-way MANOVA result of availability scores of different prompting groups

	Value	*F*	*Hypothesis df*	*Error df*	*p*
Pillai's trace	.032	1.97*	9	1611	.039
Wilks' lambda	.967	1.98*	9	1302	.037
Hotelling's trace	.033	2.00*	9	1601	.035
Roy	.033	5.96*	3	537	.000

*$p < .05$

Table 72 Analysis of variance result of availability scores of different prompting groups

Source	DFn	DFd	Sum of Squares	Mean Square	F	*p*
Availability score 1 (Time 1)	3	537	21.37	7.124	.931	.425
Availability score 2 (Time 2)	3	537	.84	.282	.041	.989
Availability score 3 (Time 3)	3	537	101.89	33.966	4.360*	.005

*$p < .05$

Table 73 Results of pairwise comparisons of group difference

Variable	Pair comparison		Mean Difference (I - J)	Std. Error	*p*
	Group (I)	Group (J)			
Availability score 3	A	B	.123	.329	1.000
	A	C	.557	.334	.579
	A	D	-.693	.340	.253
	B	C	.434	.339	1.000
	B	D	-.816	.345	.111
	C	D	-1.250*	.350	.002

*$p < .05$

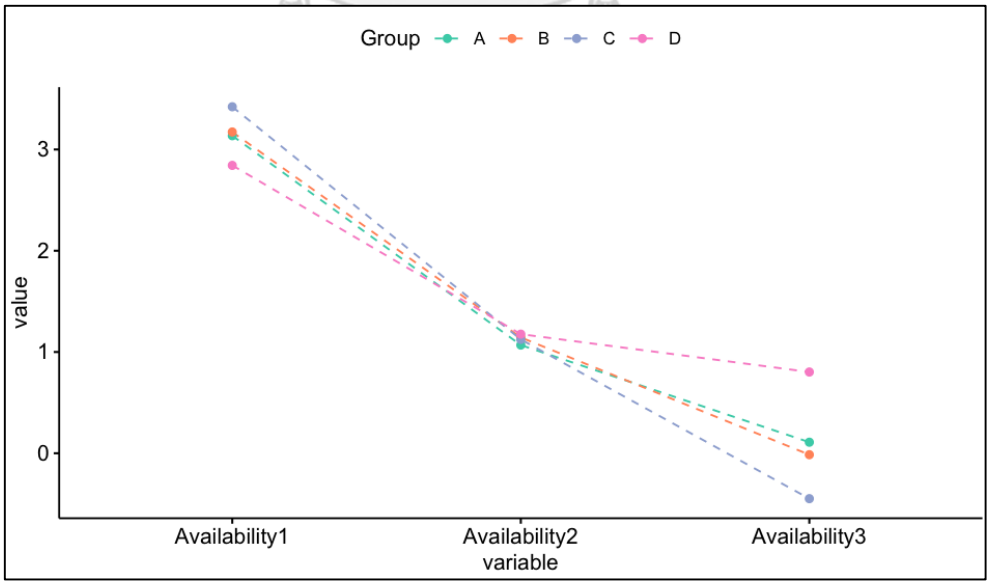


Figure 85 Line graphs of availability scores over three time points

2.3.2 Results of two-way mixed-effects ANOVA of different prompting groups and different time points on mediated scores

As seen in Table 74 and Figure 86, Group C received the highest mediated scores during the first assessment (*M* = 62.82, *SD* = 9.35), followed by Group A (*M* = 61.76, *SD* = 10.95). For the second assessment, all groups received higher mediated scores. Group C also obtained a higher mediated score than other groups (*M* = 66.37, *SD* = 8.18), followed by Group B (*M*=65.21, *SD* = 10.07). The third assessment resulted in the higher mediated scores of all groups when compared with the first two time points. Group C had the highest mediated score (*M* = 71.47, *SD* = 8.01), whereas the lowest mediated score obtained by Group D (*M* = 64.15, *SD* = 12.10), which was slightly different from the second assessment.

Table 74 Descriptive statistics of mediated scores of computerized dynamic assessment for reading literacy

Group	Mediated score of C-DA 1 (Time 1)		Mediated score of C-DA 2 (Time 2)		Mediated score of C-DA 3 (Time 3)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Group A	61.76	10.95	65.15	10.57	68.23	9.54
Group B	61.52	11.00	65.21	10.07	68.44	10.52
Group C	62.82	9.35	66.37	8.18	71.47	8.01
Group D	58.98	12.10	62.94	12.81	64.15	12.10

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

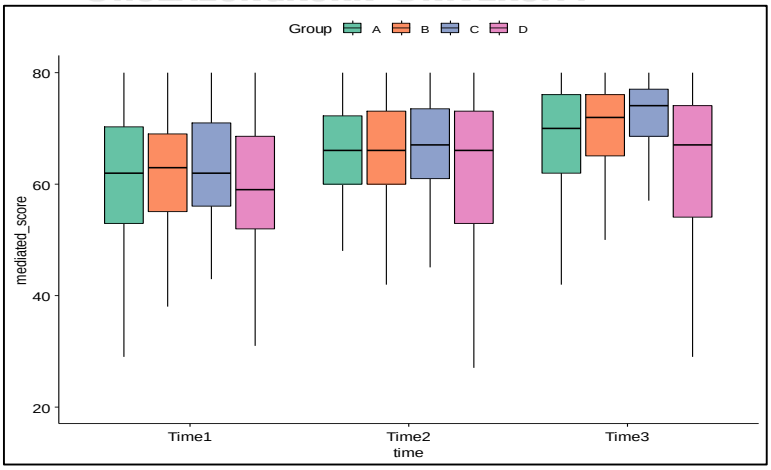


Figure 86 Box plots of mediated scores of different prompting groups over three time points

Two-way mixed ANOVA was used to evaluate the interaction effect between different prompting groups and time in explaining mediated scores. Normality assumption was checked by using QQ-plot. The Mauchly's test was used to assess the sphericity assumption of within-subject factor. Because Mauchly's test was significant and violated the sphericity assumption, the Greenhouse-Geisser sphericity correction was analyzed. As presented in Table 75 and Figure 87, it could be seen that there were no significant two-way interactions between group and time, $F(5.82, 1041.73) = 1.44$, $p = .197$. However, there were statistically significant main effects of group ($F(3, 537)=7.94$, $p < .05$) and time ($F(1.94, 1041.73) = 90.89$, $p < .05$) on the mediated score.

Table 75 Two-way mixed-effects ANOVA of mediated scores

Source	Univariate ANCOVA				
	SS	Num DF	Den DF	F	*p*
Group	4605.79	3	537	7.94*	.000
time	12468.69	1.94	1041.73	90.89*	.000
Group*time	593.27	5.82	1041.73	1.44	.197

*$p < .05$

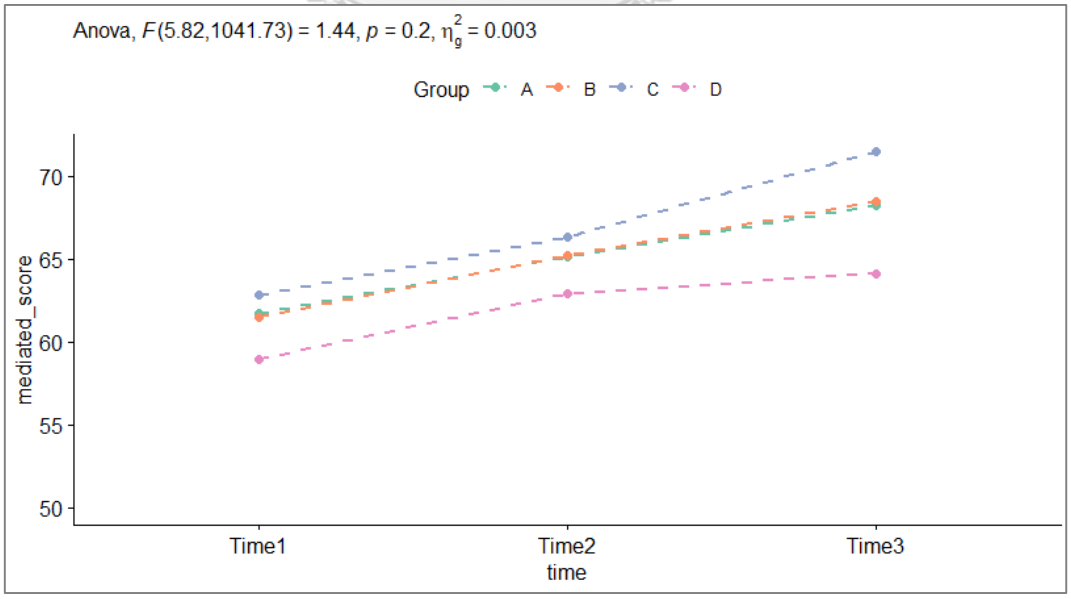


Figure 87 Line graphs of mediated scores of different prompting groups over three time points

Significant main effects

Pairwise comparison is further analyzed for the group effect (between-subjects factor) and time effect (within-subject factor). The results in Table 76 and Figure 88 revealed group effect on mediated scores was statistically significant. Group D had significantly lower mediated scores when compared with other groups. Moreover, in Table 77 and Figure 89, the time effect on mediated scores was statistically significant. Students obtained significantly higher mediated scores in each time point.

Table 76 Main effect result of pairwise comparison of random-effect (group)

Pairwise comparison		Mean Difference (I - J)	Std. Error	*p*
Group1 (I)	Group 2 (J)			
A	B	-.010	.948	1.000
A	C	-1.840	.963	.339
A	D	3.021*	.980	.013
B	C	-1.830	.978	.370
B	D	3.031*	.994	.014
C	D	4.861*	1.008	.000

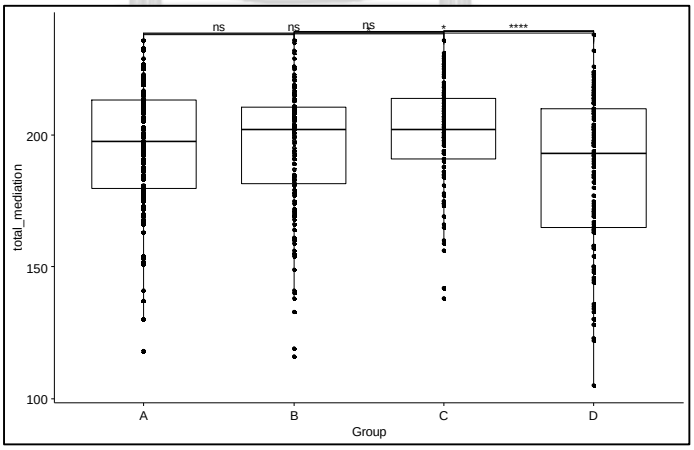


Figure 88 Box plots of group effect

Table 77 Main effect result of pairwise comparison of fixed-effect (time)

Pairwise comparison		Mean Difference (I - J)	Std. Error	*p*
Group 1 (I)	Group 2 (J)			
Time 1	Time 2	-3.646*	.482	.000
Time 1	Time 3	-6.799*	.547	.000
Time 2	Time 3	-3.153*	.482	.000

*$p < .05$

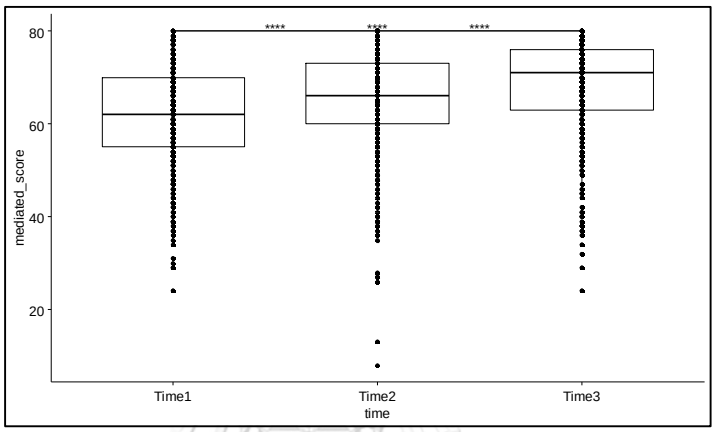


Figure 89 Box plots of time effect

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.3.3 Results of one-way MANOVA of different prompting groups on levels of promptings

2.3.3.1 Results of one-way MANOVA of different prompting groups on levels of promptings of overall reading literacy

The data in Table 78 and Figure 90 indicates that regarding the first promptings, Group D received the highest first prompting across three time points (*M* = 9.43, 7.95, 7.72, respectively). Moreover, Group A obtained the lowest first prompting during the initial assessment (*M* = 8.26), but had higher first prompting (*M* = 7.26, 5.95, respectively) when compared with Group B and Group C during the last two time points. On the contrary, Group C obtained the lowest first prompting for the second and third assessment (*M* = 6.78, 5.08, respectively).

Concerning the second prompting, Group D obtained the highest second mediation for the three time points (*M* = 5.95, 4.62, 4.13, respectively), whereas the lowest second prompting obtained by Group C over time (*M* = 4.73, 3.72, 1.92, respectively). Group A and Group B showed a little difference in receiving the second prompting.

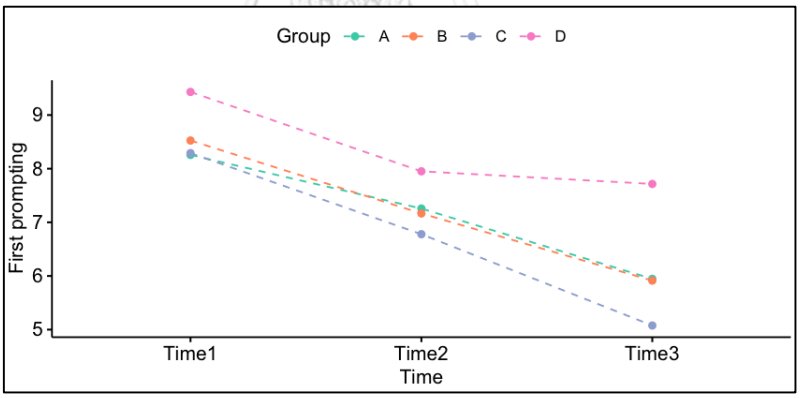
In terms of the third prompting, Group C received the lowest third prompting across three time points (*M* = 2.70, 2.05, 1.00, respectively). Group A and Group B relatively received similar third prompting during the initial assessment (*M* = 3.05, 3.04, respectively), whereas Group B obtained lower for the third time points (*M* = 1.70). Group D had the highest third prompting over time (*M* = 3.68, 2.85, 2.54, respectively) and their third prompting at the second and the third assessment had little difference.

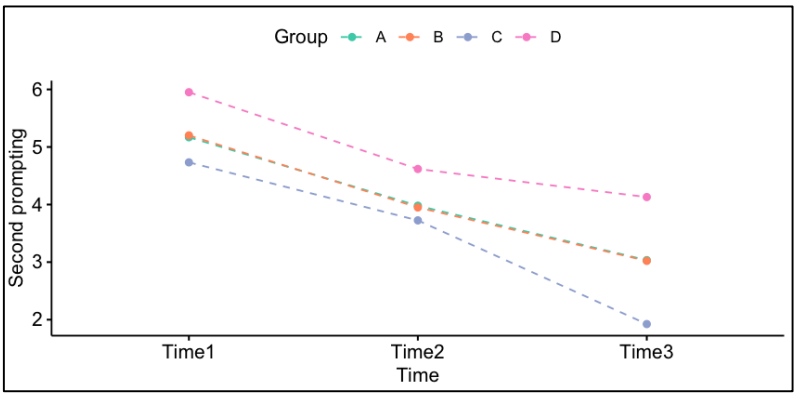
Given that the fourth prompting was the correct answer given to the test-takers, Group C obtained the lowest fourth prompting across three time points (*M* = 1.53, 1.08, 0.53, respectively), followed by Group B (*M* = 1.72, 1.29, 0.92, respectively). The highest fourth prompting was received by Group D over three time points (*M* = 1.96, 1.63, 1.47, respectively).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 78 Means of levels of promptings across three time points

Time	Group	First level		Second level		Third level		Fourth level	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
C-DA1	Group A	8.26	4.01	5.17	3.28	3.05	2.54	1.76	1.78
(Time 1)	Group B	8.52	3.94	5.20	3.42	3.04	2.64	1.72	1.87
	Group C	8.29	3.89	4.73	2.82	2.70	2.16	1.53	1.45
	Group D	9.43	4.25	5.95	3.70	3.68	2.94	1.96	1.94
C-DA 2	Group A	7.26	3.81	3.98	3.14	2.30	2.49	1.32	2.00
(Time 2)	Group B	7.16	3.91	3.95	3.13	2.39	2.29	1.29	1.40
	Group C	6.78	3.37	3.72	2.51	2.05	1.85	1.08	1.33
	Group D	7.95	4.73	4.62	3.80	2.85	2.94	1.63	2.07
C-DA3	Group A	5.95	4.08	3.03	2.89	1.82	2.05	0.97	1.38
(Time 3)	Group B	5.91	4.11	3.02	3.28	1.70	2.34	0.92	1.52
	Group C	5.08	3.65	1.92	2.35	1.00	1.64	0.53	1.04
	Group D	7.72	4.74	4.13	3.61	2.54	2.61	1.47	2.02





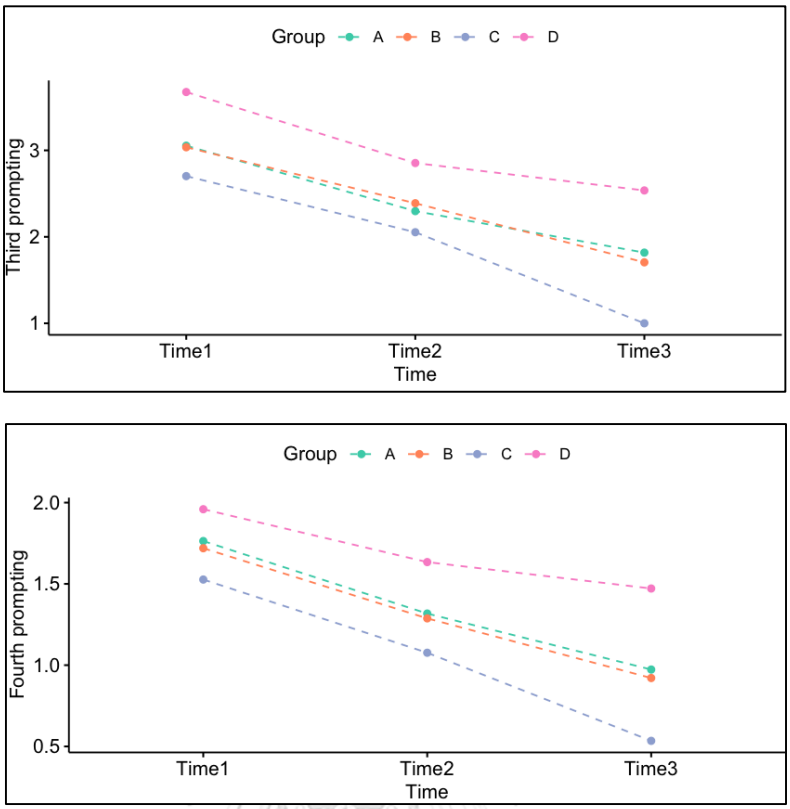


Figure 90 Line graphs of levels of prompting of different groups over three time points

One-way MANOVA was performed on four levels of prompting (first, second, third, and fourth promptings) of reading literacy. The independent variable was groups received different types of prompting, including instructional prompting (Group A), error-explanation prompting (Group B), mixed prompting (Group C), verification prompting (Group D). Because assumption of homogeneity of variance-covariance matrices was unsatisfactory, Pillai's multivariate statistic was used in this study. The multivariate group effect was significant for levels of prompting, Pillai's trace = .055, $F(12,1608) = 2.48$, $p < .05$. The results indicated a significant group difference of three levels of prompting as shown in Table 79.

In Table 80, the results of univariate tests further indicated that there were significant differences among different prompting groups in terms of first prompting, $F(3,537) = 6.740$, $p < .05$; second prompting, $F(3,537) = 7.730$, $p < .05$; third prompting, $F(3, 537) = 8.326$, $p < .05$; and fourth prompting, $F(3,537) = 6.074$, $p < .05$.

Based on Table 81, the results of follow-up pairwise comparisons using a Games-Howell post-hoc test revealed that Group D received higher levels of prompting when compared with other groups. For the first level of prompting, Group D had significantly higher prompts than other three prompting-based groups. Concerning the second level of prompting, Group D had significantly higher prompts than Group C. For the third levels of prompting, Group D received significantly higher prompts than Group A and Group C. For the fourth levels of prompting, Group D had significantly higher prompt than Group C. Moreover, the result also showed that Group A significantly required higher third level of prompting than Group C.

Table 79 One-way MANOVA result of levels of prompting of reading literacy

	Value	*F*	*Hypothesis df*	*Error df*	*p*
Pillai's trace	.055	2.48*	12	1608	.003
Wilks' lambda	.946	2.50*	12	1413	.003
Hotelling's trace	.057	2.52*	12	1598	.003
Roy	.047	6.34*	4	536	.000

*$p < .05$

Follow-up analysis

Table 80 Analysis of variance result of levels of prompting of reading literacy

Source	DFn	DFd	Sum of Squares	Mean Square	*F*	*p*
First prompting	3	537	1695.91	565.305	6.740*	.000
Second prompting	3	537	1194.23	398.078	7.730*	.000
Third prompting	3	537	700.70	233.567	8.326*	.000
Fourth prompting	3	537	237.67	79.224	6.074*	.000

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 81 Multiple comparisons in levels of prompting of reading literacy

Dependent Variable	Pair comparison		Mean Difference (I - J)	Std. Error	*p*
	Group1 (I)	Group2 (J)			
First prompting	A	B	-.14	1.044	.999
	A	C	1.31	1.009	.562
	A	D	-3.64*	1.212	.016
	B	C	1.46	1.036	.496
	B	D	-3.49*	1.235	.026
	C	D	-4.95*	1.205	.000
Second prompting	A	B	.01	.835	1.000
	A	C	1.80	.743	.075
	A	D	-2.52	.978	.052
	B	C	1.79	.778	.100
	B	D	-2.53	1.005	.060
	C	D	-4.32*	.930	.000
Third prompting	A	B	.04	.619	1.000
	A	C	1.41*	.539	.046
	A	D	-1.90*	.717	.043
	B	C	1.37	.581	.087
	B	D	-1.94	.749	.050
	C	D	-3.31*	.685	.000
Fourth prompting	A	B	.13	.418	.990
	A	C	.92	.384	.081
	A	D	-1.01	.505	.191
	B	C	.79	.374	.151
	B	D	-1.14	.498	.105
	C	D	-1.93*	.469	.000

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.3.3.2 Results of one-way MANOVA of different prompting groups on levels of promptings of locate information subscale

The results in Table 82 showed that Group D received the highest level of first prompting on locate information subscale (*M* = 6.28), followed by Group B (*M* = 5.38). Moreover, concerning the second and third levels of prompting, Group D reached the higher level of prompting use when compared with other groups (*M* = 3.55, 2.22, respectively). For the fourth level of prompting, the lowest level of prompting obtained by Group C (*M* = .65), followed by Group A (*M* = .89).

Table 82 Means of levels of promptings in locate information subscale across three time points

Group	Locate information							
	First level		Second level		Third level		Fourth level	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Group A	4.98	3.00	2.64	2.27	1.48	1.57	.89	1.22
Group B	5.38	3.17	2.88	2.47	1.63	1.81	.95	1.19
Group C	4.63	2.59	2.15	1.70	1.29	1.61	.65	.96
Group D	6.28	3.63	3.55	2.95	2.22	2.26	1.13	1.42

One-way MANOVA was performed on four levels of prompting (first, second, third, and fourth promptings) of locate information subscale. The independent variable was groups received different types of prompting (instructional prompting (Group A), error-explanation prompting (Group B), mixed prompting (Group C), verification prompting (Group D). Pillai's multivariate statistic was preferred in this study because assumption of homogeneity of variance-covariance matrices was unsatisfactory. The results revealed the significant group effect of levels of prompting, Pillai's trace = .059, $F(12, 1608) = 2.70$, $p < .05$. The results indicated a significant group difference in four levels of prompting of locate information subscale as presented in Table 83.

In Table 84, the results of univariate tests further indicated that there were significant differences among different prompting groups in terms of first prompting,

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

$F(3,537) = 6.705$, $p < .05$; second prompting, $F(3,537) = 7.690$, $p < .05$; third prompting, $F(3,537) = 6.145$, $p < .05$; and fourth prompting of locate information subscale, $F(3,537) = 3.480$, $p < .05$.

As evident from Table 85, the results of follow-up pairwise comparisons using a Games-Howell post-hoc test revealed that Group D significantly received higher first, second, and third levels of prompting in locate information subscale than Group A and Group C. For the fourth prompting, Group D had significantly higher prompt than Group C. Moreover, Group B significantly obtained larger mean of second level of prompting in locate information subscale than Group C. Levels of promptings in locate information subscale were depicted in Figure 91.

Table 83 One-way MANOVA result of levels of promptings in locate information subscale

	Value	*F*	*Hypothesis df*	*Error df*	*p*
Pillai's trace	.059	2.70*	12	1608	.001
Wilks' lambda	.941	2.72*	12	1413	.001
Hotelling's trace	.061	2.73*	12	1598	.001
Roy	.046	6.15*	4	536	.000

* $p < .05$

Follow-up analysis

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 84 Analysis of Variance result of levels of promptings in locate information subscale

Locate information	DFn	DFd	Sum of Squares	Mean Square	*F*	*p*
First prompting	3	537	194.82	64.94	6.705*	.000
Second prompting	3	537	130.69	43.56	7.690*	.000
Third prompting	3	537	61.16	20.38	6.145*	.000
Fourth prompting	3	537	15.25	5.08	3.480*	.016

* $p < .05$

Table 85 Multiple Comparisons of levels of promptings in locate information subscale

Dependent Variable	Pair comparison		Mean Difference (I-J)	Std.Error	*p*
	Group 1 (I)	Group 2 (J)			
First prompting	A	B	-.40	.365	.690
	A	C	.35	.335	.717
	A	D	-1.30*	.410	.009
	B	C	.76	.352	.141
	B	D	-.90	.424	.146
	C	D	-1.66*	.399	.000
Second prompting	A	B	-.24	.281	.823
	A	C	.49	.239	.171
	A	D	-.92*	.325	.027
	B	C	.73*	.257	.024
	B	D	-.63	.339	.194
	C	D	-1.41*	.305	.000
Third prompting	A	B	-.15	.201	.871
	A	C	.19	.191	.753
	A	D	-.74*	.242	.013
	B	C	.34	.208	.355
	B	D	-.59	.256	.103
	C	D	-93*	.248	.001
Fourth prompting	A	B	-.07	.143	.958
	A	C	.24	.131	.275
	A	D	-.24	.163	.439
	B	C	.31	.131	.091
	B	D	-.71	.163	.714
	C	D	-.48*	.153	.010

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

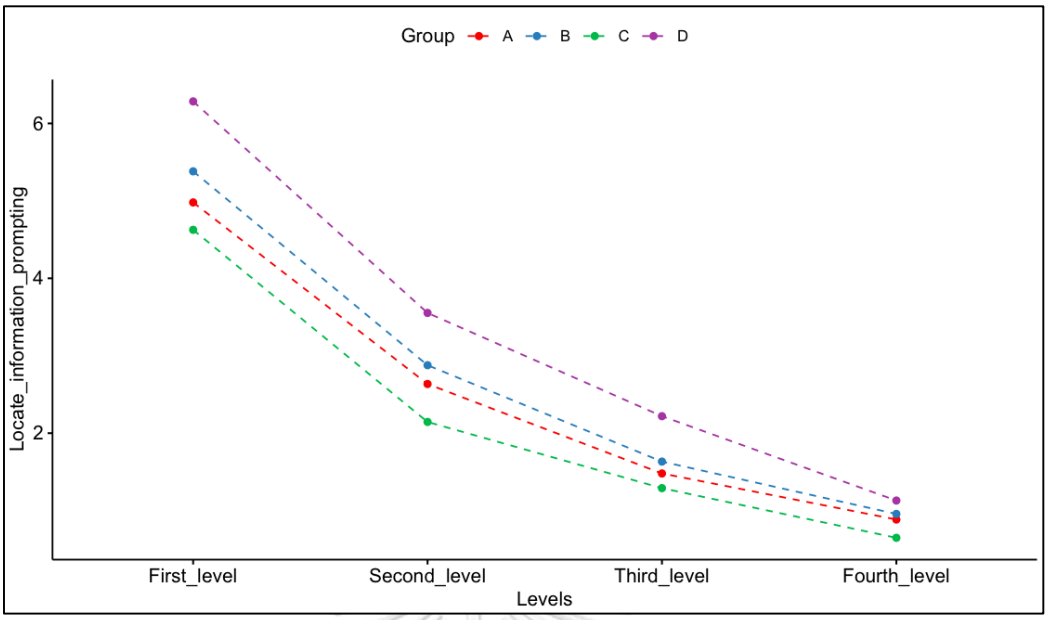


Figure 91 Line graphs of levels of promptings in locate information subscale

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.3.3.3 Results of one-way MANOVA of different prompting groups on levels of promptings of understand subscale

The results in Table 86 showed that Group D received the highest level of first prompting in understand subscale (*M* = 9.72), followed by Group A (*M* = 8.39). Moreover, concerning the second and third levels of prompting, Group C obtained the lowest use of prompting (*M* = 4.08, 2.18, respectively), whereas Group D reached the highest level of prompting use (*M*=5.65, 3.37, respectively). Also, Group B received a smaller level of prompting than Group A (*M*=4.43, 2.74, respectively). For the fourth level of prompting, the lowest level of prompting obtained by Group C (*M* = 1.20), followed by Group B (*M* = 1.45).

Table 86 Means of levels of promptings in understand subscale across three time points

Group	Understand							
	First level		Second level		Third level		Fourth level	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Group A	8.39	3.92	4.80	3.12	2.90	2.37	1.70	1.77
Group B	8.09	3.89	4.43	3.02	2.74	2.34	1.45	1.56
Group C	8.18	3.98	4.08	2.82	2.18	1.93	1.20	1.44
Group D	9.72	4.54	5.65	3.75	3.37	2.69	1.86	1.83

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

One-way MANOVA was performed on four levels of prompting (first, second, third, and fourth promptings) of understand subscale. The independent variable was groups receiving different types of prompting as follows: 1) instructional prompting (Group A), 2) error-explanation prompting (Group B), 3) mixed prompting (Group C), and 4) verification prompting (Group D). Pillai's multivariate statistic was preferred in this study because assumption of homogeneity of variance-covariance matrices was unsatisfactory. The results revealed the significant group effect of levels of prompting, Pillai's trace = .063, $F(12, 1608) = 2.89$, $p < .05$. The results indicated significant group differences in four levels of prompting in understand subscale as presented in Table 87.

In Table 88, the results of univariate tests further indicated that there were significant differences among different prompting groups in terms of first prompting, $F(3,537)=4.413$, $p < .05$; second prompting, $F(3,537)=5.683$, $p < .05$; third prompting, $F(3,537)=5.646$, $p < .05$; and fourth prompting of understand subscale, $F(3,537)=3.934$, $p < .05$.

As seen in Table 89, the results of follow-up pairwise comparisons using a Games-Howell post-hoc test revealed that Group D significantly got higher levels of first and second levels of promptings in understand subscale when compared with Group B and Group C. Furthermore, Group D significantly obtained larger mean of third and fourth levels of prompting than Group C. Moreover, Group A significantly obtained larger mean of third level of prompting in understand subscale than Group C. Levels of promptings in understand subscale as seen in Figure 92.

Table 87 One-way MANOVA result of levels of promptings in understand subscale

	Value	*F*	*Hypothesis df*	*Error df*	*p*
Pillai's trace	.063	2.89*	12	1608	.001
Wilks' lambda	.938	2.90*	12	1413	.001
Hotelling's trace	.065	2.89*	12	1598	.001
Roy	.035	4.69*	4	536	.001

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Follow-up analysis

Table 88 Analysis of Variance result of levels of promptings in understand subscale

Understand	DFn	DFd	Sum of Squares	Mean Square	F	*p*
First prompting	3	537	220.482	73.494	4.413*	.004
Second prompting	3	537	173.187	57.729	5.683*	.001
Third prompting	3	537	93.483	31.161	5.646*	.001
Fourth prompting	3	537	32.596	10.866	3.934*	.009

*$p < .05$

Table 89 Multiple Comparisons of levels of promptings in understand subscale

Dependent Variable	Pair comparison		Mean Difference (I-J)	Std.Error	*p*
	Group 1 (I)	Group 2 (J)			
First prompting	A	B	.31	.482	.911
	A	C	.21	.475	.972
	A	D	-1.33	.522	.055
	B	C	-.10	.479	.997
	B	D	-1.64*	.526	.011
	C	D	-1.54*	538	.023
Second prompting	A	B	.37	.363	.745
	A	C	.72	.356	.181
	A	D	-.85	.425	.188
	B	C	.36	.356	.751
	B	D	-1.22*	.425	.023
	C	D	-1.57*	.419	.001
Third prompting	A	B	.16	.279	.942
	A	C	.72*	.258	.028
	A	D	-.48	.312	.425
	B	C	.57	.261	.135
	B	D	-.63	.314	.185
	C	D	-1.20*	.296	.000
Fourth prompting	A	B	.24	.197	.607
	A	C	.50	.193	.051
	A	D	-.17	.221	.876
	B	C	.25	.183	.505
	B	D	-.41	.212	.220
	C	D	-.66*	.208	.009

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

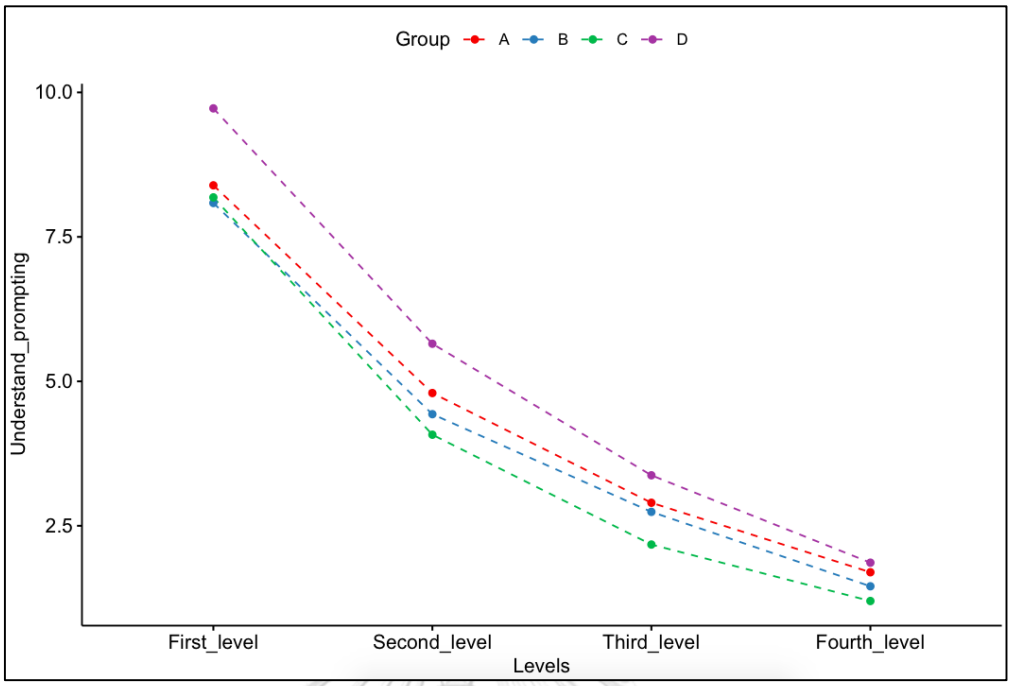


Figure 92 Line graphs of levels of prompting in understand subscale

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.3.3.4 Results of one-way MANOVA of different prompting groups on levels of promptings of evaluate and reflect subscale

The results in Table 90 showed that Group D received the highest level of first prompting in evaluate and reflect subscale (*M* = 9.09), followed by Group B (*M* = 8.14). Moreover, concerning the second and third levels of prompting, Group C obtained the lowest use of prompting (*M* = 4.16, 2.29, respectively), whereas Group D reached the highest level of prompting use (*M* = 5.50, 3.47, respectively). However, Group B received a larger level of second prompting but a smaller level of third prompting than Group A (*M* = 4.86, 2.76, respectively). For the fourth level of prompting, the lowest level of prompting obtained by Group C (*M* = 1.29), followed by Group A (*M* = 1.47).

Table 90 Means of levels of promptings in evaluate and reflect subscale across three time points

Group	Evaluate and reflect							
	First level		Second level		Third level		Fourth level	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Group A	8.09	3.01	4.75	2.65	2.79	2.07	1.47	1.60
Group B	8.14	3.12	4.86	2.77	2.76	2.19	1.52	1.52
Group C	7.34	2.98	4.16	2.41	2.29	1.86	1.29	1.23
Group D	9.09	3.55	5.50	3.20	3.47	2.66	2.07	2.08

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

One-way MANOVA was performed on four levels of prompting (first, second, third, and fourth promptings) of the evaluate and reflect subscale. The independent variable was groups receiving different types of prompting, including 1) instructional prompting (Group A), 2) error-explanation prompting (Group B), 3) mixed prompting (Group C), and 4) verification prompting (Group D). The multivariate statistic was calculated using Pillai's trace. The results revealed that there were significant group effects of levels of prompting of evaluate and reflect subscale, Pillai's trace = .051,

F(12, 1608) = 2.33, *p* < .05. The results indicated significant group differences in four levels of prompting of understand subscale as presented in Table 91.

In Table 92, the results of univariate tests further indicated that there were significant differences among different prompting groups in terms of first prompting, *F*(3,537) = 6.511, *p* < .05; second prompting, *F*(3,537) = 4.975, *p* < .05; third prompting, *F*(3,537) = 6.133, *p* < .05; and fourth prompting of evaluate and reflect subscale, *F*(3,537) = 5.421, *p* < .05.

The results of follow-up pairwise comparisons using a Games-Howell post-hoc test revealed that for all levels of promptings, Group D significantly obtained greater levels of prompting than Group C in evaluate and reflect subscale. Moreover, Group D received significantly higher level of fourth prompting than Group A. Levels of promptings in evaluate and reflect subscale as depicted in Table 93 and Figure 93.

Table 91 One-way MANOVA result of levels of promptings in evaluate and reflect subscale

	Value	*F*	*Hypothesis df*	*Error df*	*p*
Pillai's trace	.051	2.33*	12	1608	.006
Wilks' lambda	.949	2.35*	12	1413	.006
Hotelling's trace	.053	2.36*	12	1598	.005
Roy	.045	6.07*	4	536	.000

*$p < .05$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Follow-up analysis

Table 92 Analysis of Variance result of levels of promptings in evaluate and reflect subscale

Evaluate and reflect	DFn	DFd	Sum of Squares	Mean Square	*F*	*p*
First prompting	3	537	195.965	65.322	6.511*	.000
Second prompting	3	537	114.087	38.029	4.975*	.002
Third prompting	3	537	89.698	29.899	6.133*	.000
Fourth prompting	3	537	43.151	14.384	5.421*	.001

*$p < .05$

Table 93 Multiple Comparisons of levels of promptings in evaluate and reflect subscale

Dependent Variable	Pair comparison		Mean Difference (I-J)	Std.Error	*p*
	Group 1 (I)	Group 2 (J)			
First prompting	A	B	-.05	.363	.999
	A	C	.75	.360	.160
	A	D	-1.00	.405	.067
	B	C	.80	.372	.139
	B	D	-.95	.416	.103
	C	D	-1.75*	.413	.000
Second prompting	A	B	-.11	.321	.985
	A	C	.59	.303	.212
	A	D	-.75	.362	.169
	B	C	.70	.316	.119
	B	D	-.63	.373	.327
	C	D	-1.34*	.358	.001
Third prompting	A	B	.04	.253	.999
	A	C	.50	.236	.149
	A	D	-.68	.295	.099
	B	C	.47	.247	.239
	B	D	-.72	.304	.089
	C	D	-1.18*	.290	.000
Fourth prompting	A	B	-.50	.184	.995
	A	C	.18	.170	.705
	A	D	-.60*	.230	.047
	B	C	.23	.168	.527
	B	D	-.56	.228	.078
	C	D	-.78*	.217	.002

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

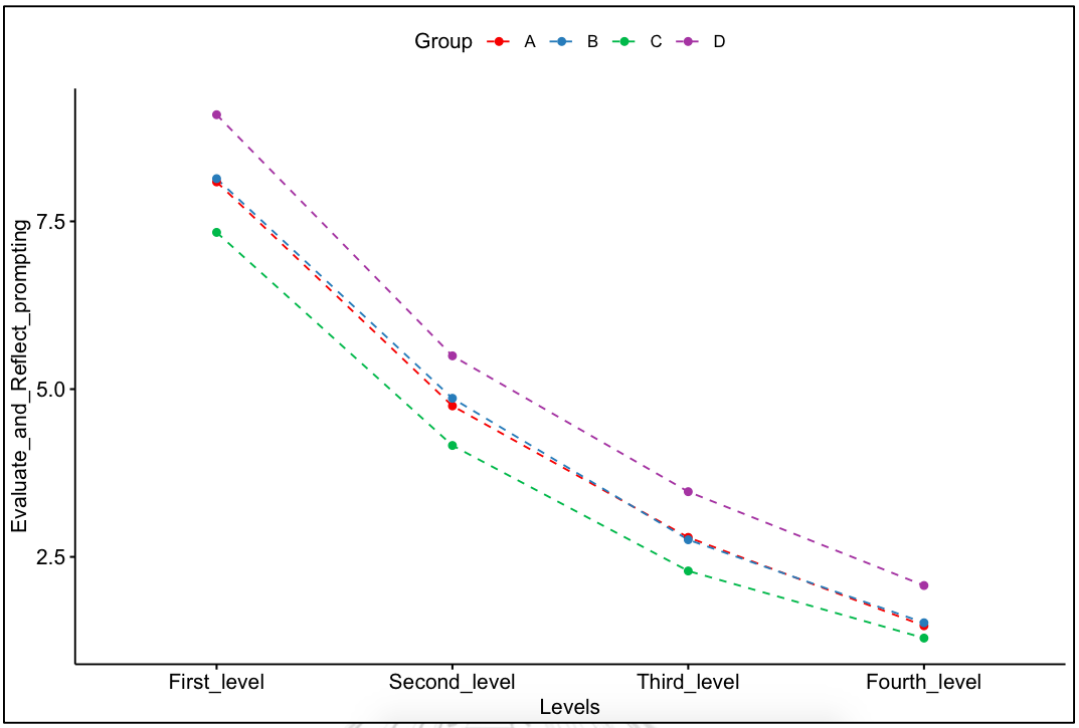


Figure 93 Line graphs of levels of prompting in evaluate and reflect subscale

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER FIVE
DISCUSSION AND CONCLUSION

This chapter is divided into four parts. The first part focuses on the summary of the findings. The second part shows the discussion of the findings. The third part is the conclusion of the research and the final part talks about the implication and recommendation for further studies.

Summary of the Study

The present study aimed at constructing and validating the computerized dynamic assessment for reading literacy and investigating the effects of computerized dynamic assessment on current performance of reading literacy, growth of learning, and learning potential of Grade 9 students.

For the initial phase of the study, reading literacy instruments for computerized dynamic assessment were constructed. The instrument differed from other computer-based tests in accordance with the application of two-tier test and the mediation procedures. All items were two-tier multiple choice questions. The first-tier items assessed student's measured constructs in reading literacy and the second-tier items measured their reasonings to avoid guessing. Each test form consisted of 20 items, along with four reading passages. The tested contents comprised three subscales as follows: 1) locate information, 2) understand, and 3) evaluate and reflect. Regarding mediation procedures, promptings were developed for each item. There were four types of promptings, including 1) instructional prompting, 2) error-explanation prompting, 3) mixed prompting, and 4) verificantion prompting. During this phase, content validity and test parallelism of the instruments were checked. Only a few items needed to be revised in accordance with expert's judgment of content and parallelism.

Concerning the second stage of the study, the data was gathered from pilot groups. For the first group, there were a total of 277 ninth grade students administered pretest and posttest. The second group was gathered from 525 ninth grade students in six secondary schools in Bangkok. The selection of participants was done through multi-stage random sampling. They were administered reading literacy instruments for computerized dynamic assessment. The data was analyzed in terms of psychometrics properties and parallelism of the instruments, including model fit, item parameter estimates, construct validity, and reliability.

Regarding the third stage of the study, computerized dynamic assessment (C-DA) for reading literacy that integrated promptings was designed and developed by using the theoretical basis of graduated prompting approach. There were 541 ninth graders from eleven secondary schools participated in this study and each individual was randomly assigned into the instructional-based group (n=148), the error-explanation group (n=139), the mixed group (n=131), and the verification group (n=123). A quasi-experimental design was adopted. They were administered C-DA across three time points. The statistical analysis was conducted to examine ninth grade students' reading literacy ability, reading literacy growth, and learning potential in reading literacy as follows:

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

(1) In terms of examining the students' reading literacy ability, one-way ANCOVA was used to investigate the effects of different prompting conditions of computerized dynamic assessment (C-DA) on reading literacy posttest score when the pretest score was served as a covariate. Moreover, one-way ANCOVA was used to examine the effects of different prompting conditions of computerized dynamic assessment (C-DA) on each reading literacy posttest subscale score when pretest subscale score was served as a covariate. Furthermore, the researcher investigated the effect of different prompting conditions of computerized dynamic assessment (C-DA) and other external variables, including gender and levels of learning

achievement, on reading literacy posttest score when the pretest score was treated as the covariate using two-way ANCOVA.

(2) With regard to reading literacy growth, univariate multiple group latent growth model was used to investigate the developmental trajectories of reading literacy and reading literacy subscales. The slope was assumed to have linear growth. Moreover, multivariate multiple group latent growth model was used to examine the relationships among changes with respect to reading literacy subscales. The model examined bidirectional associations between intercepts and slopes of comparative reading literacy subscales as follows: 1) locate information intercept and understand intercept, 2) locate information intercept and evaluate and reflect intercept, 3) understand intercept and evaluate and reflect intercept, 4) locate information slope and understand slope, 5) locate information slope and evaluate and reflect slope, 6) understand slope and evaluate and reflect slope, 7) locate information intercept and understand slope, 8) locate information intercept and evaluate and reflect slope, and 9) understand intercept and evaluate and reflect slope.

(3) In terms of learning potential in reading literacy, this research analyzed: 1) availability scores as the distance between the current ability and estimated capacity, 2) mediated scores as student's responsiveness to mediation, and 3) levels of prompting as the sums of levels of assistance needed. The researcher analyzed the differences in availability scores over three time points using one-way MANOVA. The independent variable was groups received different promptings, and the dependent variables were the availability score measured at the first time, the availability score measured at the second time, and the availability score measured at the third time. Moreover, the researcher analyzed the differences in mediated score using two-way mixed ANOVA. The dependent variables were prompting conditions and time, and the independent variable was mediated score. Additionally, levels of promptings of reading literacy as well as levels of promptings of reading literacy subscales were investigated using one-way MANOVA. The independent variable was prompting

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

condition, and the dependent variables were four levels of promptings (first, second, third, and fourth promptings).

The findings of this study consisted of two main topics with respect to the objectives of this study. It could be summarized as follows:

1. Development and validation of computerized dynamic assessment for reading literacy

1.1 Development of computerized dynamic assessment for reading literacy

This research developed computerized dynamic assessment (C-DA) for reading literacy that integrated the graduated promptings with assessment. The tested contents were composed of three dimensions of reading literacy, including 1) locate information, 2) understand, and 3) evaluate and reflect. There were twenty two-tier items combined with four reading passages in each testing session. The first-tier item was the general question to measure student's reading literacy and the second-tier item was used to avoid guessing. The tests were constructed as three parallel tests to measure in three different time points. The computerized dynamic assessment was interactive and online. It consisted of 1) user profile interface, 2) dynamic assessment interface, and 3) scoring report interface. If a student answered the first-tier question incorrectly, he or she would be provided the prompting patterns based on their group conditions from implicit to explicit. The prompting guidelines of each question were based on four types of promptings, including 1) instructional prompting, 2) error-explanation prompting, 3) fixed prompting, and 4) verification prompting.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2 Validation of computerized dynamic assessment for reading literacy

Psychometric properties were investigated for computerized dynamic assessment for reading literacy. In terms of item parameter estimates of the instruments, five measurement models were fitted and compared as follows: 1) the one-parameter logistic model, 2) the two-parameter logistic model, 3) the multidimensional item response theory model, 4) the testlet model, and 5) the bifactor model. The results indicated that the bifactor model was selected as the most fitted model to the data. The general factor of the bifactor model was assumed to represent the reading literacy demanded by all reading test items and each of the specific factors was assumed to contain one of the particular reading tasks. For the three reading literacy instruments, most items in each testlet depicted the larger discrimination parameter related to the general factor than the passage factors. Moreover, the parameter estimates of three reading literacy forms showed that a majority of the first-tier items had good multidimensional discrimination and most items produced difficulty estimates in the acceptable range. For the second-tier items, most items were good discrimination and good difficulty estimates. In terms of construct validity, confirmatory factor analysis provided evidence that the three measured constructs of the reading literacy fitted to the data. The reliability estimates for reading literacy instruments were acceptable and had high internal consistency. Regarding the statistical parallelism of the three reading literacy forms, they showed satisfactory closeness between the two assembled forms at the test and item-by-item levels.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2. Effects of different prompting conditions of computerized dynamic assessment for reading literacy on reading literacy ability, growth of learning, and learning potential of Grade 9 students

2.1 Effects of different prompting conditions of computerized dynamic assessment on reading literacy ability of Grade 9 students

(1) The results showed that different prompting conditions of computerized dynamic assessment for reading literacy had a significant effect on students' reading posttest score. Group D scored significantly lower when compared with students in any other groups.

(2) The findings revealed that different prompting conditions of computerized dynamic assessment for reading literacy had a significant effect on locate information subscale score. Group D received significantly lower locate information score when compared to other groups. Moreover, there were significant differences of different prompting conditions in understand subscale score. Group D was significantly different from other groups. In addition, significant differences were found in evaluate and reflect subscale score. Group D received significantly lower evaluate and reflect subscale score when compared to students in other groups.

(3) There was no significant interaction between prompting conditions and gender. In terms of main effects, there were significant differences in different prompting-based groups of computerized dynamic assessment on posttest score, whereas no significant differences were found between male and female.

(4) There was no significant interaction between prompting conditions and levels of learning achievement. For the main effects, there were significant differences in different prompting-based groups on posttest score, whereas no significant differences were found among students with different levels of learning achievement.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.2 Effects of different prompting conditions of computerized dynamic assessment on growth trajectory of Grade 9 students

2.2.1 Univariate latent growth curve models of changes in overall reading literacy and reading literacy subscales

(1) With regard to the overall reading literacy, all four groups had positive change in overall reading literacy over time. Group C obtained the highest score at the initial stage and the trend of their score was higher than other groups. Individual differences were found in the initial values and their rate of change for all groups. Furthermore, those students with lower starting scores performed better than did those with higher starting scores over time.

(2) Regarding the growth trajectory of locate information subscale, all four groups had positive change in locate information subscale over time. Group C received the highest score at the initial time point and their score increased over time was higher than other groups. Moreover, individual differences were found in the initial values and their rate of change. Those students with lower starting scores exhibited linear growth than did those with higher starting scores.

(3) In terms of the growth trajectory of understand subscale, Group B obtained the highest score at the initial stage. However, Group C obtained the highest growth trajectory. Group D had no change in score over time. Individual differences were found in the initial values and their rate of change for all groups, except Group A that had no differences in rate of change. Those students with lower starting scores exhibited linear growth than did those with higher starting scores.

(4) For the growth trajectory of evaluate and reflect subscale, all four groups had positive change in evaluate and reflect subscale over time. Group C obtained the highest score at the initial stage. However, Group B had the highest increased score over time. Individual differences were found in the initial values for all groups, whereas only Group D had individual differences in the rate of change.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.2.2 Multivariate latent growth curve model of the associations of change in reading literacy subscales

(1) For the associations of reading literacy subscales of total samples, the increased score in one subscale was associated with higher scores in other subscales at the initial stage. Moreover, higher growth rate in locate information was correlated with the increased growth rate in evaluate and reflect score.

(2) For the associations of reading literacy subscales of Group A, the increased score in one subscale was associated with higher scores in other subscales at the initial stage. In addition, increased understand score at the initial stage was correlated with increase evaluate and reflect score over three time points.

(3) For the associations of reading literacy subscales of Group B, the increased score in one subscale was associated with higher scores in other subscales at the initial stage. However, the growth rate in one subscale did not relate to the growth rate in other subscales.

(4) For the associations of reading literacy subscales of Group C, the increased score in one subscale was associated with higher scores in other subscales at the initial stage. Moreover, the growth rate in understand was associated with the growth rate in evaluate and reflect subscale.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

(5) For the associations of reading literacy subscales of Group D, Moreover, higher locate information at the initial stage was related to lower understand score over three time points.

2.3 Effects of different prompting conditions of computerized dynamic assessment on learning potential of Grade 9 students

2.3.1 Results of different prompting conditions of computerized dynamic assessment on availability scores at different time point

The result pointed out that no significant differences of availability scores measured at the first and second time points. However, there were significant differences of availability scores measured at the third time point among different prompting groups. Group D had a significantly higher availability score when compared with Group C.

2.3.2 Results of different prompting conditions of computerized dynamic assessment and time on mediated score

There were no significant two-way interaction between prompting conditions and time on mediated score. However, there was a significant main effect of prompting conditions (between-subject factor) on mediated score, and there was significant main effect of time (within-subject factor) on mediated score. The result found that Group D had significantly lower mediated scores when compared with other groups. Furthermore, mediated scores were obtained differently in each time point.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.3.3 Results of different prompting conditions of computerized dynamic assessment on levels of promptings

(1) Regarding the levels of promptings in overall reading literacy, the results showed that there were significant differences among prompting conditions in terms of first, second, third, and fourth promptings. The results found that Group D had a significantly higher level of prompting in overall reading literacy than other groups. Group D had a significantly higher first level of promptings than other prompting-based groups. They had a significantly larger second level of prompting than Group C and a greater third level of prompting than Group A and Group C. Also, they obtained a greater fourth level of prompting than Group C. Moreover, Group A

significantly obtained a larger third level of prompting in overall reading literacy than Group C.

(2) With respect to the levels of prompting in locate information subscale, the results pointed out that there were significant differences among prompting conditions in terms of first, second, third, and fourth promptings. For the first and second levels of prompting, Group D had significantly higher assistance in locate information when compared with Group A and Group C. Moreover, Group B significantly obtained larger second level of prompting than Group C. With respect to the third level of prompting, Group D had significant difference in acquiring third level of prompting than Group A and Group C. For the fourth level of prompting, Group D had significantly higher level of fourth prompting than Group C.

(3) Concerning the levels of prompting in understand subscale, the results revealed that there were significant differences among prompting conditions in terms of first, second, third, and fourth promptings. For the first and second levels of prompting, Group D obtained significantly higher assistance when compared with Group B and Group C. With respect to the third level of prompting, Group D required significantly higher assistance when compared with Group C. Moreover, Group A had larger promptings than Group C. Regarding the fourth level of prompting, Group D obtained higher assistance than Group C.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

(4) For the levels of prompting in evaluate and reflect subscale, the results revealed that there were significant differences among prompting conditions in terms of first, second, third, and fourth promptings. For the first, second, and third levels of prompting, Group D obtained significantly higher than Group C. For the fourth level of prompting, Group D received a significantly higher level of prompting than Group A and Group C.

Discussion

The findings are discussed into five topics including:

1. Bifactor modeling in validation for computerized dynamic assessment for reading literacy

1.1 Model fit and item parameter estimates of computerized dynamic assessment for reading literacy

Selecting the most appropriate model is necessary for the accurate assessment. Several measurement models were examined to determine which model was appropriate to the data. The results of the goodness of fit indices among the reading literacy instruments indicated that the best fitted model of all test sets was the bifactor MIRT model when compared to other unidimensional and multidimensional models. The present study corroborates the same findings with previous studies (DeMars, 2012; Dunn & McCray, 2020; Min & He, 2014; Ouyang et al., 2016; Rijmen, 2010; Toland et al., 2017) that the bifactor MIRT model was preferred rather than the traditional IRT models and testlet model in reading assessment. This might be due to the complexity and flexibility of the model structure (Markon, 2019). Despite the greater parameter estimates, a loading in the bifactor model is less constrained and more flexible than a loading in testlet model and 2PL IRT model (Markon, 2019). According to DeMars (2012), the model provided better understanding about the dimensionality of the tests. Moreover, the traditional IRT modeling might not be appropriate due to its dependency structure caused by testlet format. Sharing of a passage might lead to local item dependence in reading assessment, resulting in biased parameter estimates, overestimated reliability, and overestimated ability estimates (Eckes, 2014; Lee, 2004; Min & He, 2014).

With the parameter estimates from the bifactor model, most items had high discrimination parameters, showing that they are capable of differentiating high and low reading literacy. In contrast, few items were evaluated as weak items with the low values of discrimination parameter (i.e., Items 10, 57, and 59). Besides, most

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

items had item difficulty values between -2 and 2. The model produced a few item difficulty values higher than 2 (i.e., Items 10, 40, 44, 57, 59, and 60). Five out of six very difficult items measured evaluate and reflect subscale. It suggests that students had difficulty in answering those items in evaluate and reflect correctly. Moreover, the interesting result found that some items (i.e., Items 10, 57, and 59) did not discriminate well between low and high ability students and these items were the most difficult items. Item 10 in evaluate and reflect subscale addressed students' reasoning and understanding of detecting and handling conflict. Students were requested to read the additional text that contradicted the main reading text in task two, and then they were asked to figure out the conflicting issue represented in the additional text. In addition, item 57 in understand subscale assessed a student's ability to represent literal meaning within the text. Item 59 in evaluate and reflect subscale was a question related to asking students to figure out how the content connected to the author's point of view. This result is aligned with that of Sabbag and Zieffler (2015) that the most difficult items had very low discrimination values. One possible explanation as to why many difficult items having very low discrimination might be because of the result of students' misconception and the wording of the item. Those items measured student's understanding and reasoning to make connections between the author's point of view and the given questions. Students might have difficulty in explaining their reasoning (Sabbag & Zieffler, 2015). The wording of the question might lead students to different interpretations of what the question is asking (Clause et al., 1998; Eckes, 2014). It seems to be the case that minor changes in wording caused substantial differences in interpretation (Clause et al., 1998). Consequently, these items need to be revised or replaced in terms of content and grammatical structure.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2 Estimating the magnitude of specific dimensions of the reading literacy tests

In this study, reading passages in each test form were based on task, text, and level of complexity. With respect to level of complexity, transferability was adapted to select the reading passages. The levels of transfer comprised 1) near transfer, 2) far transfer, and 3) very far transfer. In each test form, the first passage was near transfer that related to the student's background knowledge that they were familiar with particular topics well. The second and the third reading passages were involved with far transfer that combined student's previous knowledge and new knowledge. The reading passages were general topics. The fourth reading passages were very far transfer that student hardly encountered in their daily lives but needed to know for becoming global citizen.

As opposed to the unidimensional IRT, the multidimensional bifactor model were partitioned into two common sources, general versus group factor sources (Reise, 2012). Apart from their reading literacy ability as the domain-general factor, their responses to the reading passages contribute to the variation in their performance (Min & He, 2014). The domain-general factor was assumed to represent reading literacy ability demanded by all reading literacy items, and four passage factors were assumed to contain the specific factors related to the testlets. In this study, several items in reading literacy instruments showed the higher magnitude of the discrimination parameters on the general dimension than those of the group dimensions. This suggests that the test measured what it intended to measure, which was reading literacy performance (Byun & Lee, 2016). This is consistent with Byun and Lee (2016) and Cai and Kunnan (2018) that slope parameters of the individual items of the general dimension were higher than those of the testlet dimension.

On the contrary, the result found that three out of twelve testlets had more items with the smaller magnitude of the discrimination parameter on the general dimension than on the group dimensions. The strong influence of the group effect

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

upon these items suggests that the test score variance might have been accounted for by the group factor more than the general factor (Min & He, 2014). That is to say, the major factor affecting a student's response is the influence of reading passages. Two reasons were given for these findings. Firstly, this might be because a student's background knowledge related to passage topics has an impact on students' understanding of a text (DeMars, 2006; Min & He, 2014). Items within more familiar passages might be easier, leading to local dependence (DeMars, 2006; Eckes, 2014). According to DeMars (2006), the empirical evidence shows that readers' familiarity on topic or content knowledge make association with the traits accounting for the testlet dimension of a test. Second, the input features such as format, content, or linguistic features might contribute to the testlet influences (Eckes, 2014). This is because when the passage was organized in a single continuous text, the testlet influence got stronger. It appears that students were more familiar with a single continuous text rather than the combined discrete texts for reading assessment. Thus, the bifactor model capturing the secondary dimension could result in more accurate estimates of a student's reading literacy ability. Min and He (2014) stress that caution should be taken for interpretation of the reading literacy score. In order to avoid the influence of testlet effect, some items, especially low discrimination items needed revision.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2. The replication of the parallel tests using TIF and TCC as the statistical indices

2.1 The conformity of the parallel test forms

Based on the conformity of each test set, the results yielded equivalent to one another in response to statistical indices of item response theory. Test characteristic curve (TCC) (Armstrong et al., 1998) and test information function (TIF) (van der Linden & Adema, 1998) have the potential to measure conformity in accordance with test and item pairs by using RMSD index. TCC reflects differences in difficulty, whereas TIF reflects differences between distributions across their

information functions. Accordingly, the TCC results in forms that are parallel in difficulty, but not in precision, whereas the TIF results in forms that are parallel in terms of precision (Ali & van Rijn, 2016).

The findings of the present study revealed that test forms produced similar findings for test parallelism. The results indicated that the TIF and TCC methods yielded acceptable RMSD values. The findings of statistical parallelism did not vary by test forms. Even though the TIF approach performed better than the TCC approach in terms of RMSD, the differences were not substantial. This study is coherent with the study of Armstrong et al. (1998) that there was a minimized difference in either TIF or TCC. On the other hand, the TIF seems to be more commonly chosen as the statistical target (Ali & van Rijn, 2016). Several studies have provided evidence that the TIF performed well to create parallel forms (Boekkooi-Timminga, 1990; Kantiwong & Pasiphol, 2018; Sanders & Verschoor, 1998; van der Linden & Adema, 1998; Veldkamp et al., 2012). This might be because TIF is considered as weakly parallelism that is easier to manipulate because it does not yield identical observed score distributions, requiring no scoring categories or the number of items to be the same (Lin, 2008; Samejima, 1977). Moreover, it is believed that the difficulty is implicit when the TIF of the tests are equal (Ariel et al., 2006). Consequently, the selection of statistical targets is a matter of operational convenience and the nature of test forms; there is no agreement with respect to selection of either difficulty or precision (Ali & van Rijn, 2016).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Nevertheless, some pairs of parallel tests could not attain the score equivalence. This might be explained by the non-conformity of the construct (Schmucker et al., 2019). The results showed that $RMSD_{TIF}$ and $RMSD_{TCC}$ values of the Test set 3 might not provide evidence to support parallelism. The reason might be because item difficulty estimates of Form C-DA3 had unusually high values, including Item 44 (b = 45.78), Item 58 (b =13.27), and Item 59 (b = 8.51), whereas the item difficulty estimates of Form C-DA2 ranged from 0.48-3.96. As a result, the deletion or

revision of those items could result in the parallelism of both forms in terms of TCC and TIF.

2.2 The conformity of the item-by-item parallelism

Similar to the results of the parallel of test forms, item parallelism is the case when items within test forms have equal parameters in an equal IRT model (Lin, 2008). The choice of item parallelism is based upon item information function (IF) and item characteristic curve (ICC). The results of item-by-item parallelism showed that the RMSD seemed to favor the item information function (IF) approach because it yielded higher parallelism than the item characteristic curve (ICC). However, both statistical indices interpreted the same meanings, showing that all items were parallel. The outcomes of both indices lead to a conclusion that the ICC performed as well as the IF approaches. Thus, this method has the potential to produce the alternate test forms that are statistically equivalent, without great loss of difficulty and precision.

3. Effectiveness of computerized dynamic assessment for reading literacy on reading literacy abilities of ninth graders

3.1 Effects of different prompting conditions of computerized dynamic assessment for reading literacy on student's reading literacy posttest

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

With respect to prompting conditions, different promptings relied upon the guided contents. This experiment included four types of promptings as follows: (1) instructional prompts that guided on the necessary strategies to solve the problem, (2) error-explanation prompts that explained why a specific response to item was not correct, (3) mixed prompts that combined the explanation of the reason why a specific response was not correct and provided necessary strategies to solve the problem, and (4) verification prompts or control condition provided whether the item was correct or incorrect. In accordance with Dorfler et al. (2017), content is assumed to be the most influential feature of prompting.

This study fully supported the hypothesis that treatment conditions received higher performance when compared with the control condition. It can be explained that promptings may provide students in treatment conditions with more strategies and contextual clues, whereas control condition provided verification prompts that only indicated whether the answer is correct or incorrect. This result is consistent with previous studies working on promptings of computer-based assessment in cognitive skills (Poehner & Lantolf, 2013; Resing & Elliott, 2011; Stevenson et al., 2016; Teo, 2012; Wang & Chen, 2016; Wang, 2011; Wu et al., 2017; Zhang & Lu, 2019; Zhang et al., 2017). Those studies reported that computerized dynamic assessment generally produced positive outcomes on the achievement of students. According to Zhang et al. (2017), computerized dynamic assessment may reduce cognitive difficulty. The graduated prompt approach helped students solve problems by providing repeated practice with more specific hints that can guide students to discover or apply some principles to solve problems and learn more (Ai, 2017; Poehner & Lantolf, 2013; Wang, 2011; Zhang et al., 2017). Resing and Elliott (2011) states that the graduated prompt approach enable students to evaluate their progress when solving cognitive tasks. Tzuriel and Shamir (2002) said that the treatment condition provided a more motivating and supportive environment than the control condition, leading to the improvement of cognitive capacities of students.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

In addition, the results showed that prompting methods of the treatment groups were not significantly different in terms of their posttest scores. When considering reading literacy subscales, the results showed similar findings that there were no significant differences in terms of posttest scores in each subscale among treatment groups. This suggests that a wide range of prompts of computerized dynamic assessment for reading literacy could be used for enhancing reading literacy. With the inclusion of graduate assistance, the graduated prompting techniques in computerized dynamic assessment for reading literacy could help students to construct their own learning in a graduated way.

3.2 Effects of prompting conditions and external variables on student's reading literacy posttest

The results revealed that there was no significant interaction between prompting conditions and gender. The main effect analysis indicated no statistically significant differences in the posttest scores between male and female students. Moreover, there was no significant interaction between prompting conditions and levels of learning achievement. The findings showed that no statistically significant differences in the posttest scores among students with different levels of learning achievement, but significant differences were found in different prompting conditions. This supports evidence that computerized dynamic assessment for reading literacy were applicable for most male and female students with a wide range of ability. Moreover, promptings of computerized dynamic assessment for reading literacy had influence on posttest scores. The finding is consistent with Wang, (2011) and Wang & Chen (2016) that different types of promptings of computer-based assessment were applicable for most students with a wide range of ability. The interesting findings also found that although the mean differences were not statistically significant, low achievers had higher posttest scores than high achievers. This indicates that the different types of promptings of computerized dynamic assessment might provide low achievers with more guidance and assistance in the supportive environment that would increase the level of reading ability. Low achievers might have more opportunity to learn and have more solutions to find the correct answer. Wang (2010) and Wang (2011) found that the web-based dynamic assessment could enable low-level learners to experience more effective learning and improve their learning compared with their counterparts. Wang & Chen (2016) also pointed out that the low-initial reading ability had higher gain than other groups due to the supportive environments that help them construct their learning. Lin (2016) said that scaffolding in dynamic assessment assists students with low prior knowledge to reduce cognitive load (Lin, 2016). Therefore, the

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

promptings of computerized dynamic assessment were suitable for students with a wide range of ability and may particularly help low achievers to improve their performance in reading literacy.

4. Effectiveness of computerized dynamic assessment for reading literacy on reading literacy growth of ninth graders

4.1 Effects of different prompting conditions of computerized dynamic assessment for reading literacy on growth trajectory in overall reading literacy and each reading literacy subscales

4.1.1 Growth trajectory in reading literacy

Multiple-group latent growth curve modeling was performed to examine the sample mean and individual differences in the initial level and trajectories of the overall reading literacy across ninth graders with different prompting conditions.

Our research extended the findings by investigating that latent growth curve can be used to differentiate the trajectory of reading literacy between groups with different promptings. The intercept and slope of overall reading literacy of ninth graders among different prompting conditions were compared. The model identified that all prompting conditions were significantly associated with overall reading literacy at the baseline level and rate of change in reading literacy over time. The growth rate of reading literacy for treatment groups were higher than those for the control group. It can be concluded that computerized dynamic assessment had a long term impact on reading literacy and all students benefited from computerized dynamic assessment for reading literacy. This study is coherent with Wang and Chen (2016) and Yang and Qian (2020). Wang and Chen (2016) said that their web-based dynamic assessment had significant effects on the change or growth in students' reading ability. Yang and Qian (2020) pointed out that although students received the increased difficulty level of the items after four weeks of learning, students in

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

computerized dynamic assessment obtained higher scores than the control group. They were more capable of integrating the mediation included in the intervention even when the context changed. However, this result is not consistent with that reported by Zhang et al. (2017) that the computerized dynamic assessment led to higher performance of the experimental group than the control group over time. This contradiction with Zhang et al. (2017) might be a result of the shorter period of time and the different types of statistical analyses. Moreover, the present findings revealed that the largest increase in overall reading literacy was mixed prompting. This suggests that students in mixed prompting benefitted the most from the promptings of the computerized dynamic assessment for reading literacy over time. This can be explained by the prompt given in mixed prompting group that helped students perceived their own reading difficulties and offerred strategies or hints that they could use to deliver a correct response and transfer their knowledge to the next question (Golke et al., 2015).

4.1.2 Growth trajectory in locate information subscale

The model identified that all prompting conditions were significantly associated with locate information subscale at the baseline level and rate of change over time. All prompting groups had a significant increase in locate information over time. The results showed that the larger increase in locate information subscale score was mixed prompting group. Error-explanation prompting group had high values in locate information score at the baseline, but they showed the lower rate of increase in locate information subscale score over time when compared with other groups. According to OECD (2019a), locate information draws upon readers' understanding of the task demand, text organization and the relevance of text. The reasons for the lower locate information subscale score in Error-explanation prompting group might be because the given prompt in this group did not provide sufficient support in the long-term growth. The error-explanation prompt might not focus on student's awareness of their information needs and

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

might not increase students' ability to scan through a text to retrieve specific pieces of information as provided by instructional prompt.

4.1.3 Growth trajectory in understand subscale

The model identified that all prompting conditions, except Verification prompting group were significantly associated with understand subscale at the baseline level and rate of change over time. The results indicated a significant increase in understand subscale in the sample of treatment groups, whereas no significant change was observed in Verification prompting or control group. This indicated that Verification prompting group not only received lower growth rate in understand score, but they also had no change in understand subscale over three assessment sessions. This might be because students in Verification prompting group experienced difficulties in conveying and making inferences in understand subscale. OECD (2019a) said that understand is based on the situation model proposed by Kintsch (1998) that a situation model consists of two core processes: the representation of literal meaning and the integration of literal text through mapping and inference processes. Both processes require students to comprehend sentences and identify the implicit main idea of the text. Thus, the verification prompt given in Verification prompting group might not be sufficient in forming an understanding of the meaning conveyed in the passage.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.1.4 Growth trajectory in evaluate and reflect subscale

The model identified that all prompting conditions were significantly associated with evaluate and reflect subscale at the baseline level and rate of change over time. All prompting groups had a significant increase in evaluate and reflect subscale over time. Error-explanation prompting group had the highest rate of increase in evaluate and reflect subscale over time. This may be due to the fact that prompting in Error-explanation prompting group explained why a specific response was incorrect. Students in Error-explanation prompting group benefitted from the prompt providing them the mistake explanation and they were less likely

to make the same mistake again in the subsequent sessions (Dorfler et al., 2017; Golke et al., 2015). OECD (2019a) pointed out that evaluate and reflect requires students to reflect on the content and form of the text and assess the quality and validity of information. Error-explanation prompting group may have the opportunity to reflect on their own experience or knowledge to compare and contrast the chosen response whether it is correct or incorrect that help them improve their awareness of the information provided within the text.

4.2 The association of reading literacy subscales

The multivariate analyses extended the univariate analyses of reading literacy and reading literacy subscales to examine the interrelatedness of locate information, understand, and evaluate and reflect subscales across three time points. The results of the multivariate model showed that locate information, understand, and evaluate and reflect subscales are associated during the initial stage of assessment for all prompting groups. Higher initial level of locate information was categorized by higher understand and evaluate and reflect subscales. It indicated that students with higher scores in one subscale had higher scores on other subscales. That might be supported by the framework of PISA reading literacy that reading as a cognitive skill involves specific reading processes that competent readers use when engaging with the texts to achieve the goals. A wide range of reading processes that acknowledge the goal-driven nature of reading literacy comprise 1) ability to locate information, 2) ability to understand, and 3) ability to evaluate and reflect (OECD, 2019a). The constructs involved having basic and complex reading processes. The higher basic reading processes work, the better high-level reading functions process (OECD, 2019a).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

The results also pointed out the differences in interrelatedness of reading literacy subscales among different prompting conditions that lead to their performance in reading literacy over time. Currently, no longitudinal study has

identified the association of changes in reading literacy subscales by different prompting groups.

Regarding the multivariate model results in Instructional prompting group, the rate of increase in one subscale was not related to the rate of increase in other subscale over time. On the other hand, the understand subscale at the initial stage was correlated with the rate of increase in evaluate and reflect subscale over three testing sessions. It indicates that students who started with higher understand had steeper growth trajectories for evaluate and reflect subscale. This cross-domain interrelation identified that early understand predicts evaluate and reflect subscale in Instructional prompting group. Content of prompt might play a role in driving development. The reasons might be because 1) the strategies given in the instructional prompt in understand subscale might be useful to items in evaluate and reflect subscales and 2) both understand and evaluate and reflect subscale required a higher level of reading processes beyond searching and selecting relevant information. According to OECD (2019a), understand requires students to represent the explicit meaning of texts and integrate information and generate inferences, and evaluate and reflect requires students to assess the quality and credibility of information, reflect on content and form, and detect and handle conflict within the texts.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

With respect to the multivariate model result in Error-explanation prompting group, the interesting findings is that the rate of increase in one subscale was not related to the developmental trajectories in other subscale over time. That might be because Error-explanation prompting group received the error-explanation prompt. In each item, the error-explanation prompt was more contextual and specific for the item (Golke et al., 2015; Ting & Kuo, 2016). The given prompt in items within each subscale depended on the error pattern corresponding to the chosen response that not related to items within other subscales

Furthermore, the results of the multivariate model in Mixed prompting group found the associations between the rates of increase in understand subscale and evaluate and reflect subscale over time. The data indicated significant changes in both domains over time. The prompts provided in Mixed prompting group were the combination between the instructional and error-explanation prompts. The guided prompts in understand and evaluate and reflect might help each other in developmental trajectory.

For the multivariate model result in Verification prompting group, there was negative correlation between the intercept factor in locate information and the slope factor in understand subscale. This indicates that students who with higher initial locate information value had lower growth trajectories for understand subscale over time. This cross-domain interrelation identified that early locate information predicts understand subscale.

Our findings suggest that each reading literacy subscale is intertwined differently in accordance with the given prompting conditions. It is likely to believe that understand and evaluate and reflect continued to intertwine from each other. Instructional prompting group provides evidence that early understand score supported evaluate and reflect subscale across the developmental period. Moreover, Mixed prompting group pertains that the rate of increase in understand subscale might facilitate the developmental period in evaluate and reflect subscale. This supports the view that the development of understand score positively affects the evaluate and reflect score over time.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

5. Effectiveness of computerized dynamic assessment for reading literacy on learning potential in reading literacy of ninth graders

5.1 Effects of different prompting conditions of computerized dynamic assessment for reading literacy on availability scores

Theory of zone of proximal development of Vygotsky (1980) provides an insight into the difference between what students currently know and what they may realize in the future. In order to measure availability, the student-specific learning trajectories of reading literacy were computed using the Michaelis-Menten model. Availability is the distance between an individual's current ability and maximum capacity. The capacity score was the prediction of the student's capacity to learn reading literacy over three time points (McNeish & Dumas, 2017; McNeish et al., 2020). Practically, Poehner and Lantolf (2013) said that availability is indicated by how much investment in instructional activity is required in the classroom in the future. The findings from this study revealed that no significant differences among different prompting groups were found in availability scores during the first two assessment sessions. This might be because their distances between ability score and capacity score were similar among all prompting conditions. However, the findings showed that there were differences between Mixed prompting group and Verification prompting group in terms of availability score measured at the third testing session. This may be because Mixed prompting group reached the predicted reading literacy performance, whereas Verification prompting group did not attain their predicted capacity during their third time point. This suggests that in order to help students to reach their predicted potential, repeated measures with Mixed prompting group achieved the attainment faster than those in the Verification prompting group. They needed only three assessment time points to reach their predicted capacity, whereas Verification prompting group required more than three testing sessions. Consequently, the present study has some theoretical contributions. Following Vygotsky, the zone of proximal development discerns how close individuals are to

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

successful independent functioning (Poehner et al., 2015). Thus, Mixed prompting group might be the most effective prompt to help students attain their potential in reading literacy ability.

5.2 Effects of different prompting conditions of computerized dynamic assessment for reading literacy on mediated scores

Development of computerized dynamic assessment was based on the cake format defined by Sternberg and Grigorenko (2002) and the graduated prompt approach proposed by Campione and Brown (1985). Mediated score is a measure of the amount of mediation learners required as they worked through the test and how well learners responded to mediation (Poehner & Lantolf, 2013; Poehner et al., 2015). Following the graduated prompt approach, students can obtain mediation and be guided progressively when they have difficulties in solving problems. Several scholars have defined the way students profit from given prompts as learning potential (Dorfler et al., 2017; Resing & Elliott, 2011). The term 'learning potential' is used to observe students' responsiveness to the given support to foster learning and understanding in the focused domain (Dorfler et al., 2017; Poehner et al., 2015). In accordance with Poehner and Lantolf (2013), the number of prompts required provides an indication of the amount of instruction students need to reach a potential performance level. As a result, it is an essential part of diagnosis and more relevant to future teaching and learning (Poehner et al., 2015).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

The results revealed that there was no significant interaction between the two variables on mediated score, prompting conditions and time. Thus, the effects of the two factors were considered separately. The data indicated statistically significant differences among different prompting conditions with respect to mediated scores. The result showed that Verification prompting group obtained significantly lower mediated scores when compared with other groups. This suggests that Verification prompting group required more prompts to solve the problems during the training session compared with other prompting conditions. The result is consistent with

Resing and Elliott (2011) that students who scored lower tended to need more prompts. Group differences are evident because some groups needed a considerable number of prompts and scaffolds, while others needed little help. The results of Stevenson et al. (2016) indicated that students with lower skills may require more specific prompts such as scaffolds to effectively benefit from instruction. It might be due to Verification prompting group being provided either correct or incorrect. This may enhance the opportunity for students' guessing correct answers and need more assistance.

Moreover, the time effects on mediated scores were significant. The mediated scores were higher across three time points. This result can be explained by students demonstrating their knowledge and strategies learned from the given prompts from the previous assessment; as a result, they required fewer prompts. This result is consistent with previous studies (Resing & Elliott, 2011; Wang, 2011). The results of Resing and Elliott (2011) revealed that there was a decrease in the number of prompts students needed to solve the items from the first training session to the second training session. Wang, (2011) found that students needed significantly fewer prompts to answer the item correctly. This indicates that students clearly profited from the prompts.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

This suggested that the students were capable of using effective reading strategies and demonstrated awareness of their reading skill (Teo, 2012). Several researchers mentioned that the graduated prompt appears to be effective in improving student performance (King et al., 2015; Kozulin & Garb, 2002; Teo, 2012; Wang, 2011). Teo (2012) found that during the C-DA process, students became more strategic and reflective readers that could control their reading processes. King et al. (2015) mentioned that performance during dynamic assessment sessions might be predictive of future performance on a similar task. The less support required in dynamic assessment, the better the performance on the experimental task (King et al., 2015). Thus, following the graduated prompt approach, this study addresses

mediated scores as an effective measure in determining students' responsiveness to mediation, resulting in the learning potential of ninth graders' reading literacy ability.

5.3 Effects of different prompting conditions of computerized dynamic assessment for reading literacy on levels of prompting

Levels of prompting were separated into four levels of promptings as follows: 1) the first level of prompting, the second level of prompting, 3) the third level of prompting, and 4) the fourth level of prompting. Traditionally, learners were given an amount of assistance based on their incorrect attempt when the mediation is presented in a gradually implicit to explicit design (Teo, 2012). In this study, the researcher was interested in measuring the level of mediation students needed during the training session. A sequence of prompts ranged from 1) very general, 2) abstract and implicit, 3) explicit and specific, and 4) provide the correct answer. The levels of prompts indicated what kind of instruction students need to reach a potential performance level (Ai, 2017; Teo, 2012). The prompting approach plays a role as the teacher guiding or assisting learners. The type of prompts students need during training may guide choices of the most appropriate classroom instructions (Resing & Elliott, 2011; Stevenson et al., 2016).

In this part, the researchers paid attention to the second and the third levels of promptings. The second level of prompt determines the abstract and implicit guidance, whereas the third level of prompt is related to the more specific and explicit guidance. The results of the present study revealed that only Verification prompting group required significantly greater second level of promptings compared to Mixed prompting group in overall reading literacy. It should be noted that Verification prompting group received a verification prompt that did not provide assistance with respect to the content. This indicated that there were no differences in terms of implicit guidance among those who received either instructional prompt or error-explanation prompt or those who obtained both. However, Instructional prompting group required a significantly higher third level of promptings when

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

compared to Mixed prompting group. That might be because the explicit instructional guidance might not be enough to support students' solving problems in reading literacy. Even though prompts were narrowed down to guide students to focus on certain paragraphs or sentences, the explanation given at this level was not sufficient. This is coherent with the study of Teo (2012) that indicated some students repeatedly arrived at mediation level 3. They needed more mediation at this level than what was provided in the C-DA program. Teachers then observe the lacking and modify by adding a more explicit and more understandable assistance to the learner.

For locate information domain, Error-explanation prompting group received significantly greater second levels of prompting in locate information when compared to Mixed prompting group. Moreover, Verification prompting group got significantly higher prompts than Instructional prompting group and Mixed prompting group. This might be because instructional prompt provided in Instructional prompting group and Mixed prompting group might help students discover some principles or reading strategies that are necessary to handle problems (Poehner & Lantolf, 2013). This suggests that the instructional prompt might be sufficient in providing a solution for locate information domain.

In terms of understand domain, Verification prompting group required greater second levels of promptings compared to Error-explanation prompting group and Mixed prompting group. Moreover, there were significant differences in receiving the third level of prompting between Instructional prompting group and Mixed prompting group in understand subscale. It might be because the error-explanation prompt helped students to uncover their misunderstanding concepts. Moreover, the explicit instructional guidance might not be enough to support students' solving problems in understand subscale. Even though students were told to focus on certain paragraphs or sentences and were given explanations about the meaning of the texts, it seemed that they needed more powerfully explicit prompts in understand subscale. This suggests that the error-explanation prompt provided by Error-explanation prompting

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

group and Mixed prompting group was beneficial in solving the items in understand subscale. Golke et al. (2015) states that students would benefit from the prompt when they understand the errors and try to resolve it; as a result, the mistake would be less likely to occur on the subsequent assessment.

With respect to evaluate and reflect domain, Verification prompting group required greater second and third levels of promptings compared to Mixed prompting group. This result showed that both implicit and explicit guidance related to either instruction or error-explanation was sufficient for solving evaluate and reflect domain. Particularly the guidance combined instruction and error-explanation was more effective in supporting students' solving problems in evaluate and reflect subscale. As a result, students in different prompting conditions required varying levels of support in different reading literacy subscales when they were taken computerized dynamic assessment.

Overall, different prompting conditions of computerized dynamic assessment for reading literacy showed similarities and differences in some aspects. With respect to reading literacy, Verification prompting group significantly differed from other prompting-based groups, whereas treatment groups were not significantly different from each other. This indicated that computerized dynamic assessment for reading literacy with a wide range of promptings could be used to improve current performance in reading literacy. Graduated prompting technique was necessary.

For reading literacy growth, Mixed prompting group received the highest growth trajectory in reading literacy as well as locate information and understand subscales. This might be because mixed prompting helped students to perceive their own reading difficulties and offered reading strategies they were missing in the reading process that could be used to deliver a correct response (Golke et al., 2015). Moreover, students in Mixed prompting with higher growth in understand subscale received the rate of increase in evaluate and reflect subscale. This showed that they could transfer their knowledge to the next questions. Error-explanation prompting group had higher rate of growth in evaluate and reflect subscale, whereas they showed the lower rate of growth in locate information over time when compared with other

groups. This might be because Error-explanation prompting group might have opportunities to reflect on their own experience or knowledge to compare and contrast the response that might help them impove their awareness of the information provided within the text. Those skills reflected on higher-order reading process. Error-explanation prompt is preferred for tasks requiring higher-order cognitive processes (Golke et al., 2015; Petrović et al., 2017). However, their prompt might not increase their ability to scan through a text to retrieve specific pieces of information in locate information subscale. As a result, mixed prompting method might be effective to assess rate of growth in reading literacy. Moreover, error-explanation might be suitable for rate of growth in higher-order reading skills.

In terms of learning potential in reading literacy, Verification prompting group significantly required more prompts to solve problems when compared to other groups. This might be because Verification prompting group was provided either correct or incorrect that may enhance the opportunities for guessing correct answers and needed more assistance. The verification prompt might not sufficient in forming knowledge to solve problems, whereas other groups required less support and were able to use effective reading strategies and demonstrated their reading awareness (Teo, 2012). Levels of prompting could guide instruction needed in classroom. The results showed that Instructional prompting group required higher third level of promptings in overall reading literacy and understand subscale than Mixed prompting group. This might be because the explicit instructional guidance in Instructional prompting group might not be enough to support students' solving problems in reading literacy, especially in understand subscale. On the other hand, Error-explanation prompting group received significantly greater second levels of prompting in locate information when compared to Mixed prompting group. As a result, in order to assess learning potential, Mixed prompting was appropriate to assess and enhance learning potential in reading literacy. Instructional prompt was not sufficient for understand subscale and Error-explanation prompt was not suitable for locate information subscale.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Conclusion

We conclude that this research provides evidence for the psychometric properties of the parallel reading literacy tests and an insightful information pertaining to the effectiveness of computerized dynamic assessment on current performance of reading literacy, growth trajectory of learning, as well as responsiveness to mediation predicted as students' learning potential. Bifactor model was chosen as the most fitted model in estimating parameters of the reading literacy. Each reading literacy instrument was determined by two factors: a domain-general reading literacy factor for all items and four testlet factors for four reading passages. The developed parallel instruments may be confidently used as equivalent measures of reading literacy performance among ninth grade students. Evidence firmly supported the equivalence of scores from the developed parallel versions of reading literacy instruments using the statistical index of RMSD. In terms of current performance of reading literacy, the computerized dynamic assessment was developed in accordance with the theoretical basis of the graduated prompt approach proposed by Campione and Brown (1985). The prompting obtained by learners plays an important role in improving their learning effectiveness. Evidence showed that prompting-based groups received the higher posttest scores when compared to the control condition. For the growth of learning, a latent growth curve model was analyzed to investigate the influence of computerized dynamic assessment on students' reading literacy in multiple timeframes. Our study has identified that computerized dynamic assessment for reading literacy led to rate of growth in reading literacy and reading literacy subscales of ninth graders with different promptings. With respect to students' learning potential, computerized dynamic assessment for reading literacy acts as a mediator to promote learners' potential development and assess the learners' levels of prompting in the process of learning. The design of mixed promptings should be applied in the classroom for enhancing the growth rate and learning potential of ninth graders. Consequently, this

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

study represents a stepping stone for future research in dynamic assessment and its application in educational settings.

Implication and Recommendation of the Study

The current study can draw several implications for further research as follows:

1. Test construction and validity study of the reading literacy

It is noteworthy that the application of the bifactor model could be more adequate for passage-based reading literacy assessment when compared to other models and may be recommended to resolve the problem of local dependence as well as bias in analyzing the test and individual items of the reading literacy tests. Moreover, this research contributes to provide the empirical evidence supporting the validity of the instruments for measuring reading literacy as the bifactor model has capacity to consider diverse factors involved in reading literacy performance. Consequently, the results and suggestions provide a broad range of application and utility that can be used to develop a more appropriate version of a test, to have more accuracy in selection and construction for test items, as well as to improve the quality of the instrument as a whole.

For further studies, the computerized dynamic assessment for reading literacy seems limited in the item type and the number of items contained in each testlet. In this research, the instruments were multiple-choice questions, that each reading passage or testlet consisted of five multiple-choice items. Further studies can be designed to include items located in varying lengths within testlets. The research suggests that more different types of test items, such as fill in blank and open-ended test items, should be used to further understand students' learning conditions. Mixed format should be adapted to be closely matched the PISA test. Moreover, this study did not construct test items in accordance with the level of reading proficiency defined by OECD (2019). Further studies should consider the level of reading proficiency and task difficulty as guideline in construction of test items. This study

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

selected reading passages by using task, text, and level of transferability. Each test consisted of four reading passages that might not achieve a full range of texts associated with real-life situations. Hence, a wide range of reading materials both online and offline such as email, catalogues, or social media posts and comments, should be selected as reading passages for further research.

2. Statistical parallelism of the reading literacy instruments

The findings provide evidence that the test forms are parallel at the level of the individual items as well as at the level of the overall test. Both TCC and TIF yielded satisfactory degrees of test parallelism. Consequently, the three parallel test sets can be used as interchangeable instruments for reading literacy assessment. Moreover, this study provides evidence that both TCC and TIF could be used as statistical indices for parallelism of reading literacy. We hope that this leads to a broader awareness of the statistical methods necessary for establishing the parallel test forms. Finally, this research is beneficial to researchers and practitioners who are attempting to develop the alternate measures using both content and statistical considerations in order to preserve the validity of the test forms. Thus, the reading literacy instruments meet the requirements of parallel tests and can be used as an instrument for longitudinal data to track students' progress in learning and assessment.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

This study is limited only to RMSD in measuring conformity, a different study may be conducted with the different types of statistical indices such as the mean square deviation (MSD), the mean relative deviation (MRD), and the maximum relative information gap (MRIG).

3. Contribution of computerized dynamic assessment to measure student's reading literacy

This study generated evidence to support the notion that computerized dynamic assessment could assess student's reading literacy performance. The system can be administered to large numbers of learners and reports of the learners'

performance can be automatically generated. Practically, computerized dynamic assessment could be developed and widely used in classroom to assist teachers and educators in construction of an assessment-centered learning that engages assessment as teaching and learning strategy. Teachers can integrate the computerized dynamic assessment, along with the prompting design, to develop the system and prompting in other subject areas into their teaching in classrooms. They can also be performed in remedial classrooms in order to improve low-achiever learners' learning with the help of prompting approach.

There is a need to determine the effectiveness of different types of promptings of computer-based assessment of other grouping factors such as other grades or other subjects. Future studies are needed to investigate how these individually varying factors might affect reading literacy. Another issue that needs further study concerns students' attitudes towards computerized dynamic assessment. More qualitative and quantitative studies of the effectiveness of the computerized dynamic assessment for reading literacy is necessary. In addition, the design of computerized dynamic assessment for reading literacy should concern with user experience (UX) design to make the system more appropriate to the test-takers.

4. The study of growth trajectory of reading literacy

Our findings contribute to the reading literacy concept by specifying how reading literacy may change over time. Our result adds an interesting perspective to previous literature on comparing the initial value and rate of growth of different prompting conditions. Moreover, the results contribute to knowledge about the relationships among intercepts and changes with respect to reading literacy subscales of different prompting conditions. Consequently, results of the current study provide support for the computerized dynamic assessment as an effective tool for assessing the developmental reading literacy of ninth grade students in multiple timeframes. We expect this study to serve as a starting point and generate interest in enhancing reading literacy in longitudinal study.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

This study emphasized the immediate experiment effect without further investigating the delay effect. Future research could extend the timeframe of the experiment to several weeks after the testing sessions end. The study found the growth trajectory of students with different prompting methods. It seemed to have some unobserved factors responsible the change across three time points. Thus, longitudinal studies should investigate the factors affecting the growth trajectory of reading literacy and show the growth patterns of change over time. In our study, the researchers conducted three repeated measures for the growth model. Conducting the study with a longer interval and with more repeated measures might provide further information on growth in reading literacy and reading literacy subscales.

5. The estimation of responsiveness to mediation as the learning potential

This study confirmed the important role of graduated prompting assessment in a student's performance. The findings extend our understanding of how students responded to assistance and the level of mediation they used. We suggest that mixed prompting, which is guidance combined both instruction and error-explanation, could be used as an effective prompt to strengthen students' responsiveness to mediation. The present study has contributed to education by providing empirical evidence of the potential of the different types of promptings of computerized dynamic assessment to support students' reading literacy and to serve as the basis for future development of reading literacy field. For educators, the finding of different types of promptings of computer-based assessment to enhance reading literacy ability provide information to support teachers in classroom assessment and encourage them to make use of the graduated prompting assessment to promote students' learning potential in classroom and encourage students' self-learning. It is also helpful for educators to obtain information during the assessment session about the individual's learning processes. It is our hope that

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

the present study may offer a promising direction for future work in graduated prompting approach.

This study is focused on group differences in prompts given. Further work needed to be done on how individually varying factors, such as gender, age, and education level, have impact on students' responsiveness to mediation. Moreover, there is a need for qualitative study to determine the effectiveness of prompting approach on individual learning. Further study could also examine the design of prompts that encourage self-assessment and self-directed learning.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

REFERENCES

Ai, H. (2017). Providing graduated correcitve feedback in an intelligent computer-assisted language learning environment. *European Association for Computer Assisted Language Learning*, *29*(3), 313-334.

Ali, U. S., & van Rijn, P. W. (2016). An evaluation of different statistical targets for assembling parallel forms in item response theory. *Applied Psychological Measurement*, *40*(3), 163-179.

Ariel, A., Veldkamp, B. P., & Breithaupt, K. (2006). Optimal testlet pool assembly for multistage testing designs. *Applied Psychological Measurement*, *30*(3), 204-215.

Armstrong, R. D., Jones, D. H., & Kunce, C. S. (1998). IRT test assembly using network-flow programming. *Applied Psychological Measurement*, *22*, 237-472.

Baghaei, P., & Ravand, H. (2016). Modeling local item dependence in cloze and reading comprehension tests using testlet response theory. *Psicologica*, *37*, 85-104.

Baker, F. B. (2001). *The basics of item response theory* (Second ed.). ERIC Clearinghouse on Assessment and Evaluation.

Blozis, S. A., Conger, K. J., & Harring, J. R. (2007). Nonlinear latent curve models for multivariate longitudinal data. *International Journal of Behavioral Development*, *31*(4), 340-346.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, *46*(4), 443-459.

Boekkooi-Timminga, E. (1990). The construction of parallel tests from IRT-based item banks *Journal of Educational Statistics 15*(2), 129-145.

Bordens, K. S., & Abbott, B. B. (2011). *Research design and methods: A process approach*. McGraw-Hill.

Bowles, R. P., & Montroy, J. J. (2013). Latent growth curve modeling using structural equation modeling. In Y. Petscher, C. Schatschneider, & D. Compton (Eds.), *Applied quantitative analysis in education and the social sciences*. Routledge.

Bradlow, E. T., Wainer, H., & Wang, X. (1999). A Bayesian random effects model for testlets. *Psychometrika*, *64*(2), 153-168.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Byun, J.-H., & Lee, Y.-W. (2016). The latent trait modeling of passage-based reading comprehension test: Testlet-based MIRT approach. *English Language Assessment*, *11*, 25-45.

Cai, Y., & Kunnan, A. J. (2018). Examining the inseparability of content knowledge from LSP reading ability: an approch combining bifactor-multidimensional item response theory and structural equation modeling. *Language Assessment Quarterly*, *15*(2), 109-129.

Cameron, C. E., Grimm, K. J., Steele, J. S., & Castro-Schilo, L. (2015). Nonlinear Gompertz curve models of achievement gaps in mathematics and reading. *Journal of Educational Psychology*, *107*(3), 789-804.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally

Campione, J. C., & Brown, A. L. (1985). *Dynamic assessment: One approach and some initial data*. Bolt Beranek and Newman.

Chalmers, R. P. (2012). MIRT: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*(6).

Chandai, S. (2016). *Development of an instructional model based on scaffolded reading experiences approach and self-regulated learning for enhance reading literacy of lower secondary school students*. (Doctoral dissertation), Chulalongkorn University.

Clause, C. S., Mullins, M. E., Nee, M. T., Pulakos, E., & Schmitt, N. (1998). Parallel test form development: a procedure for alternate predictors and an example. *Personnel psychology*, *51*, 193-208.

Cotrus, A., & Stanciu, C. (2014). A study on dynamic assessment techniques, as a method of obtaining a high level of learning potential, untapped by conventional assessment. *Procedia-Social and Behavioral Sciences*, *116*(2014), 2616-2619.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson.

Davin, K. (2013). Integration of dynamic assessment and instructional conversations to promote development and improve assessment in the language classroom.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Language Teaching Research, *17*(3), 303-322.

Davis, E. A. (2009). Prompting middle school science students for productive reflection: generic and directed prompts. *The Journal of the Learning Sciences*, *12*(1), 91-142.

Debeer, D., Ali, U. S., & Rijn, P. W. V. (2017). Evaluating statistical targets for assembling parallel mixed-format test forms. *Journal of Educational Measurement*, *54*(2), 218-242.

Dechsri, P. (2016). Students' performance in PISA and the adequacy of teaching and learning In L. M. Thien, N. A. Razak, & J. P. Keeves (Eds.), *What can PISA2012 data tell us?* . Rotterdam: Sense.

DeMars, C. E. (2006). Application of the bi-factor multidimensional item response theory model to testlet-based tests. *Journal of Educational Measurement*, *43*(2), 145-168.

DeMars, C. E. (2012). Confirming testlet effects. *Applied Psychological Measurement*, *36*(2), 104-121.

Diowvilai, D., Samranjai, J., Sommano, B., Janwong, V., & Mookham, T. (2012). *Development of elementary children Grade 4-6 with reading literacy through Lampang learning enrichment network.* Lampang Rajabhat University.

Dorfler, T., Golke, S., & Artelt, C. (2017). *Evaluating prerequisites for the development of a dynamic test of reading competence: Feedback effects on reading comprehension in children*. Springer International Publishing.

Dumas, D. G., & McNeish, D. M. (2017). Dynamic measurement modeling: Using nonlinear growth models to estimate student learning capacity. *Educational Researcher*, *46*(6), 284-292.

Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy*, *35*, 333-363.

Dunn, K. J., & McCray, G. (2020). The place of the bifactor model in confirmatory factor analysis investigations into construct dimensionality in language testing. *Frontiers in Psychology*, *11*, 1357.

Ebadi, S., Weisi, H., Monkaresi, H., & Bahramlou, K. (2018). Exploring lexical inferencing as a vocabulary acquisition strategy through computerized dynamic assessment

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

and static assessment. *Computer Assisted Language Learning*, *31*(7), 790-817.

Eckes, T. (2014). Examining testlet effects in the TestDaF listening section: A testlet response theory modeling approach *Language Testing*, *31*(1), 39-61.

Eleje, L. I., & Esomonu, N. (2018). Test of achievement in quantitative economics for secondary schools: Construction and validation using item response theory. *Asian Journal of Education and Training*, *4*(1), 18-28.

Elliott, J. G. (2003). Dynamic assessment in educational settings: Realizing potential. *Educational Review*, *55*(1), 15-32.

Feuerstein, R., Falik, L. H., & Feuerstein, R. (2003). The learning potential assessment device. In R. Feuerstein (Ed.), *Feuerstein's theory and applied systems: A reader.* Retrieved from http://www.icelp.info/media/328542/Learning-Potential-Assessment-Device-Reader.pdf

Geiser, C., Keller, B. T., & Lockhart, G. (2013). First-versus second-order latent growth curve models: Some insights from latent state-trait theory. *Structural Equation Modeling: A Multidisciplinary Journal 20*(3), 478-503.

Gellert, A., & Elbro, C. (2017). Try a little bit of teaching: A dynamic assessment of word decoding as a kindergarten predictor of word reading difficulties at the end of Grade 1. *Scientific Studies of Reading*, *21*(4), 277-291.

Gibbons, R. D., & Hedeker, D. R. (1992). Full-information item bi-factor analysis. *Psychometrika*, *57*, 423-436.

Golke, S., Dorfler, T., & Artelt, C. (2015). The impact of elaborate feedback on text comprehension within a computer-based assessment. *Learning and Instruction*, *39*, 123-136.

Grigorenko, E., & Sternberg, R. (1998). Dynamic testing. *Psychological Bulletin*, *124*(1), 75-111.

Grimm, K. J., & Ram, N. (2009). Nonlinear growth models in Mplus and SAS. *Structural Equation Modeling*, *16*(4), 676-701.

Grimm, K. J., & Ram, N. (2011). Nonlinear Growth Curves in Developmental Research. *Child Development*, *82*(5), 1357-1371.

Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia*

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Journal of Mathematics, Science & Technology Education, *11*(5), 989-1008.

Guthke, J., & Wingenfeld, S. (1992). The learning test concept: origins, state of the art, and trends. In H. C. Haywood & D. Tzuriel (Eds.), *Interactive Assessment* (pp. 64-93). Springer.

Ha, D. T. (2017). Applying multidimensional three-parameter logistic model (M3PL) in validating a multiple-choice test. *International Journal of Scientific and Research Publications*, *7*(2), 175-183.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective* (Seventh ed.). Pearson Prentice Hall.

Hancock, G. R., Kuo, W.-L., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling*, *8*(3), 470-489.

Harring, J. R., Kohli, N., Silverman, R. D., & Speece, D. L. (2012). A second-order conditionally linear mixed effects model with observed and latent variable covariates. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(1), 118-136.

Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: clinical and educational applications*. Cambridge University Press

Hong, S., & Ho, H.-Z. (2005). Direct and indirect longitudinal effects of parental involvement on student achievement: Second-order latent growth modeling across ethnic groups. *Journal of Educational Psychology*, *97*(1), 32-42.

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55.

Ifenthaler, D. (2012). Determining the effectiveness of prompts for self-regulated learning in problem-solving scenarios. *Educational Technology & Society*, *15*(1), 38-52.

Jumnaksarn, S. (2013). The development of causal relationship model of factors influencing on reading literacy of 15-year-old students in Thailand. *Journal of education Research, Faculty of Education, Srinakharinwirot University*, *8*(2), 213-230.

Kanjanawasee, S. (2012). *Modern test theory* (Fourth ed.). Chulalongkorn University

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Printery.

Kantiwong, A. (2017). *Comparison of item and test parallel index between test forms that selected by different characteristics of experts: Application of formal cultural consensus theory*. (Master's thesis), Chulalongkorn University.

Kantiwong, A., & Pasiphol, S. (2018). Comparison of item and test parallel index between test forms among different characteristics of experts: Application of formal cultural consensus theory. *An Online Journal of Education*, *13*(4), 286-300.

Kelloway, E. K. (2015). *Using Mplus for structural equation modeling: A researcher's guide*. SAGE Publications Asia-Pacific

Kilic, D., & Saglam, N. (2009). Development of a two-tier diagnostic test to determine students' understanding of concepts in Genetics. *Eurasian Journal of Educational Research*(36), 227-244.

Kim, W. H. (2017). *Application of the IRT and TRT models to a reading comprehension test.* (Doctoral dissertation), Middle Tennessee State University

King, M. R., Binger, C., & Kent-Walsh, J. (2015). Using dynamic assessment to evaluate the expressive syntax of children who use augmentative and alternative communication. *Augmentative and Alternative Communication*, *31*(1), 1-14.

Kozulin, A. (2014). Dynamic assessment in search of its identity. In A. Yasnitsky, R. V. Veer, & M. Ferrari (Eds.), *The cambridge handbook of cultural-historical psychology,* Cambridge University.

Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension. *School Psychology International*, *23*(1), 112-127.

Ku, D. T., Shih, J.-l., & Hung, S.-H. (2014). The integration of concept mapping in a dynamic assessment model for teaching and learning accounting. *Educational Technology & Society*, *17*(1), 141-153.

Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: Bridging the past into the future. *Journal of Applied Linguistics*, *1*(1), 49-72.

Lantolf, J. P., & Poehner, M. E. (2014). Dynamic assessment of L2 development: Bridging the past into the future. *Journal of Applied Linguistics*, *1*(1), 49-72.

Lee, Y.-W. (2004). Examining passage-related local item dependence (LID) and

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

measurement construct using Q3 statistics in an EFL reading comprehension test. *Language Testing*, *21*(1), 74-100.

Lin, J.-W. (2016). Development and evaluation of the diagnostic power for a computer-based two-tier assessment. *J Sci Educ Technol*, 497-511.

Markon, K. E. (2019). Bifactor and hierarchical models: Specification, inference, and interpretation. *Annu. Rev. Clin. Psychol*, *15*, 51-69.

McNeish, D., & Dumas, D. (2017). Nonlinear growth models as measurement models: A second-order growth curve model for measuring potential. *Multivariate Behavioral Research*, *52*(1), 61-85.

McNeish, D., Dumas, D. G., & Grimm, K. J. (2020). Estimating new quantities from longitudinal test scores to improve forecasts of future performance. *Multivariate Behavioral Research*, *55*(6), 894-909.

Medawela, R. M. S., Ratnayaka, D., Abeyasinghe, W., Jayasinghe, R., & Marambe, K. (2018). Effectiveness of "fill in the blanks" over multiple choice questions in assessing final year dental undergraduates. *Educacion Medica*, *19*(2), 72-76.

Milenkovic, D. D., Segedinac, M. D., Hrin, T. N., & Horvat, S. (2016). The impact of instructional strategy based on the triplet model of content representation on elimination of students' misconceptions regarding inorganic reactions. *Journal of the Serbian Chemical Society*, *81*(6), 717-728.

Min, S., & He, L. (2014). Applying unidimensional and multidimensional item response theory models in testlet-based reading assessment. *Language Testing*, *31*(4), 453-477.

Mullis, I. V. S., Martin, M. O., & Sainsbury, M. (2015). *PIRLS 2016 assessment framework*. Retrieved from https://timssandpirls.bc.edu/pirls2016/framework.html

National Assessment Governing Board. (2017). *The NAEP reading assessment framework*. Retrieved from https://www.nationsreportcard.gov/reading_2017/about/framework?grade=4

Nilsawang, P. (2011). *Factors related to reading literacy of grade 9 students under primary education service area office in Sisaket province*. (Master's thesis), Mahasarakham University.

Nunnally, J. C. (1978). An Overview of Psychological Measurement. In Wolman B.B. (Ed.),

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Clinical Diagnosis of Mental Disorders. Springer.

OECD. (2016). *PISA2018 reading literacy framework*. OECD. Retrieved from https://www.iprase.tn.it/documents/20178/344196/Pisa+2018+reading+literacy+framework+_final.pdf/14f3abfc-966c-46b1-a8d8-4d962193ecfd

OECD. (2019a). *PISA 2018 assessment and analytical framework*. PISA. OECD Publishing.

OECD. (2019b). *PISA 2018 Released Field Trial New Reading Items.* Retrieved from http://www.oecd.org/pisa/test/PISA-2018-Released-New-REA-Items.pdf

OECD. (2019c). *PISA 2018 Technical Report*. Retrieved from https://www.oecd.org/pisa/data/pisa2018technicalreport/PISA2018%20TecReport-Ch-15-Proficiency-Scales.pdf

OECD/UNESCO. (2016). *Education in Thailand: An OECD-UNESCO Perspective*. OECD Publishing

Office of the Basic Education Commission. (n.d.). *The educational policies of the Office of the Basic Education Commission in the financial year 2018-2019*. Retrieved from https://www.obec.go.th/about/นโยบายสพฐ-ปีงบประมาณ-พ-ศ-2561

Office of the Education Council. (2017). *The National Education Plan of Thailand 2017-2036*. Retrieved from http://www.moe.go.th/moe/th/news/detail.php?NewsID=47936&Key=news20

Ouyang, X., Xin, T., & Chen, F. (2016). Construct validity of the Children's Coping Strategies Scale (CCSS): A bifactor model approach. *Psychological Reports*, *118*(1), 199-218.

Petrović, J., Pale, P., & Jeren, B. (2017). Online formative assessments in a digital signal processing course: Effects of feedback type and content difficulty on students learning achievements. *Education and Information Technologies*, *22*(6), 3047-3061.

Poehner, M. E. (2008). *Dynamic Assessment: A Vygotskian Approach to Understanding and Promoting L2 Development*. Springer

Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during computerized dynamic assessment (C-DA). *Language Teaching Research*, *17*(3), 323-342.

Poehner, M. E., Zhang, J., & Lu, X. (2015). Computerized dynamic assessment (C-DA):

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing 32*(3), 337-357.

Praputtakun, P., Dahsah, C., Tambunchong, C., & Mateapinikul, P. (2013). The case study of prethomsuksa 6 students' scientific literacy and reading ability. *Journal of Education Thaksin University*, *13*, 127-140.

Qiu, T., & Van Compernolle, R. A. (2021). computerized dynamic assessment of implicature comprehension in L2 Chinese. *Language Learning & Technology*, *25*(2), 55-74.

Ravand, H. (2015). Assessing testlet effect, impact, differential testlet, and item functioning using cross-classified multilevel measurement modeling. *SAGE Open*(2015), 1-9.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667-696.

Resing, W. C. M., & Elliott, J. G. (2011). Dynamic testing with tangible electronics: Measuring children's change in strategy use with a series completion task. *The British Psychological Society*, *81* (4), 579-605.

Rijmen, F. (2010). Formal relations and an empirical comparison among the bi-factor, the testlet, and a second-order multidimensional IRT model. *Journal of Educational Measurement*, *47*(3), 361-372.

Rovinelli, R. J., & Hambleton, R. K. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research*, *2*, 49-60.

Sabbag, A. G., & Zieffler, A. (2015). Assessing learning outcomes: An analysis of the goals-2 instrument. *Statistics Education Research Journal 14*(2), 93-116.

Sanders, P. F., & Verschoor, A. J. (1998). Parallel test construction using classical item parameters. *Applied Psychological Measurement*, *22*(3), 212-223.

Schmucker, A., Abberger, B., Boecker, M., & Baumeister, H. (2019). Parallel short forms for the assessment of activities of daily living in cardiovascular rehabilitation patients (PADL-cardio): Development and validation. *Disability and Rehabilitation*, *41*(7), 826-832.

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation*

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

modeling. Lawrence Erlbaum Associates.

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (Fourth ed.). Routledge.

Shih, J.-L., Ku, D. T., & Hung, S.-H. (2013). Enhancing the learning achievement and attitudes of Taiwan vocational school students in accounting with the dynamic assessment system. *Educational Computing Research*, *49*(3), 345-362.

Shrestha, P. N. (2017). Investigating the learning transfer of genre features and conceptual knowledge from an academic literacy course to business studies: Exploring the potential of dynamic assessment. *Journal of English for Academic Purposes*, *25*, 1-17.

Sternberg, R., & Grigorenko, E. (2002). *Dynamic testing: The nature and measurement of learning potential*. Cambridge University Press.

Stevenson, C., Heiser, W., & Resing, W. (2016). Dynamic testing: Assessing cognitive potential of children with culturally diverse backgrounds. *Learning and Individual Differences*, *47*, 27-36.

Teo, A. (2012). Promoting EFL students' inferential reading skills through computerized dynamic assessment. *Language Learning & Technology*, *16*(3), 10-20.

The Institute for the Promotion of Teaching Science and Technology. (2015). *PISA 2015 assessment and analytical framework*. Success Publication.

The Institute for the Promotion of Teaching Science and Technology. (2018). *PISA 2015 results in Science, Reading, and Mathematics: Excellence and equity in education.* Success Publication.

The Institute for the Promotion of Teaching Science and Technology. (2020). *PISA 2018 results: What student know and can do*. Retrieved from https://pisathailand.ipst.ac.th/issue-2019-48/

Thomson, S., Hillman, K., & Bortoli, L. D. (2013). *A teacher's guide to PISA reading literacy*. ACER.

Ting, M., & Kuo, B. (2016). A knowledge-structure-based adaptive dynamic assessment system for calculus learning. *Journal of Computer Assisted Learning*, *32*(2), 105-119.

Toland, M. D., Sulis, I., Giambona, F., Porcu, M., & Campbell, J. M. (2017). Introduction to

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

bifactor polytomous item response theory analysis. *Journal of School Psychology*, *60*, 41-63.

Tomas, J. M., Pinazo-Hernandis, S., Donio-Bellegarde, M., & Hontangas, P. M. (2017). Validity of the de Jong Gierveld loneliness scale in Spanish older population: Competitive structural models and item response theory. *Eur J Ageing*, *14*, 429-437.

Tzuriel, D. (2001). *Dynamic assessment of young children*. Plenum.

Tzuriel, D., & Shamir, A. (2002). The effects of mediation in computer assisted dynamic assessment. *Journal of Computer Assisted Learning*, *18*(1).

UNESCO Bangkok. (2016). *2017/2018 Global education monitoring report: Thailand highlights*. Retrieved from https://bangkok.unesco.org/content/20172018-global-education-monitoring-report-thailand-highlights

United Nations. (2018). *Sustainable development goals*. Retrieved from https://www.un.org/sustainabledevelopment/education/

van der Linden, W. J., & Adema, J. J. (1998). Simultaneous assembly of multiple test forms. *Journal of Educational Measurement*, *35*(3), 185-198.

Veldkamp, B. P., Matteucci, M., & Jong, M. G. d. (2012). Undertainties in the item parameter estimates and robust automated test assembly. *Applied Psychological Measurement*, *37*(2), 123-139.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, J., & Chen, S. (2016). Development and validation of an online dynamic assessment for raising students' comprehension of science text. *International Journal of Science and Mathematics Education*, *14*(3), 373-389.

Wang, T.-H. (2010). Web-based dynamic assessment: Taking assessment as teaching and learning strategy for improving students' e-Learning effectiveness. *Computers & Education*, *54*(4), 1157-1166.

Wang, T. (2011). Implementation of Web-based dynamic assessment in facilitating junior high school students to learn mathematics. *Computers & Education*, *56*(4), 1062-1071.

Wang, W.-C., & Wilson, M. (2005). The Rasch testlet model. *Appilted Psychological*

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Measurement, *29*(2), 126-149.

Wu, H.-M., Kuo, B.-C., & Wang, S.-C. (2017). Computerized dynamic adaptive tests with immediately individualized feedback for primary school mathematics learning. *Educational Technology & Society*, *20*(1), 61-72.

Yang, T.-C., Fu, H.-T., Hwang, G.-J., & Yang, S. J. H. (2017). Development of an interactive mathematics learning system based on a two-tier test diagnostic and guiding strategy. *Australasian Journal of Educational Technology*, *33*(1), 62-77.

Yang, Y., & Qian, D. D. (2020). Promoting L2 English learners' reading proficiency through computerized dynamic assessment. *Computer Assisted Language Learning*, *33*(5-6), 628-652.

Zhang, J., & Lu, X. (2019). Measuring and supporting second language development using computerized dynamic assessment. *Language and Sociocultural Theory*, *6*(1), 92-115.

Zhang, R.-C., Lai, H.-M., Cheng, P.-W., & Chen, C.-P. (2017). Longitudinal effect of a computer-based graduated prompting assessment on students' academic performance. *Computers & Education*, *110*, 181-194.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

APPENDICES

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Appendix A

The List of Experts

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

รายนามผู้ทรงคุณวุฒิที่ตรวจความตรงเชิงเนื้อหาและความเป็นคู่ขนาน
ของแบบวัดการรู้เรื่องการอ่านฉบับก่อนเรียนและหลังเรียน

1. รองศาสตราจารย์ ดร.สุวรรณี ยหะกร	อาจารย์ประจำสาขาหลักสูตรและการสอน(การสอนภาษาไทย) มหาวิทยาลัยสุโขทัยธรรมาธิราช
2. ผู้ช่วยศาสตราจารย์ ดร.เด่นดาว ชลวิทย์	อาจารย์ประจำสาขาวิชาเอกภาษาไทย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง
3. ผู้ช่วยศาสตราจารย์ ดร.ภาณุพงศ์ อุดมศิลป์	อาจารย์ประจำภาควิชาภาษาไทยและภาษาตะวันออก คณะมนุษยศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
4. อาจารย์ ดร.สันติวัฒน์ จันทร์ได	อาจารย์สอนวิชาภาษาไทย ระดับมัธยมศึกษา โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม
5. ผู้ช่วยศาสตราจารย์ ดร.ศศิธร บัวทอง	อาจารย์ประจำแขนงวิชาการวัดและประเมินผลการศึกษา มหาวิทยาลัยสุโขทัยธรรมาธิราช

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

รายนามผู้ทรงคุณวุฒิที่ตรวจความตรงเชิงเนื้อหาและความเป็นคู่ขนาน ของแบบวัดการรู้เรื่องการอ่านฉบับที่หนึ่ง (Test set 1), ฉบับที่สอง (Test set 2), และฉบับที่สาม (Test set 3)

กลุ่มที่หนึ่ง รายนามของผู้ทรงคุณวุฒิที่ตรวจสอบความตรงเชิงเนื้อหาและความเป็นคู่ขนานของแบบวัดการรู้เรื่องการอ่าน (ข้อสอบข้อที่ 1-5 ข้อที่ 21-25 และข้อที่ 41-45)

1. อาจารย์ ดร. สง่า วงค์ไชย	อาจารย์ประจำสาขาวิชาเอกภาษาไทย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง
2. อาจารย์ ดร.ทัศนีย์ เศรษฐพงษ์	อาจารย์ประจำสาขาวิชาภาษาไทย คณะครุศาสตร์ มหาวิทยาลัยราชภัฏสวนสุนันทา
3. ครูปองภพ ปะวันเนย์	ครูสอนวิชาภาษาไทย ระดับมัธยมศึกษา โรงเรียนชิโนรสวิทยาลัย
4. อาจารย์เชรษฐรัฐ กองรัตน์	อาจารย์สอนวิชาภาษาไทย ระดับมัธยมศึกษา โรงเรียนสาธิตมหาวิทยาลัยขอนแก่น ฝ่ายมัธยม
5. อาจารย์ ดร.วารุณี ลัภนโชค	อาจารย์สาขาวิชาวิจัยและประเมินทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

กลุ่มที่สอง รายนามของผู้ทรงคุณวุฒิที่ตรวจสอบความตรงเชิงเนื้อหาและความเป็นคู่ขนานของแบบวัดการรู้เรื่องการอ่าน (ข้อสอบข้อที่ 6-10 ข้อที่ 26-30 และข้อที่ 46-50)

6. ผู้ช่วยศาสตราจารย์ ดร.อภิรักษ์ อนะมาน	อาจารย์ประจำวิชาหลักสูตรและการสอน(การสอนภาษาไทย) มหาวิทยาลัยสุโขทัยธรรมาธิราช
7. อาจารย์ ดร.พิณพนธ์ คงวิจิตต์	อาจารย์ประจำสาขาวิชาการสอนภาษาไทย มหาวิทยาลัยศิลปากร
8. ครูมินตรา ลายสนิทเสรีกุล	ครูสอนวิชาภาษาไทย ระดับมัธยมศึกษา โรงเรียนเตรียมอุดมศึกษาน้อมเกล้า
9. ครูปัญจพล บุลาลม	ครูสอนวิชาภาษาไทย ระดับมัธยมศึกษา โรงเรียนชิโนรสวิทยาลัย
10. ผู้ช่วยศาสตราจารย์ ดร.ฉัตรศิริ ปิยะพิมลสิทธิ์	อาจารย์สาขาวิชาวิจัยและประเมินทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

กลุ่มที่สาม รายนามของผู้ทรงคุณวุฒิที่ตรวจสอบความตรงเชิงเนื้อหาและความเป็นคู่ขนานของแบบวัดการรู้เรื่องการอ่าน (ข้อสอบข้อที่ 11-15 ข้อที่ 31-35 และข้อที่ 51-55)

11. ผู้ช่วยศาสตราจารย์เจษฎา บุญมาโฮม	อาจารย์ประจำวิชาหลักสูตรและการสอน (ภาษาไทย) มหาวิทยาลัยราชภัฏนครปฐม
12. อาจารย์เจษฎา ศรีวัฒนเมธากุล	อาจารย์ประจำสาขาภาษาไทย คณะศิลปศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช
13. อาจารย์ ดร.เฉลิมลาภ ทองอาจ	อาจารย์สอนวิชาภาษาไทย ระดับมัธยมศึกษา โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม
14. ครูจริยา ประเสริฐนู	ครูสอนวิชาภาษาไทย ระดับมัธยมศึกษา โรงเรียนภูเขียว
15. อาจารย์ ดร.เกียรติยศ กุลเดชชัยชาญ	อาจารย์ประจำสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ

กลุ่มที่สี่ รายนามของผู้ทรงคุณวุฒิที่ตรวจสอบความตรงเชิงเนื้อหาและความเป็นคู่ขนานของแบบวัดการรู้เรื่องการอ่าน (ข้อสอบข้อที่ 16-20 ข้อที่ 36-40 และข้อที่ 56-60)

16. ผู้ช่วยศาสตราจารย์ ดร.กมลพัทธ์ ใจเยือกเย็น	อาจารย์สอนภาษาไทย คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏพระนคร
17. อาจารย์ธนชพร พุ่มภชาติ	อาจารย์ประจำสาขาวิชาเอกภาษาไทย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง
18. อาจารย์มณิการ์ ชูทอง	อาจารย์สอนวิชาภาษาไทย ระดับมัธยมศึกษา โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม
19. ครูนนทชัย บุตรสอน	ครูสอนวิชาภาษาไทย ระดับมัธยมศึกษา โรงเรียนนครพนมวิทยาคม
20. ผู้ช่วยศาสตราจารย์ ดร.พินดา วราสุนันท์	อาจารย์ภาควิชาครุศึกษา คณะศึกษาศาสตร์และพัฒนศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Appendix B

The Results of IOC Index of the Reading Literacy Pretest and Posttest

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1. The results of IOC index of the reading literacy pretest

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	1	1	1	0	1	1	4	1	0	0.80		ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	2	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	3	1	1	-1	0	1	3	1	1	0.40	▪ ตัวเลือกที่ถูกโดดมาก จากคำถามจะเป็นการเดาใจผู้เขียนมากกว่าการหาใจความสำคัญและสรุปความ ▪ คำถามจะเป็นการถามความรู้สึก	เปลี่ยนคำถาม
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	4	1	1	1	1	0	4	1	0	0.80	▪ ปรับคำ ▪ ข้อ ก และ ง ดูจะไม่ต่างกันนัก	ใช้ได้
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	5	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	6	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	7	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	8	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	9	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	10	1	0	0	1	0	2	3	0	0.40	▪ คำถามน่าจะเป็นความเข้าใจหรือไม่ ? ▪ งงคำถาม	เปลี่ยนคำถาม
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	11	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	12	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	13	1	1	1	1	1	5	0	0	0.66	▪ ปรับคำ	ใช้ได้

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	14	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	15	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	16	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	17	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	18	1	0	1	1	1	4	1	0	0.80		ใช้ได้
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	19	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	20	1	1	1	0	1	4	1	0	0.80	▪ ข้อ ข มีประเด็นแค่ 1 ประเด็นเลยไม่สามารถบอกได้ว่าสอดคล้องกับอะไร ข้ออื่น ๆ มี 2 ประเด็น คำตอบไม่แสดงความสอดคล้อง ▪ ควรปรับ ข้อ ก.รถยนต์ไร้คนขับ ไม่สอดคล้องกับรถยนต์ที่ควบคุมโดยมนุษย์	ใช้ได้

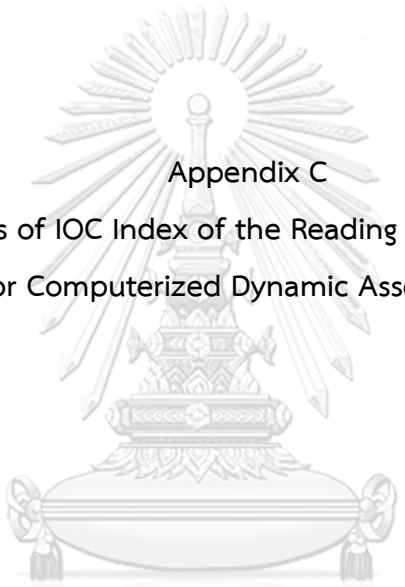
2. The results of IOC index of the reading literacy posttest

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	21	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	22	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	23	1	1	0	1	0	3	2	0	0.60	▪ ถ้าคำตอบที่ถูกคือ ข ควรปรับตัวเลือก จ เพราะตัวเลือก จ ก็ถูก	ใช้ได้
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	24	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ ▪ คำถามไม่ตรงกับวัตถุประสงค์	ใช้ได้
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	25	1	1	1	1	0	4	1	0	0.80	▪ ข้อนี้อาจเป็นความรู้นอกเหนือจากบทอ่านมากจนเกินไป จะทราบได้อย่างไรว่าเว็บไซค์นี้น่าเชื่อถือ	เปลี่ยนคำถาม
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	26	1	1	1	1	1	5	0	0	1.00	▪ ควรเขียนใต้ภาพให้ชัดเจนว่ารูปภาพที่ 1 เพราะเลข 1 ด้านบทจะเขียนว่าแนวโน้มผู้ประกอบอาชีพอิสระในไทย ซึ่งไม่สัมพันธ์กัน	ใช้ได้
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	27	1	1	1	1	1	5	0	0	1.00	▪ ควรบอกใต้ภาพให้ชัดเจนว่าภาพที่ 2 ▪ ชื่อภาพคือ soft skills ที่คนทำงานควรมี แต่ในคำถามคือถามทักษะที่นายจ้างต้องการ คำถามนี้เป็นการตีความ จะไม่ตรงกับสมรรถนะการรู้เรื่องการอ่าน	ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	28	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	29	1	1	1	0	1	4	1	0	0.80	▪ ในรูปไม่ได้เขียนบอกว่าเป็นรูปที่ 1 และรูปที่ 2 อย่างชัดเจน ▪ นักการตลาดออนไลน์มีการสื่อสารแค่ทางเดียว ไม่น่าจะมีการสื่อสารที่ดี สื่อสารน่าจะเป็นการติดต่อสองทางแบบปัจจุบัน ▪ รูปที่ 1 เป็นเรื่องของคนๆเดียว แต่รูปที่ 2 เป็นเรื่องของการทำงานร่วมกับคนอื่น ไม่น่าจะสัมพันธ์กัน ▪ ข้อ ข น่าจะถูกต้องด้วยไหม ต้องเปลี่ยนตัวเลือกใหม่	ใช้ได้
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	30	1	1	0	0	0	2	3	0	0.40	▪ ในรูปที่ 1 และ 2 คือรูปไหน ควรระบุด้วย ▪ ข้อ ก ควรปรับคำตอบเพราะมีแค่ 1 ประเด็น ▪ คำถามไม่ชัดเจนเพราะรูปไม่ชัดเจน ▪ ข้อ ข ไม่ใช่ข้อมูลจากรูป	เปลี่ยนคำถาม

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
											▪ ภาพ 1 เป็นเหมือนอาชีพอิสระ แต่ภาพที่ 2 เป็นการทำงานในองค์กร ดังนั้นต้องการถามอะไร คำตอบข้อ ค, ง, จ เหมือนกัน	
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	31	1	1	1	1	0	4	1	0	0.80	▪ จะดูได้อย่างไร? ควรมีการทำรูปมือ click	ใช้ได้
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	32	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	33	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	34	1	1	1	1	0	4	1	0	0.80	▪ แก้ไขให้เป็นไปในทิศทางเดียวกัน คือการดูจากคะแนนและ percentile ไม่ใช่ปีพศ.	ใช้ได้
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	35	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	36	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	37	1	0	1	1	1	4	1	0	0.80	▪ ลักษณะคำถามคล้ายข้อ 39	ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	38	1	1	1	1	1	5	0	0	1.00	▪ ปรับภาษาของตัวเลือกข้อ จ เพื่อให้เป็นบูรณาการและอนุมานมากขึ้น	ใช้ได้
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	39	1	0	1	1	1	4	1	0	0.80	▪ ลักษณะคำถามคล้ายข้อ 37 เป็นการถาม “ทัศนะ” หรือไม่ ▪ ปรับคำ	ใช้ได้
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	40	1	1	1	0	1	4	1	0	0.80	▪ คำตอบแต่ละข้ออ่านแล้วไม่เข้าใจว่าความสอดคล้องคืออะไร หรืออะไรไม่ใช่ความสอดคล้อง ควรเขียนขยายความให้ชัดเจน เช่น การไม่ใช่ครีมกันแดด ไม่สอดคล้องกับ การเป็นโรคมะเร็งผิวหนัง	ใช้ได้

Appendix C

The Results of IOC Index of the Reading Literacy Test Sets for Computerized Dynamic Assessment



จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1. The results of IOC index of the reading literacy (Test Set 1)

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	1	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำถาม ▪ ไม่ต้องใช้คำว่าเธอในตัวเลือก	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	0	4	1	0	0.80	▪ ปรับคำ ▪ เรียงตัวเลือกสั้น-ยาว	ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	2	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	3	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	▪ ตัวเลือกโดด 1 ตัว ▪ ถ้าจะเล่นคำที่สงท้ายด้วย “ใจ”และมีอารมณ์ในทิศทางเดียวกัน ควรใช้ทุกคำ	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	4	ข้อสอบระดับที่หนึ่ง	1	1	1	0	1	4	1	0	0.80	▪ ปรับคำถาม จากทำไม เป็น “เพราะเหตุใด”	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	5	ข้อสอบระดับที่หนึ่ง	1	1	1	1	0	4	1	0	0.80	▪ ปรับคำ เว็บไซต์ และอีเมล ▪ ปรับภาษาในข้อคำถาม ▪ ไม่แน่ใจว่า ก ถูก? ▪ ข้อ ข. เปลี่ยนจากชื่อเป็นประวัติ	ใช้ได้
		ข้อสอบระดับที่สอง	0	1	0	1	1	3	2	0	0.60	▪ ความทันสมัยกับช่วงเวลาต่างกันอย่างไร ▪ จะใช้ “ท่าน” หรือ “นักเรียน” ให้ใช้เหมือนกัน ▪ แก้ไขตัวเลือกตามที่ผู้ทรงเสนอ	ใช้ได้
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	6	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	▪ เด็กจะรู้ไหมว่านกคือทวิตเตอร์	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	7	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	8	ข้อสอบระดับที่หนึ่ง	1	0	1	1	0	4	2	0	0.60	▪ ปรับคำว่า “ได้กล่าว” เป็น “ข้อความ” ว่า ▪ สิ่งที่ให้อ่านยังไม่มีลักษณะของบทความ ตัวเลือกไม่มีลักษณะเป็น “ความหมาย”	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	0	4	1	0	0.80	▪ ข้อ ก มีความใกล้เคียง	ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	9	ข้อสอบระดับที่หนึ่ง	1	0	0	1	0	2	3	0	0.40	▪ ไม่มีข้อมูลมาสรุปได้ว่าวัยรุ่นเหมาะสมกับไลน์ อีกทั้งนิยามวัยรุ่นคืออะไร วัยเริ่มทำงานเป็นวัยรุ่นได้หรือไม่ วัยผู้สูงอายุอาจะเล่นไลน์มากกว่าวัยรุ่นก็ได้ ไม่ทราบว่ารูปที่ 1,2 หมายถึงรูปใด (ใต้ภาพไม่ได้ระบุ)	เปลี่ยนคำถาม
		ข้อสอบระดับที่สอง	1	0	0	1	0	2	3	0	0.40		เปลี่ยนคำถาม
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	10	ข้อสอบระดับที่หนึ่ง	1	0	0	1	0	2	3	0	0.40	▪ สรุปยาก	เปลี่ยนคำถาม
		ข้อสอบระดับที่สอง	1	0	0	1	0	2	3	0	0.40	▪ สรุปยาก ▪ ข้อ ก บอกไว้ที่ใด ข้อมูลไม่ควรใช้การอนุมานเอาเอง เพราะอาจจะผิดได้ ดังนั้นจากภาพข้อมูลอาจจะไม่เพียงพอที่จะตอบให้ชัดเจน	เปลี่ยนคำถาม
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	11	ข้อสอบระดับที่หนึ่ง	1	1	1	-1	1	4	0	1	0.60	▪ อาจเรียงตัวเลือกตามลำดับความสั้นยาว ▪ ไม่มีข้อความระบุว่า ง. เป็นผู้จัดโครงการ จึงไม่มีหลักฐานเพียงพอที่จะรุปได้ว่าหน่วยงานใดเป็นผู้จัดการค้นหาและเลือกสารสนเทศต้องมีคำตอบที่ถูกต้องชัดเจน	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	-1	1	4	0	1	0.60	▪ ตัวเลือก ค. อาจใช้คำให้คล้ายกับ ก. และ ข. เช่น เป็นชื่อหน่วยงานที่แสดงไว้อย่างชัดเจนภายใต้ประกาศของกิจกรรมถุงผ้าสื่อธรรม ▪ ถ้าใช้เหตุผลนี้ อาจมีผู้แย้งได้ จากข้อมูลที่มีไม่ได้สรุปว่าใครจัด การมีสัญลักษณ์ใต้ชื่อไม่ใช่ข้อสรุปได้	ใช้ได้
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	12	ข้อสอบระดับที่หนึ่ง	1	1	1	0	1	4	1	0	0.80	▪ น่าจะเพิ่มตัวเลือกอื่น ๆ เช่น จ. ไม่ปรากฏจำนวนผู้สนใจร่วมโครงการ ▪ เพิ่มคำว่า “ทั้งหมด” ลงไปในข้อคำถาม	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ ตัวเลือก ค. อาจใช้คำว่าผู้ที่ให้ เหมือนกับ ก. และ ข. เพราะพิจารณาจากจำนวนผู้ที่ตอบรับโครงการจำนวน 29 คนและผู้ที่รอตอบรับอีก 9 คน	ใช้ได้

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
												▪ ปรับคำ ▪ ขอให้ปรับตัวเลือก ก และ ข ให้เป็นตัวลวงที่หลอกให้คนมาตอบได้ เช่น เป็นตัวเลขที่เกินจากการลงทะเบียนออนไลน์	
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	13	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	▪ ถ้ามีภาษาที่ซ้ำกันทุกข้อให้นำไปรวมไว้ในคำถาม ตัดคำว่า “จิตอาสา”ออก	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ เพิ่มเติม ก.เพราะเป็นกิจกรรมที่เน้นความสำคัญของการภาวนาใจ ข. เพราะเป็นกิจกรรมที่เน้นการสอนหลักธรรมให้แก่ผู้ที่เข้าร่วม ค. เพราะเป็นกิจกรรมที่ต้องการให้ผู้เข้าร่วมปักผ้าเป็น ▪ ให้เพิ่มความยาวของตัวเลือกอื่น ๆ ให้ใกล้เคียงกับตัวเลือกที่ถูกต้อง	ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	14	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	▪ ตัวเลือกข้อ ก, ง, จ ใช้คำว่าต้องการ แต่ ค ใช้อยากควรเลือกใช้คำให้ไปในทิศทางเดียวกัน ▪ เปลี่ยนชื่อคน ▪ ปรับ ค เป็นสมใจอยากรู้จักเพื่อนใหม่และทำกิจกรรมร่วมกับผู้อื่น ▪ ปรับคำ	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	0	0	3	2	0	0.60	▪ ปรับแก้ตัวเลือกให้สอดคล้องกับข้อ ง เพื่อให้เป็นตัวลองที่มีคนมาเลือกตอบ	ใช้ได้
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	15	ข้อสอบระดับที่หนึ่ง	1	0	0	0	1	2	3	0	0.40	▪ ตัวเลือกต่างๆควรสอดคล้องกับโจทย์ ทั้งไวยากรณ์และรูปแบบโดยภาพรวม ใช้ภาษาระดับทางการแต่ตัวเลือก ง. ใช้คำว่า “เบอร์โทร” ควรแก้ไขเป็น “หมายเลขโทรศัพท์ ” ▪ จากตัวเลือก ก ข ค สามารถเลือกได้ทั้ง 3 ข้อ เสนอว่าปรับโจทย์คำว่า “ถูกที่สุด” น่าจะชัดกว่านี้ ▪ ข้อ ข เป็นตัวเลือกที่ยาวที่สุดเมื่อเทียบกับข้ออื่นๆ ควรปรับความยาวให้เท่ากัน ▪ ตัวลวงยังหลอกคนตอบไม่ได้	เปลี่ยนคำถาม
		ข้อสอบระดับที่สอง	1	1	1	1	0	4	1	0	0.80	▪ อาจจะเปลี่ยนตำว่า “ดู” เป็น “พิจารณา”	เปลี่ยนคำถาม

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	16	ข้อสอบระดับที่หนึ่ง	1	0	1	1	1	4	1	0	0.80	▪ เหมือนข้อสอบวัดความเข้าใจในการอ่าน ไม่ได้วัดตรงตามสมรรถนะที่ต้องการ	ใช้ได้
		ข้อสอบระดับที่สอง	1	0	1	1	1	4	1	0	0.80	▪ ควรหลีกเลี่ยงการใช้ภาษาปาก “แบน” ให้ใช้คำว่า “งดใช้” หรือ “ห้ามใช้” ดีกว่า	ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	17	ข้อสอบระดับที่หนึ่ง	0	1	0	1	1	3	2	0	0.60	▪ ข้อความ“ ”อยู่ในย่อหน้าสุดท้ายซึ่งมีประโยคสนับสนุน ทำให้เข้าใจว่าคำตอบคือข้ออื่น ▪ ข้อ ก น่าจะถูกต้องด้วยหรือไม่	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	18	ข้อสอบระดับที่หนึ่ง	1	1	1	0	1	4	1	0	0.80	▪ ข้อสอบง่ายไปหากพิจารณาตามวัตถุประสงค์เชิงพฤติกรรม ▪ ข้อคำถามสอดคล้องแต่เฉลยอาจจะกำกวมระหว่างข้อ ข และ ค	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	0	1	4	1	0	0.80		ใช้ได้
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	19	ข้อสอบระดับที่หนึ่ง	1	1	0	0	1	3	2	0	0.60	▪ ตัวเลือก ข ประโยคสื่อความไม่ชัดเจน ▪ ปรับคำถาม ▪ ข้อคำถามไม่สอดคล้องกับวัตถุประสงค์เชิงพฤติกรรม ควรปรับภาษาเป็นการถามจุดประสงค์ของผู้เขียน น้ำเสียง หรือมุมมอง	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	0	1	4	1	0	0.80		ใช้ได้
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	20	ข้อสอบระดับที่หนึ่ง	0	1	1	1	1	4	1	0	0.80	▪ เนื้อหาในข้อคำถามไม่ครบถ้วน ▪ ปรับคำถาม เรียกว่า “การกระทำที่ขัดแย้งกัน” หรือ “ไม่สัมพันธ์กัน” ดีหรือไม่	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้

2. The results of IOC index of the reading literacy (Test Set 2)

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	21	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำถาม	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	22	ข้อสอบระดับที่หนึ่ง	1	1	1	0	1	4	1	0	0.80	▪ ปรับคำ ▪ ปรับคำถาม	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	23	ข้อสอบระดับที่หนึ่ง	1	0	1	1	1	4	1	0	0.80	▪ ใช้คำว่าคุณลักษณะ แทนคุณสมบัติดีไหม ▪ ให้แทนคำว่าผู้พูด ควรใช้คำว่า พระอาจารย์ หรือระบุให้ชัดเจน ▪ ปรับคำถาม คุณสมบัติที่เหมาะสม เปลี่ยนเป็นหลักปฏิบัติที่สำคัญ ▪ ปรับภาษา	ใช้ได้
		ข้อสอบระดับที่สอง	1	0	1	1	1	4	1	0	0.80		ใช้ได้
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	24	ข้อสอบระดับที่หนึ่ง	1	1	1	1	0	4	1	0	0.80	▪ ปรับคำ	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	25	ข้อสอบระดับที่หนึ่ง	1	1	1	0	0	3	2	0	0.60	▪ ความคิดเห็นน่าจะเป็นข้อมูล ไม่ใช่สารสนเทศหรือไม่? ▪ ควรใช้คำเหมือนกันว่าจะใช้ ท่าน หรือ นักเรียน ▪ ตัวลวงผิดชัดเจนและสร้างความสับสน เพราะไม่มีปรากฏในบทความ ▪ ข้อ ก,ข ถูก ▪ ปรับภาษาในคำถาม	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ ตัวเลือกข้อ ข แก้เป็น เพราะผู้พูดเป็นบุคคลต้นแบบในการสอน	ใช้ได้
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	26	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	27	ข้อสอบระดับที่หนึ่ง	0	1	1	1	1	4	1	0	0.80	▪ มีข้อความว่า “หรือส่งพัสดุมาได้” ทำให้เข้าใจได้ว่า ส่งทางไปรษณีย์ใกล้บ้านได้ไหม? ▪ ปรับคำ ▪ น่าจะถามถึงที่รับทั้ง 2 ที่	ใช้ได้
		ข้อสอบระดับที่สอง	0	1	1	1	1	4	1	0	0.80	▪ ข้อความระบุ จุดรับ 2 ที่	ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	28	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	29	ข้อสอบระดับที่หนึ่ง	1	1	0	1	0	3	2	0	0.60	▪ ตัวลวงง่าย ▪ เหตุผลขาดน้ำหนัก	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	0	1	0	3	2	0	0.60	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	30	ข้อสอบระดับที่หนึ่ง	1	0	1	1	0	3	2	0	0.60	▪ ไม่แน่ใจว่าคำถามคืออะไรอาจต้องปรับภาษา	เปลี่ยนคำถาม
		ข้อสอบระดับที่สอง	1	0	0	1	0	2	3	0	0.40	▪ ไม่ค่อยสัมพันธ์กัน	เปลี่ยนคำถาม
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	31	ข้อสอบระดับที่หนึ่ง	0	1	1	-1	0	3	2	0	0.60	▪ ข้อมูลในบทอ่านไม่ได้แสดงอย่างชัดเจนว่าโครงการเมล็ดพันธุ์แห่งความดีอยู่ในแหล่งใด อาจจะอยู่ในข่าวเด่นหรือข่าวประชาสัมพันธ์ซึ่งไม่มีในตัวเลือกก็ได้ ▪ คำถามอาจจะเพิ่มเป็น ถ้าต้องการทราบข้อมูลเพิ่มเติมเกี่ยวกับการประกวด... ต้องสืบค้นจากแหล่งใดของเว็บไซต์ ▪ คำตอบไม่สามารถค้นหาได้จากข้อมูลที่มีอยู่ทั้งสองภาพ ▪ คำตอบสามารถตอบได้มากกว่า 1 ข้อ ควรหาคำถามใหม่ที่มีคำตอบชัดเจน	ใช้ได้
		ข้อสอบระดับที่สอง	0	1	1	-1	0	3	2	0	0.60	▪ คำถามเปลี่ยนจากแหล่งใด เป็นหัวข้อใดบนเว็บไซต์ ▪ คนที่ไม่เคยเข้าจะตอบไม่ได้จากข้อมูลที่	ใช้ได้แต่มีการปรับคำ

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
												กำหนดให้	
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	32	ข้อสอบระดับที่หนึ่ง	0	1	1	1	0	3	2	0	0.60	▪ ไม่ใช่ค้นคืนแต่น่าจะเป็นเข้าใจมากกว่า เพราะไม่ได้มีข้อมูลแสดงอย่างชัดเจนว่าครั้งแรกปีใด ▪ ถ้าคำตอบทุกข้อเป็น พศ อาจเปลี่ยนโจทย์เป็นครั้งแรกในปีใด หรือเมื่อใด ▪ อาจจะเพิ่มตัวเลือกไม่ปรากฏ	ใช้ได้
		ข้อสอบระดับที่สอง	0	1	1	1	0	3	2	0	0.60	▪ แม้ปี 62 จะเป็นครั้งที่ 4 แต่โครงการนี้ไม่ได้ระบุว่าจัดขึ้นทุกปี ดังนั้นครั้งที่ 1 อาจจะไม่ใช่ปี 2559 ก็ได้ ▪ ตัวเลือกโดดมาก ข้อ ข,ค ไม่ค่อยเกี่ยวข้อง ▪ ลองหาตัวเลือกที่เป็นตัวลวงอื่น ๆ	ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	33	ข้อสอบระดับที่หนึ่ง	1	1	1	-1	1	4	0	1	0.60	▪ เสนอ ก. รวมกลุ่ม ข. เลือกใช้คำใดคำหนึ่ง “อนุรักษ์” หรือ “รักษา” ▪ มีคำตอบที่ถูกต้องมากกว่า 1ข้อ ข้อ ง ตามเงื่อนไข 1.6 และข้อ จ ตามเงื่อนไข 1.7	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	-1	1	4	0	1	0.60	▪ ข. เพราะโครงการนี้ต้องการพัฒนาให้เยาวชนเติบโตเป็นผู้ใหญ่ที่มีความสามารถ ค. เพราะโครงการนี้....	ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	34	ข้อสอบระดับที่หนึ่ง	1	1	1	-1	0	3	2	0	0.60	▪ ใครเป็นผู้ที่มีคุณสมบัติเหมาะสมในการเข้าร่วมโครงการ... มากที่สุด ▪ ปรับคำ ▪ เรียงลำดับความสั้น-ยาวของตัวเลือก	ใช้ได้

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
		ข้อสอบระดับที่สอง	1	1	1	-1	0	3	2	0	0.60	▪ ปรับคำ ส่วนร่วม ▪ ควรให้เหตุผลที่ตรงกับเกณฑ์ที่เป็นเงื่อนไขพิจารณาทุนและเกี่ยวข้องกับคำตอบ ง	ใช้ได้
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	35	ข้อสอบระดับที่หนึ่ง	1	1	0	-1	0	2	2	1	0.20	▪ ปรับคำ ▪ ความน่าเชื่อถือคืออะไร? นิยาม? อาจจะต้องนำเงื่อนไขที่ไม่ได้อยู่ในประกาศมาใช้	เปลี่ยนคำถาม
		ข้อสอบระดับที่สอง	1	1	0	-1	0	2	2	1	0.20	▪ ตัวเลือกทั้ง 3 ไม่เป็นไปในทิศทางเดียวกัน ▪ ตอบได้มากกว่า 1 ข้อ	เปลี่ยนคำถาม
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	36	ข้อสอบระดับที่หนึ่ง	1	0	1	1	1	4	1	0	0.80	▪ ควรเติม “ตามลำดับ” ในบทอ่าน	ใช้ได้
		ข้อสอบระดับที่สอง	1	0	0	1	1	3	2	0	0.60	▪ คิดว่าบทความไม่ได้ระบุลำดับที่ชัดเจนเพียงพอ ▪ ควรตัดคำว่า “มากที่สุด” เพราะมากที่สุดคืออินเดีย	ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	37	ข้อสอบระดับที่หนึ่ง	1	1	1	0	1	4	1	0	0.80	▪ ปรับภาษา “พยายาม” อาจจะไม่ได้หมายถึง “ไม่ทำเลย”	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	38	ข้อสอบระดับที่หนึ่ง	1	1	0	0	1	3	2	0	0.60	▪ สามารถตอบข้อ จ ได้หรือไม่	ใช้ได้แ
		ข้อสอบระดับที่สอง	1	1	1	0	1	4	1	0	0.80		ใช้ได้
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	39	ข้อสอบระดับที่หนึ่ง	1	1	1	0	1	4	1	0	0.80	▪ ปรับข้อคำถาม เสนอแนะให้ใช้ “จากข้อความข้างต้นสะท้อนถึงจุดมุ่งหมายของผู้เขียนอย่างไร”	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	40	ข้อสอบระดับที่หนึ่ง	0	1	1	1	1	4	1	0	0.80		ใช้ได้
		ข้อสอบระดับที่สอง	0	1	1	1	1	4	1	0	0.80		ใช้ได้

3. The results of IOC index of the reading literacy (Test Set 3)

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	41	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	▪ ควรเรียงลำดับตัวเลือกจากสั้นไปยาว	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ ตัวเลือกมีทั้ง ไรน้ำ และ ไรน้ำนางฟ้า ควรระบุให้เหมือนกัน ▪ ในบทอ่านประโยคที่ว่า “เนสนใจการเลี้ยงปลาสวยงามอยู่เป็นทุนเดิม” สามารถนำมาเป็นเหตุผลได้หรือไม่?	ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	42	ข้อสอบระดับที่หนึ่ง	1	1	1	0	1	4	1	0	0.80	▪ ปรับคำ ▪ เปลี่ยนข้อความที่ยกมาในคำถาม	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	43	ข้อสอบระดับที่หนึ่ง	1	1	1	1	0	4	1	0	0.80	▪ เรียงลำดับความสั้นยาวของตัวเลือก ▪ แก้คำถามเป็น หลักคิดใดที่เนนำมาใช้เพื่อเพิ่มความมั่นใจให้กับตนเอง	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ แก้ตัวเลือก ก เป็น เพราะเนเป็นเด็กเก็บตัวมาโดยตลอด จึงกลัวคนอื่นไม่ยอมรับในสิ่งที่เธอเป็น ▪ ปรับคำในตัวเลือกที่ถูก	ใช้ได้
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	44	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	▪ ไม่ควรขึ้นต้นประโยคด้วยคำบุพบท ▪ ปรับคำ	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00	▪ ใช้คำฟุ่มเฟือย หากตัดคำว่า “ผู้เขียนกล่าวว่าความหมายยังคงเดิมหรือไม่	ใช้ได้
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	45	ข้อสอบระดับที่หนึ่ง	1	1	1	1	0	4	1	0	0.80	▪ ปรับคำ ▪ ให้ใช้คำว่า ท่าน หรือ นักเรียน ให้ใช้เหมือนกัน ▪ พิจารณาตัวเลือก ▪ ปรับภาษาที่คำถาม	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	46	ข้อสอบระดับที่หนึ่ง	1	0	1	1	1	4	1	0	0.80	▪ นักเรียนต้องมีความรู้ทางภูมิศาสตร์ด้วย จึงไม่แน่ใจว่าเป็นข้อสอบวัดที่ดี	ใช้ได้
		ข้อสอบระดับที่สอง	1	0	1	1	1	4	1	0	0.80		ใช้ได้

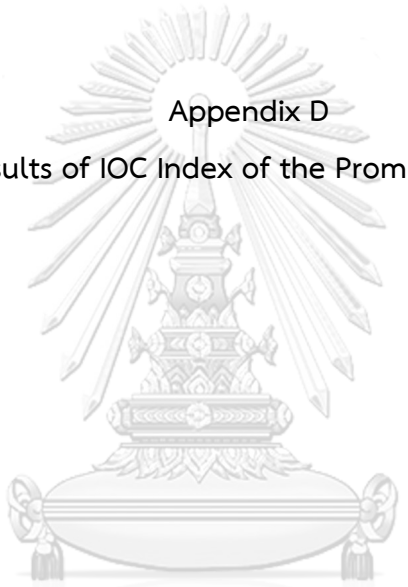
วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	47	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	48	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	▪ แก้จากคำว่า บทความ เป็นบทอ่าน	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	49	ข้อสอบระดับที่หนึ่ง	0	0	1	1	1	3	2	0	0.60		ใช้ได้
		ข้อสอบระดับที่สอง	0	0	1	1	1	3	2	0	0.60		ใช้ได้
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	50	ข้อสอบระดับที่หนึ่ง	0	1	0	1	1	3	2	0	0.60	▪ วัยผู้ใหญ่หรือวัยสูงอายุ ▪ ควรปรับคำถาม เช่น ข้อใดสรุปไม่ถูกต้อง ซึ่งง่ายและชัดเจนกว่า ▪ ตัวเลือก ก.ข ไม่ชัดเจน	เปลี่ยนคำถาม
		ข้อสอบระดับที่สอง	0	1	0	1	1	3	2	0	0.60	▪ วัยทำงานก็ถือเป็นวัยผู้ใหญ่ไหม?	เปลี่ยนคำถาม
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	51	ข้อสอบระดับที่หนึ่ง	0	0	1	0	1	2	3	0	0.40	▪ ควรเป็นสมรรถนะเข้าใจมากกว่าค้นคืน ▪ เรียงตัวเลือกตามลำดับสั้น-ยาว ▪ ควรเปลี่ยนข้อคำถาม เช่น สนใจทางด้านรูปทรง ▪ คำถามนี้คนทำไม่จำเป็นต้องอ่านเนื้อหาก็สามารถตอบได้ จึงดูเป็นคำถามที่แม้เนื้อหาเกี่ยวข้องแต่ไม่เหมาะสมกับระดับชั้นของผู้ทำข้อสอบ เสนอให้เปลี่ยนคำถามใหม่	เปลี่ยนคำถาม
		ข้อสอบระดับที่สอง	1	1	1	0	1	4	1	0	0.80	▪ ตัวเลือกไม่ดึงดูดให้เลือก ▪ ข้อ ข กับ ค สามารถเป็นคำตอบได้ แต่ข้อ ข ชัดเจนกว่าเท่านั้น ▪ ตัวเลือกที่เป็นตัวลวงไม่เกี่ยวกับศิลปะ ควรปรับตัวลวงใหม่ให้สามารถหลอกให้กกลุ่มอ่อนมาตอบ	เปลี่ยนคำถาม
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	52	ข้อสอบระดับที่หนึ่ง	1	1	1	1	0	4	1	0	0.80	▪ คำถามลักษณะนี้หากไม่อ่านก็อาจตอบได้ อาจผิดวัตถุประสงค์ที่ต้องการเน้นวิเคราะห์ ▪ ควรหาคำตอบที่ถูกเป็นอย่างอื่น เพราะการระบายสี (การใช้สี) ตัวเลือกอาจ	ใช้ได้

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
												ตีความได้หลายอย่าง	
		ข้อสอบระดับที่สอง	1	1	1	1	0	4	1	0	0.80	▪ ปรับคำ	ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	53	ข้อสอบระดับที่หนึ่ง	1	1	1	0	1	4	1	0	0.80	▪ เรียงตัวเลือกจากสั้นยาว ▪ ข้อถูกมีคำว่า “วาดลายเส้น” เหมือนโจทย์ ต้องปรับแก้ตัวเลือก ให้มีคำว่าภาพลายเส้น ทุกข้อหรือไม่เช่นนั้นก็ตัดคำนี้ออกจากตัวเลือก ข	ใช้ได้แต่มีการปรับคำ
		ข้อสอบระดับที่สอง	1	1	1	0	1	4	1	0	0.80	▪ ปรับคำ ▪ เสนอให้ลองตัดคำว่า การวาดภาพลายเส้นออก ถ้าตัดออก แล้วใจความยังเหมือนเดิม ก็ควรตัดออกเพราะเหตุผลที่ให้มาพอจะหลอกกลุ่มอ่อนได้แล้ว	ใช้ได้
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	54	ข้อสอบระดับที่หนึ่ง	1	1	1	-1	0	3	2	0	0.60	▪ เรียงตัวเลือกจากสั้นยาว ▪ ควรปรับคำถาม เพราะ ง ก็เป็นคำตอบที่ถูก ▪ ตัวเลือกมันยาวไปน่าจะแยกเป็นตามเงื่อนไข	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	0	0	3	2	0	0.60	▪ ปรับแก้ตัวเลือก ก	ใช้ได้
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	55	ข้อสอบระดับที่หนึ่ง	1	1	0	-1	0	3	2	0	0.60	▪ คำถามไม่ค่อยสอดคล้องกับคำตอบอาจแก้ไขเป็น “ข้อใดไม่ใช่สิ่งที่สนับสนุนให้ประกาศการอบรมวาดภาพลายเส้นมีความน่าเชื่อถือ ▪ ปรับคำถามความน่าเชื่อถือ หรือไม่ก็ปรับตัวเลือก โดยเอาสิ่งที่นอกเหนือจากในประกาศมาใส่ในตัวเลือก ▪ แก้โจทย์ใหม่ อ่านแล้วไม่เข้าใจว่าต้องการถามอะไรกับผู้อ่าน โจทย์ในข้อนี้อนุมานว่าจะถามว่าข้อใดที่ไม่เกี่ยวข้องกับความน่าเชื่อถือของประกาศ หรือข้อใดมราส่งผลต่อความน่าเชื่อถือของประกาศน้อยที่สุด	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	0	0	0	3	2	0	0.60	▪ ตัวลวงไม่ดึงดูดให้เลือก ▪ ปรับแก้ตัวเลือก ก และ ข ให้สามารถเป็นตัวลวงได้	ใช้ได้
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	56	ข้อสอบระดับที่หนึ่ง	1	0	1	1	1	4	1	0	0.80		ใช้ได้
		ข้อสอบระดับที่สอง	1	0	1	1	1	4	1	0	0.80		ใช้ได้
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	57	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
		ข้อสอบระดับ	1	1	1	1	1	5	0	0	1.00		ใช้ได้

วัตถุประสงค์เชิงพฤติกรรม	ข้อที่		คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	การแปลผล
			คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
		ที่สอง											
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	58	ข้อสอบระดับที่หนึ่ง	1	1	1	1	1	5	0	0	1.00	ข้อ ง ช่วยเพิ่มมูลค่าของสัตว์หรือไม่	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	59	ข้อสอบระดับที่หนึ่ง	1	1	1	0	1	4	1	0	0.80	ปรับคำถาม	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	60	ข้อสอบระดับที่หนึ่ง	0	1	1	0	1	3	2	0	0.80	▪ ปรับแก้ภาษา ▪ ใช้คำว่า “การโต้แย้ง” “ความคิดเห็นที่ขัดแย้ง” ในข้อคำถามน่าจะชัดเจนกว่า	ใช้ได้
		ข้อสอบระดับที่สอง	1	1	1	1	1	5	0	0	1.00		ใช้ได้

Appendix D

The Results of IOC Index of the Prompting Methods



จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1. The results of IOC index of the prompting methods of reading literacy (Test Set 1)

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		คนที่ 1	**คนที่ 2**	**คนที่ 3**	**คนที่ 4**	**คนที่ 5**	**+ 1**	**0**	**- 1**		
1	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
2	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
3	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ตอบผิดครั้งที่ 2 ให้พิจารณาว่าผู้พูดกล่าวในเชิงบวกหรือเชิงลบ โดยสังเกตจากน้ำเสียงของผู้พูดว่ามีน้ำเสียงเศร้าใจ กังวลใจ ห่วงใย เป็นต้น
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
4	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ เปลี่ยนเป็นย่อหน้าใดของบทอ่าน และ สไตล์=รูปแบบ? ▪ เพิ่มคำว่า “ซึ่ง”
5	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		คนที่ 1	**คนที่ 2**	**คนที่ 3**	**คนที่ 4**	**คนที่ 5**	**+ 1**	**0**	**- 1**		
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
6	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ ▪ นักเรียนบางคนอาจจะไม่ทราบสัญลักษณ์ของโซเชียลมีเดีย
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	0	1	1	1	1	4	1	0	0.80	▪ ข้อความของ ข และ ง เหมือนกันหรือไม่?
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ ▪ ข้อความค่อนข้างยาว เด็กอาจไม่อ่าน
7	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	-1	4	0	1	0.60	▪ ปรับคำ ▪ คำเฉลยไม่มีข้อมูล
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ มีศัพท์เฉพาะที่เขียนเป็นคำภาษาอังกฤษ
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	-1	4	0	1	0.60	▪ ปรับคำ
8	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
9	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	0	0	1	1	3	2	0	0.60	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	0	0	1	0	2	3	0	0.40	▪ แก้ไขตัวช่วยคำชี้แนะ
	ตัวช่วยคำชี้แนะแบบผสม	1	0	0	1	0	2	3	0	0.40	▪ แก้ไขตัวช่วยคำชี้แนะ
10	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	0	0	1	1	3	2	0	0.60	▪ ตัวเฉลยไม่มีข้อมูลมายืนยัน
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	0	0	1	1	3	2	0	0.60	▪ ปรับคำ

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		คนที่ 1	**คนที่ 2**	**คนที่ 3**	**คนที่ 4**	**คนที่ 5**	**+ 1**	**0**	**- 1**		
	ตัวช่วยคำชี้แนะแบบผสม	1	0	0	1	1	3	2	0	0.60	▪ ปรับคำ
11	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	0	0	1	3	2	0	0.60	▪ การสะกดคำ “เว็บไซต์” “ปรากฏ” ▪ ตอบผิดครั้งที่ 2 เป็นการค้นหาและเลือก ไม่น่าจะใช่การกวาดสายตา ข้อความกวาดสายตาควรอยู่ครั้งที่ 1 ▪ ปรับคำ ▪ ถ้าไม่ได้บอกว่าใครจัดชัดเจน นักเรียนจะตอบตามสัญลักษณ์ที่มีในเนื้อหา คำอธิบายจะไม่ได้ช่วยให้ตอบได้ในครั้งต่อไป จะดูเป็นการตีความหรืออนุมานจากเนื้อเรื่องที่กำหนดให้
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	0	1	4	1	0	0.80	▪ ปรับคำ ▪ การเฉลยที่มีคำแนะนำแบบนี้เป็นการอธิบายที่ใช้ข้อมูลนอกเหนือจากเนื้อหาที่กำหนดให้ซึ่งเป็นข้อมูลที่ผู้อ่านไม่ได้รับ ผู้เขียนรู้คนเดียว การอธิบายต้องอยู่บนพื้นฐานข้อมูลเดียวกันกับที่ผู้อ่านได้รับข้อมูล จึงจะถือว่ายุติธรรมกับผู้ทำข้อสอบ
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	0	1	4	1	0	0.80	▪ ปรับคำ ▪ เนื้อหาที่อธิบายเกินกว่าสิ่งที่ผู้อ่านได้รับ
12.	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
13	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	0	4	1	0	0.80	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	0	4	1	0	0.80	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	0	4	1	0	0.80	▪ ปรับคำ
14	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	0	4	1	0	0.80	▪ คำว่า พระรัตนตรัย สะกดผิด ▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ จากกิจกรรมหลัก เปลี่ยนเป็น เป้าหมายหลัก
15	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะ	1	1	0	1	0	3	2	0	0.60	▪ ปรับคำ เว็บไซต์ จิตอาสา กฎหมาย

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		คนที่ 1	**คนที่ 2**	**คนที่ 3**	**คนที่ 4**	**คนที่ 5**	**+ 1**	**0**	**- 1**		
	อธิบายข้อผิดพลาด										▪ กิจกรรมที่น่าเชื่อถือมีอะไรบ้าง? ถูกต้องตามกฎหมายมีความน่าเชื่อถือมั้ย? ▪ ตอบผิดครั้งที่ 3 ต้องพิจารณาปัจจัยอื่นร่วมด้วย คือต้องมีกี่ปัจจัย (อาจปรับคำถาม)
	ตัวช่วยคำชี้แนะแบบผสม	1	1	0	1	0	3	2	0	0.60	▪ ปรับคำ เบอร์โทร เว็บไซต์ ▪ ปรับคำถามว่า ถามถึงเรื่องอะไร เป็น มีความน่าเชื่อถือหรือไม่ ▪ ตอบผิดครั้งที่ 2 ความถูกต้องของข้อมูล ต้องระบุให้ชัดว่าความถูกต้องคืออะไร พิจารณาจากอะไร
16	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ในส่วนของเฉลยควรปรับรูปประโยคเพื่อลดความซับซ้อน
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
17	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ลักษณะคำถามคล้ายข้อ 39 ▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	▪ แยกไม่ค่อยออกกับแบบการสอน
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ รายละเอียดยาวมาก
18	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับภาษาของตัวเลือกข้อ จ เพื่อให้เป็นบูรณาการและอนุมานมากขึ้น
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	0	1	4	1	0	0.80	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	0	1	4	1	0	0.80	
19.	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	0	1	4	1	0	0.80	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	0	1	4	1	0	0.80	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	0	1	4	1	0	0.80	
20	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	

2. The results of IOC index of the prompting methods of reading literacy (Test Set 2)

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ	
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1			
21.	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	■	ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00		
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00		
22	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	■	ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	■	ปรับคำ
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	■	เทคนิคการเดาความจากบริบท (context clue)
23	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	■	ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00		
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00		
24	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	■	ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00		
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00		
25	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00		ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	■	ปรับคำ ตรวจตำสะกดให้ถูกต้อง เช่น เว็บไซต์ อีเมล
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00		
26	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00		
	ตัวช่วยคำชี้แนะที่มีลักษณะ	0	1	1	1	1	4	1	0	0.80	■	ตอบผิดครั้งที่ 3 มีข้อความปฏิเสธซ้ำ

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1		
	อธิบายข้อผิดพลาด										
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
27	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	0	1	1	1	1	4	1	0	0.80	▪ ตอบผิดครั้งที่ 2 และ 3 มีที่ไหนบ้าง เนื้อความมีมากกว่า 1 ที่
28	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
29	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	0	1	1	4	1	0	0.80	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	0	1	-1	3	1	1	0.60	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	0	1	1	4	1	0	0.80	
30	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ คำเฉลยยังไม่สมบูรณ์ ▪ คำเฉลยมีประเด็นที่ขัดแย้งกัน
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
31.	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	0	-1	0	2	2	1	0.20	ปรับคำ ▪ ตอบผิดครั้งที่ 1 การชี้แนะให้มากกว่าที่จะให้
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	0	0	3	2	0	0.60	▪ ตอบผิดครั้งที่ 2 ให้ตัวเลือกนี้ ข้อมูลเพิ่มเติมหาไม่ได้ใน ง จัดซื้อจัดจ้าง นอกนั้นหาข้อมูลเพิ่มเติมได้หมด
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	0	4	1	0	0.80	▪ การสะกดคำ ▪ การอธิบายเพิ่มเติมไม่ได้ช่วยให้ผู้ตอบมีโอกาสตอบมากขึ้นเพราะข้อมูลที่เป็นคำตอบอยู่นอกเหนือจากข้อมูลที่กำหนดให้ ไม่ได้เป็นทักษะการสืบค้นแต่เป็นการอนุมานตามประสบการณ์เดิม

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1		
32	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	0	1	0	1	1	3	2	0	0.60	▪ ตอบผิดครั้งที่ 1 การชี้แจงละเอียดเกินกว่าที่คำถามถามมั้ย?
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	0	1	1	1	0	3	2	0	0.60	▪ ปรับคำ เช่น ปรากฏ
	ตัวช่วยคำชี้แนะแบบผสม	0	1	1	1	0	3	2	0	0.60	
33	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	-1	1	4	0	1	0.80	▪ ตอบผิดครั้งที่ 3 ควรดูเงื่อนไขให้ครบถ้วนและละเอียด
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	-1	1	4	0	1	0.80	▪ คำตอบสามารถตอบได้มากกว่า 1 ข้อ
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	-1	0	3	1	1	0.40	▪ ตอบผิดครั้งที่ 3 เป็นการเฉลยที่ไม่ยุติธรรม
34	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	-1	1	4	0	1	0.60	▪ ตอบผิดครั้งที่ 3 ควรใช้หลักฐานที่เป็นเงื่อนไขของการให้ทุนสนับสนุน คำโฆษณาใน FB หลักฐานอ่อนกว่า
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	0	-1	1	3	1	1	0.40	▪ ตอบผิดครั้งที่ 3 การชี้แนะไม่ค่อยขยายความมากเท่าใดนัก
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	-1	1	4	0	1	0.60	▪ ใช้เงื่อนไขเป็นเกณฑ์ตัดสิน
35	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	-1	1	4	0	1	0.60	▪ เฉลยได้มากกว่า 1 ข้อ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	-1	1	4	0	1	0.60	▪ ตัวเลือกตอบได้มากกว่า 1 ข้อ
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	-1	1	4	0	1	0.60	▪ ตัวเลือกตอบได้มากกว่า 1 ข้อ
36	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	0	1	1	4	1	0	0.80	▪ ตอบผิดครั้งที่ 2 ข้อ ง และ จ ไม่แน่ใจว่าสามารถสรุปดีความได้เลยหรือไม่
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
37	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ลักษณะคำถามคล้ายข้อ 39 ▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1		
38	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	0	1	4	1	0	0.80	▪ ปรับภาษาของตัวเลือกข้อ จ เพื่อให้เป็นบูรณาการและอนุมานมากขึ้น
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	0	1	4	1	0	0.80	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	0	1	4	1	0	0.80	
39	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
40.	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	

จุฬาลงกรณ์มหาวิทยาลัย CHULALONGKORN UNIVERSITY

3. The results of IOC index of the prompting methods of reading literacy (Test Set 3)

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1		
41	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
42	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	ตอบผิดครั้งที่ 3 ขาดคำเชื่อมประโยค “แต่” ที่เป็นเหตุเป็นผล
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
43	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
44	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
45	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ
46	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	0	1	1	1	4	1	0	0.80	▪ เพิ่มรายละเอียดของคำเฉลย
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	0	1	1	1	4	1	0	0.80	
	ตัวช่วยคำชี้แนะแบบผสม	1	0	1	1	1	4	1	0	0.80	
47	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1		
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	-1	4	0	1	0.60	
48	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
49	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	0	0	1	1	1	3	2	0	0.60	▪ ตอบผิดครั้งที่ 3 เปลี่ยนจากอีบุ๊ค เป็นอื่นๆ
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	0	0	1	1	1	3	2	0	0.60	
	ตัวช่วยคำชี้แนะแบบผสม	0	0	1	1	1	3	2	0	0.60	
50	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	0	1	1	4	1	0	0.80	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	0	1	1	4	1	0	0.80	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	0	1	1	4	1	0	0.80	
51	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	0	1	4	1	0	0.80	▪ ตอบผิดครั้งที่ 1 กล่าวถึงเรื่องอะไร มีเนื้อความอย่างไร อาจจะดูเจาะลึกเกินไป ▪ คำตอบที่ถูกต้อง อาจจะต้องอธิบาย/ขยายความเพิ่มเติม ▪ คำถามไม่เหมาะสมกับนักเรียน เพราะดูจากการชี้แนะต้องตอบได้แน่นอน
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	0	0	3	2	0	0.60	▪ ปรับคำ
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	0	0	3	2	0	0.60	▪ อาจต้อง click เข้าไปในหมวดศิลปะ แล้วดูข้างใน จากนั้นนำมาขยายความตรงที่ต้องการชี้แนะผู้ทำแบบทดสอบ
52	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับคำ เว็บไซต์
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบาย0ข้อผิดพลาด	1	1	1	1	0	4	1	0	0.80	▪ มีคำปฏิเสธซ้อน เช่น ไม่ใช่ ไม่ได้
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	0	4	1	0	0.80	
53	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	0	4	1	0	0.80	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	0	4	1	0	0.80	
54	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ คำตอบควรบอกเกณฑ์ด้วยว่า อายุเท่าไหร่ วันอะไร เงินเท่าไหร่ และทำไมคนที่อยู่ชลบุรีอาจจะไม่ใช่
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบาย	1	1	0	1	0	3	2	0	0.60	

ข้อคำถาม	ลักษณะของตัวช่วยคำชี้แนะ	คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1		
	ข้อผิดพลาด										
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	0	4	1	0	0.80	
55	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	0	0	3	2	0	0.60	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	0	0	3	2	0	0.60	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	0	0	3	2	0	0.60	▪ ปรับคำ เบอร์โทรเป็นหมายเลขโทรศัพท์ ▪ เว็บไซต์เป็นเว็บไซต์
56	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ไม่สอดคล้องกับสมรรถนะการอ่าน
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	▪ ตอบผิดครั้งที่ 3 ควรแก้เป็น งาของสัตว์เป็นอวัยวะที่ถูกนำมาใช้ประโยชน์
57	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ลักษณะคำถามคล้ายข้อ 39
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	-1	1	1	1	1	4	0	1	0.60	▪ ไม่สอดคล้องกับข้อคำถาม เพราะเป็นคำชี้แนะของข้อ 58
	ตัวช่วยคำชี้แนะแบบผสม	-1	1	1	1	1	4	0	1	0.60	▪ ไม่สอดคล้องกับข้อคำถาม เพราะเป็นคำชี้แนะของข้อ 58
58	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	▪ ปรับภาษาของตัวเลือกข้อ จ เพื่อให้เป็นบูรณาการและอนุมานมากขึ้น
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
59	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	
60	ตัวช่วยคำชี้แนะที่มีลักษณะการสอน	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะที่มีลักษณะอธิบายข้อผิดพลาด	1	1	1	1	1	5	0	0	1.00	
	ตัวช่วยคำชี้แนะแบบผสม	1	1	1	1	1	5	0	0	1.00	

Appendix E

The Results of IOC Index of the Parallelism of Reading Literacy Test Sets for Computerized Dynamic Assessment

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

วัตถุประสงค์เชิงพฤติกรรม	คู่ที่	ความเป็นคู่ขนาน			คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนน	ข้อเสนอแนะ
		ฉบับที่ หนึ่ง (ข้อที่)	ฉบับที่ สอง (ข้อที่)	ฉบับที่ สาม (ข้อที่)	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1	เฉลี่ย (IOC)	
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	1.1	1	21		1	1	0	1	1	4	1	0	0.80	
	2.1	1		41	1	1	0	1	1	4	1	0	0.80	
	3.1		21	41	1	1	1	1	1	5	0	0	1.00	
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	1.2	2	22		1	1	1	1	1	5	0	0	1.00	
	2.2	2		42	1	1	1	1	1	5	0	0	1.00	
	3.2		22	42	1	1	1	1	1	5	0	0	1.00	
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	1.3	3	23		0	0	0	1	1	2	3	0	0.40	▪ คำถามข้อ 23 มีความคลุมเครืออาจจะใช้วัดสมรรถนะการสะท้อนและประเมิน
	2.3	3		43	1	0	1	1	1	4	1	0	0.80	▪
	3.3		23	43	0	0	0	1	1	2	3	0	0.40	▪
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	1.4	4	24		1	1	1	1	1	5	0	0	1.00	
	2.4	4		44	1	1	1	1	1	5	0	0	1.00	
	3.4		24	44	1	1	1	1	1	5	0	0	1.00	
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	1.5	5	25		1	1	1	1	1	5	0	0	1.00	
	2.5	5		45	1	1	1	1	1	5	0	0	1.00	
	3.5		25	45	1	1	1	1	1	5	0	0	1.00	
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	1.6	6	26		1	1	1	1	1	5	0	0	1.00	
	2.6	6		46	1	1	1	1	1	5	0	0	1.00	
	3.6		26	46	1	1	1	1	1	5	0	0	1.00	
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	1.7	7	27		1	1	1	1	1	5	0	0	1.00	
	2.7	7		47	1	1	1	1	1	5	0	0	1.00	
	3.7		27	47	1	1	1	1	1	5	0	0	1.00	
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	1.8	8	28		0	1	1	1	1	4	1	0	0.80	▪ ความยาวของข้อความและตัวเลือกมีขนาดต่างกัน ส่งผลต่อการใช้เวลาในการอ่านไม่เท่ากัน
	2.8	8		48	0	1	1	1	1	4	1	0	0.80	▪ ความยาวของข้อความและตัวเลือกมีขนาดต่างกัน ส่งผลต่อการใช้เวลาในการอ่านไม่เท่ากัน
	3.8		28	48	1	1	1	1	1	5	0	0	1.00	
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	1.9	9	29		0	1	1	1	1	4	1	0	0.80	▪ ความยาวของข้อความและตัวเลือกมีขนาดต่างกัน ส่งผลต่อการใช้เวลาในการอ่านไม่เท่ากัน

วัตถุประสงค์เชิงพฤติกรรม	คู่ที่	ความเป็นคู่ขนาน			คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		ฉบับที่หนึ่ง (ข้อที่)	ฉบับที่สอง (ข้อที่)	ฉบับที่สาม (ข้อที่)	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1		
	2.9	9		49	0	1	1	1	1	4	1	0	0.80	▪ ความยาวของข้อความและตัวเลือก1มีขนาดต่างกัน ส่งผลต่อการใช้เวลาในการอ่านไม่เท่ากัน
	3.9		29	49	1	1	1	1	1	5	0	0	1.00	▪
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	1.10	10	30		1	1	1	1	1	5	0	0	1.00	▪
	2.10	10		50	1	1	1	1	1	5	0	0	1.00	▪
	3.10		30	50	1	1	1	1	1	5	0	0	1.00	▪
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	1.11	11	31		-1	1	1	1	1	4	0	1	0.60	▪ ข้อ 11สามารถหาคำตอบได้เลยจากการค้นคืนสาระ แต่ข้อ 31ต้องวิเคราะห์เพิ่มเติมเพราะไม่มีบอกไว้อย่างชัดเจน
	2.11	11		51	-1	0	1	-1	1	2	1	2	0.00	▪ ข้อ 11 สามารถหาคำตอบได้เลยจากการค้นคืนสาระ แต่ข้อ 51 ต้องใช้การวิเคราะห์เพิ่มเติม ▪ ข้อ 51 อาจตอบจากความรู้เดิม ▪ ข้อ 51 เป็นความจำ
	3.11		31	51	1	0	1	-1	1	3	1	1	0.40	▪ ข้อ 31 และ 51 เป็นไปในทิศทางเดียวกัน ต้องมีภูมิรู้ว่าข้อมูลควรอยู่ในหมวดหมู่ใด ▪ ข้อ 51 อาจตอบจากความรู้เดิม
นักเรียนสามารถค้นหาและเลือกสารสนเทศที่เกี่ยวข้องได้	1.12	12	32		-1	1	1	1	1	4	0	1	0.60	▪ ความซับซ้อนต่างกัน ▪ ข้อ 12 สามารถนำข้อมูลมาใช้ตอบซึ่งมีข้อมูลที่บอกไว้ชัดเจน ▪ ข้อ 52 ไม่มีข้อมูลที่ชัดเจนบอก ต้องอาศัยทักษะการวิเคราะห์เพิ่มเติม
	2.12	12		52	0	1	-1	-1	1	2	1	2	0.00	▪ ข้อ 12 และข้อ 52 สามารถหาคำตอบได้จากบทอ่านแม้เป็นคนละลักษณะ
	3.12		32	52	-1	1	-1	0	1	2	1	2	0.00	▪ ข้อ 32 จะต้องอาศัยทักษะความเข้าใจวิเคราะห์เพิ่มเติม นอกเหนือจากการค้นคืน ▪ ข้อ 52 สามารถหาคำตอบได้จากการค้นคืน
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	1.13	13	33		1	1	1	0	1	4	1	0	0.80	▪ ปรับเฉลยข้อที่ 13 ▪ 13 เป็นความจำ 33 เข้าใจ
	2.13	13		53	1	1	1	1	1	5	0	0	1.00	▪ เป็นความจำทั้งคู่
	3.13		33	53	1	1	1	0	1	4	1	0	0.80	▪ 33 เป็นความจำ 34 เข้าใจ
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	1.14	14	34		1	1	1	1	1	5	0	0	1.00	▪ เปลี่ยนเฉลยข้อที่ 34
	2.14	14		54	1	1	1	1	1	5	0	0	1.00	▪
	3.14		34	54	1	1	1	1	1	5	0	0	1.00	▪
นักเรียนสามารถประเมินคุณภาพและความน่าเชื่อถือของเนื้อเรื่องได้	1.15	15	35		0	1	1	0	-1	2	2	1	0.20	▪ ข้อ 15 มีการประเมินพร้อมบอกเหตุผล แต่อีก 2 ข้อเป็นเพียงเหตุผล

วัตถุประสงค์เชิงพฤติกรรม	คู่ที่	ความเป็นคู่ขนาน			คะแนนของผู้เชี่ยวชาญ					คะแนนรวม			คะแนนเฉลี่ย (IOC)	ข้อเสนอแนะ
		ฉบับที่หนึ่ง (ข้อที่)	ฉบับที่สอง (ข้อที่)	ฉบับที่สาม (ข้อที่)	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	+1	0	-1		
														ประกอบ
	2.15	15		55	0	1	1	0	-1	2	2	1	0.20	
	3.15		35	55	1	1	1	0	1	4	1	0	0.80	
นักเรียนสามารถเข้าถึงและค้นคืนสารสนเทศที่อยู่ในเนื้อเรื่อง	1.16	16	36		0	0	1	1	1	3	2	0	0.60	▪ ตอบผิดครั้งที่ 2 ข้อที่ 3 โดยส่วนตัวคิดว่าข้อนี้ไม่ชัดเจน
	2.16	16		56	0	0	1	1	1	3	2	0	0.60	▪ เปลี่ยนคำจาก “กิจกรรมหลัก” เป็น “เป้าหมายหลัก”
	3.16		36	56	1	0	1	1	1	4	1	0	0.80	
นักเรียนสามารถเข้าใจความหมายที่แท้จริงของเนื้อเรื่อง	1.17	17	37		1	1	1	1	1	5	0	0	1.00	
	2.17	17		57	1	1	1	1	1	5	0	0	1.00	
	3.17		37	57	1	1	1	1	1	5	0	0	1.00	
นักเรียนสามารถบูรณาการและลงข้อสรุปจากสารสนเทศหลาย ๆ ส่วนที่อยู่ในเนื้อเรื่อง	1.18	18	38		1	1	1	1	1	5	0	0	1.00	
	2.18	18		58	1	1	1	1	1	5	0	0	1.00	
	3.18		38	58	1	1	1	1	1	5	0	0	1.00	
นักเรียนสามารถสะท้อนความคิดเห็นเกี่ยวกับเนื้อหาและรูปแบบของเรื่องที่ได้อ่านอย่างมีวิจารณญาณ	1.19	19	39		1	1	1	1	1	5	0	0	1.00	
	2.19	19		59	1	1	1	1	1	5	0	0	1.00	
	3.19		39	59	1	1	1	1	1	5	0	0	1.00	
นักเรียนสามารถตรวจสอบข้อมูลที่ขัดแย้งกันและหาวิธีจัดการข้อขัดแย้งนั้น	1.20	20	40		1	1	1	1	1	5	0	0	1.00	▪ คำถามข้อ 20 และ 40 มีความคู่ขนานกัน แต่คำตอบอ่านแล้วไม่เข้าใจ ▪ งงคำถาม
	2.20	20		60	1	1	1	1	1	5	0	0	1.00	▪
	3.20		40	60	1	1	1	1	1	5	0	0	1.00	▪

Appendix F

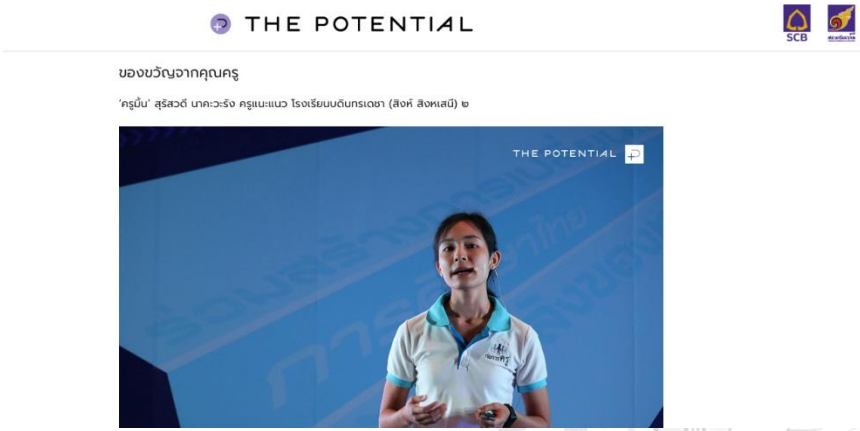
Example of Reading Literacy for Computerized Dynamic Assessment

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตัวอย่างข้อสอบการรู้เรื่องการอ่าน ฉบับที่หนึ่ง

บทอ่านที่หนึ่ง: ของขวัญจากคุณครู

สถานการณ์: เด็กหญิงชมพู่ได้อ่านเรื่องราวของคุณครูมิ้นซึ่งมาพูดบทเวที TEP Forum 2019 ใน หัวข้อ 'ภาพใหม่การศึกษาไทยเพื่อการสร้างเสริมสมรรถนะเด็กไทย'



วันนี้อยากพาทุกคนมารับฟังเรื่องราวของเด็กคนหนึ่ง เป็นเด็กผู้หญิงที่มีพฤติกรรมชอบส่งงานช้า งานที่ส่งไม่ค่อยมีคุณภาพ ดูเหมือนจะไม่ค่อยตั้งใจเรียน ชอบคุยเล่นกับเพื่อนบ้าง มีความเสี่ยงที่เขาอาจจะเรียนไม่ผ่านวิชานี้ เพราะเขาส่งงานที่ไม่ค่อยดีเท่าไหร่ ในความคิดของคุณครูทุกท่าน เด็กคนนี้เป็นยังไงบ้าง อาจจะรู้สึกว่าเธอขี้เกียจจังเลย เธอไม่น่ารักจังเลย หรืออาจจะมองเขาว่าเขาไม่มีความรับผิดชอบ หลายท่านก็น่าจะคิดแบบนี้ แต่มีคุณครูอยู่ท่านหนึ่งที่ไม่ได้คิดแบบนั้น เขาคิดว่าน่าจะลองเข้าไปคุยสอบถามกับเด็กคนนี้ดูก่อน จนได้คำตอบว่าเด็กคนนี้รู้สึกแย่ รู้สึกอายทุกครั้งที่เขาทำงานที่ครูมอบหมายให้ออกมาไม่ดี กลัวคุณครูมองว่าตัวเขา เป็นเด็กที่ไม่ฉลาด ไม่เก่ง แม้เขาจะคิดแล้วว่าจะทำอย่างไรต่อกับงานที่ครูสั่งมา ทำอย่างไรให้งานออกมาดีที่สุดแล้วก็ตาม มัวแต่คิด ๆ แล้วก็คิด ๆ จนหมดเวลาและไม่ได้ส่งงาน ก็ต้องเร่งทำให้เสร็จ ทำให้งานไม่มีคุณภาพ สิ่งที่เกิดขึ้นมันมีแต่ความเครียด ความกดดัน ความคาดหวังในตัวเอง อยากเป็นเด็กที่ทำให้ครูประทับใจ พอครูรู้เช่นนี้ ครูคิดแล้วว่าเด็กคนนี้ต้องการความช่วยเหลือ ครูจึงปลุกปั้นพลัง ให้กำลังใจ และพูดเพียงประโยคเดียวว่า ครูเชื่อว่าเราทำได้ จากเด็กที่ดูจะเรียนไม่จบพร้อมเพื่อน ดูทำงานไม่ทัน สุดท้ายเขาก็เรียนผ่านมาได้ แถมยังจบไปพร้อมกันกับเพื่อนอีกด้วย

มันก็เลยทำให้มิ้นคิดอย่างหนึ่งว่า สิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน ทำให้เด็กที่คิดว่าชีวิตเขามีแต่ความมืด ถูกตัดสินว่าเป็นเด็กที่ไม่ดี เพียงแค่ครูคนหนึ่งเชื่อมั่นในตัวเขา มันช่วยปลุกปั้นแสงสว่างในตัวเขาขึ้นมา ถ้าคิดในทางกลับกัน เด็กคนนี้ไม่ได้เจอกับครูคนนี้ ชีวิตเขาจะเป็นอย่างไร?
มิ้นคิดว่าเด็กคนนั้นคงรู้สึกไร้ค่า รู้สึกว่าตัวเองไม่สามารถทำอะไรได้อีกต่อไป

ทุกคนอาจจะคิดว่ามิ้นเป็นครูคนนั้นใช่ไหมหรือเปล่า – แต่ไม่ใช่ มิ้นเป็นนักเรียนคนนั้น มิ้นเคยเป็นนักเรียนที่ล้มเหลวเคยถูกมองว่าเป็นคนที่ไม่มีความรับผิดชอบ เป็นนักเรียนที่ดูไม่ตั้งใจเรียน แต่เพราะของขวัญที่ครูท่านนั้นยื่นมาให้ ทำให้มิ้นรู้ว่าเรายังมีคุณค่าอะไรบางอย่างอยู่ ของขวัญชิ้นนี้ทำให้เราเรียนรู้ความผิดพลาดของเรา และทำให้มิ้นรู้ว่าถ้ามิ้นไปสอนเด็กคนไหนก็ตาม ไม่ว่าเขาจะเรียนเก่ง/ไม่เก่ง เป็นเด็กดี/ไม่ดี ในสายตาของใคร สิ่งที่มิ้นจะทำกับเขา คือการหยิบยื่นโอกาสมองเห็นแสงสว่างในตาของเขา และเชื่อมั่นในตัวเขา "เพราะกว่าเด็กหนึ่งคนจะเติบโตขึ้นมา ไม่มีใครหรอกค่ะที่ไม่เคยผิดพลาด ทุกคนล้วนผ่านการผิดพลาดมาแล้ว แต่จะมีครูสักกี่คนที่ให้โอกาสกับความผิดพลาดของ เด็ก ๆ"

"มันเป็นเรื่องน่ากลัว ที่เด็ก ๆ ในโรงเรียนกลัวที่จะผิด กลัวที่จะพลาด เขาอยากเป็นที่หนึ่ง แม้ในโรงเรียนจะมีความหลากหลายมากมาย แต่มีเพียงสองเครื่องหมายเท่านั้นคือถูกกับผิด ซึ่งมันไปปิดกั้นความอยากเรียนรู้ของนักเรียน"

ที่มา: The Potential (2019)

1.1 เพราะเหตุใดเด็กผู้หญิงในเรื่องนี้จึงมีพฤติกรรมส่งงานช้า

ก. เป็นเด็กขี้เกียจ

ข. ไม่มีความรับผิดชอบต่องาน

ค. รู้สึกแย่ที่ทำงานออกมาไม่ดี

ง. ไม่ค่อยตั้งใจเรียนเวลาครูสอน

จ. คิดว่าคงเรียนไม่ผ่าน จึงไม่สนใจทำงานส่ง

1.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะเธอไม่มีเวลาทำงานที่ครูมอบหมายได้ดี

ข. เพราะเธอคิดว่าตนเองไม่สามารถทำงานได้ดีเหมือนเพื่อนคนอื่น ๆ ในห้องเรียน

ค. เพราะมีคุณครูท่านหนึ่งเข้าไปสอบถามและพบว่าเธอมีความเครียดและกดดันในการทำงานส่งครู

2.1 จากข้อความ ***"สิ่งเหล่านี้มันเหมือนของขวัญที่ครูให้กับนักเรียน"*** มีความหมายว่าอย่างไร

ก. ครูได้มอบของขวัญให้กับนักเรียน

ข. ครูให้กำลังใจและเชื่อมั่นในตัวของนักเรียน

ค. ครูให้โอกาสนักเรียนส่งงานช้ากว่ากำหนดได้

ง. ครูทำให้นักเรียนไม่เครียดเวลาที่สอนในชั้นเรียน

จ. ครูเปิดโอกาสให้นักเรียนได้ทำในสิ่งที่ตนเองสนใจ

2.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะครูทำหน้าที่เป็นผู้ช่วยเหลือให้แก่นักเรียนเมื่อมีปัญหาเดือดร้อน

ข. เพราะครูเป็นผู้ที่ให้โอกาสแก่นักเรียนยามนักเรียนทำผิดพลาด

ค. เพราะครูเป็นผู้ที่รักและคอยดูแลนักเรียนเหมือนสมาชิกในครอบครัว

3.1 บทความเรื่องนี้ ครูมิ้นรู้สึกอย่างไรกับนักเรียนที่ล้มเหลว

ก. เหนื่อยใจ

ข. เห็นอกเห็นใจ

ค. หงุดหงิดใจ

ง. ทุกข์ใจ

จ. เศร้าใจ

3.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะครูมิ้นเคยเป็นนักเรียนที่ล้มเหลวมาก่อน

ข. เพราะครูมิ้นรู้สึกสงสารลูกศิษย์ที่เป็นนักเรียนที่ล้มเหลว

ค. เพราะครูมิ้นเคยมีลูกศิษย์ที่ล้มเหลวมาก่อน

4.1 เพราะเหตุใดผู้พูดจึงเน้นย้ำว่า ***"แม้ในโรงเรียนจะมีความหลากหลายมากมาย แต่มีเพียงสองเครื่องหมายเท่านั้นคือถูกกับผิด ซึ่งมันไปปิดกั้นความอยากเรียนรู้ของนักเรียน"***

ก. เพื่อแสดงความคิดเห็นว่านักเรียนควรได้รับโอกาสในการเรียนรู้จากการลองผิดลองถูกด้วยตนเอง

ข. เพื่อเปรียบเทียบความแตกต่างระหว่างสิ่งที่นักเรียนทำถูกและสิ่งที่นักเรียนทำผิด

ค. เพื่อเสนอแนะวิธีการเรียนรู้ที่ถูกต้องให้แก่นักเรียนนำไปปฏิบัติ

ง. เพื่อสนับสนุนครูในการส่งเสริมนักเรียนให้เรียนรู้ได้ด้วยตนเอง

จ. เพื่อเป็นกำลังใจให้แก่คุณครูในการเรียนรู้จากความผิดพลาดของตนเอง

4.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะผู้พูดไม่ต้องการให้นักเรียนเป็นเด็กไม่ดี

ข. เพราะผู้พูดต้องการสะท้อนปัญหาของเด็กนักเรียนที่ล้มเหลวในโรงเรียน

ค. เพราะผู้พูดต้องการให้โอกาสนักเรียนในการเรียนรู้จากความผิดพลาดของตน

5.1 นักเรียนคิดว่าบทความเรื่องนี้ควรปรับปรุงด้านใดเพื่อให้มีความน่าเชื่อถือมากยิ่งขึ้น

ก. ควรระบุวันเดือนปีที่เผยแพร่บทความ

ข. ควรระบุประวัติของผู้ให้ข้อมูล

ค. ควรมีการระบุชื่อของเว็บไซต์ที่เผยแพร่บทความ

ง. ควรมีการอ้างอิงแหล่งที่มาของข้อมูล

จ. ควรให้ที่อยู่อีเมลผู้เขียนเพื่อที่ผู้อ่านสามารถติชมได้

5.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะการระบุวันเดือนปีที่เผยแพร่ทำให้อ้างอิงบทความได้

ข. เพราะการระบุวันเดือนปีที่เผยแพร่ทำให้เข้าใจบริบทและสถานการณ์ในอดีตมากขึ้น

ค. เพราะการระบุวันเดือนปีที่เผยแพร่ทำให้รู้ว่าบทความมีความทันสมัยมากน้อยเพียงใด

CHULALONGKORN UNIVERSITY จุฬาลงกรณ์มหาวิทยาลัย

บทอ่านที่สอง: โซเชียลมีเดีย

สื่อสารกับคนวัยไหน ใช้มีเดียใดถึงเหมาะ?

ในปี 2018 ที่ผ่านมา ก็ถือว่าเป็นอีกหนึ่งปีที่ตลาดของผู้ใช้อินเทอร์เน็ตเติบโตไปอย่างก้าวกระโดด ซึ่งปัจจุบัน **สถิติการใช้อินเทอร์เน็ตของคนไทยมีอัตราเฉลี่ยอยู่ที่วันละ 10 ชั่วโมง 5 นาที/ต่อคน** จึงทำให้หลาย ๆ ธุรกิจมีการปรับตัวไปตามพฤติกรรมผู้บริโภค และยังมีธุรกิจออนไลน์ใหม่ ๆ ผุดขึ้นอีกเพียบ

หลายแบรนด์ หลายธุรกิจ ก็เลือกเอา **"คอนเทนต์" (Content)** มาเป็นปัจจัยหลักในการสื่อสารกับผู้บริโภค เพื่อเข้าถึงผู้คน และให้คนมาร่วมมีปฏิสัมพันธ์กับแบรนด์ได้มากขึ้น แต่การเข้าถึงกลุ่มเป้าหมายถูกที่ ถูกเวลา ก็เป็นอีกหนึ่งกลยุทธ์ที่สำคัญไม่แพ้กัน

จากสถิติการใช้อินเทอร์เน็ตของคนไทยปี 2018 พบว่า...

รูปที่ 1 สื่อสารกับคนวัยไหน ใช้มีเดียใดถึงเหมาะ?

ที่มา: Creative Talk (2019)

สถิติผู้ใช้ LINE JOBS

LINE JOBS ถือว่าเป็นแพลตฟอร์มใหม่ล่าสุดของ LINE แต่ก็มีจำนวนผู้ใช้เพิ่มขึ้นอย่างรวดเร็ว โดยจะสังเกตได้ว่าผู้ใช้ส่วนใหญ่ก็ยังเป็นกลุ่ม Millennial ที่มีความคุ้นชินกับเทคโนโลยีสูง ที่ต้องการใช้ช่องทางค้นหางานใหม่ ๆ นอกเหนือจากการค้นหาผ่านเว็บไซต์ตรงของบริษัท หรือแพลตฟอร์มกลางในการสมัครงาน ซึ่งเว็บไซต์บางเว็บจะคิดค่าบริการในการหางาน

รูปที่ 2 สถิติผู้ใช้ LINE JOBS

ที่มา thumbsup (2019)

6.1 จากรูปภาพที่ 1 “สื่อสารกับคนวัยไหน ใช้มีเดียใดถึงเหมาะ?” คนในกลุ่มใดบ้างที่ใช้ น้อยที่สุด

ก. Gen Z และ Gen Y

ข. Gen Y และ Gen X

ค. Gen X และ Baby boomer

ง. Baby boomer และ Gen Y

จ. Gen Z และ Gen X

6.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะกลุ่มวัยดังกล่าวไม่นิยมเล่น

ข. เพราะ เป็นโซเชียลมีเดียที่ไม่เหมาะกับกลุ่มวัยดังกล่าว

ค. เพราะเส้นสีฟ้ามีผู้ใช้ 19% และเส้นสีแดงมีผู้ใช้ 4%

7.1 มีเดียใดที่ได้รับความนิยมสูงสุดในทุกกลุ่มวัย

ก.

ข.

ค.

ง.

จ.

7.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะมีเดียดังกล่าวเป็นโซเชียลมีเดียที่ใช้งานได้ง่ายที่สุด

ข. เพราะมีเดียดังกล่าวเป็นโซเชียลมีเดียที่มีการใช้งานที่หลากหลายในทุกช่วงวัย

ค. เพราะเมื่อดูจำนวนผู้ใช้โซเชียลมีเดียในทุกกลุ่มวัยพบว่าในทุกกลุ่มวัยมีผู้ใช้มีเดียดังกล่าวมากที่สุด

8.1 จากข้อความว่า ***"LINE JOBS ถือว่าเป็นแพลตฟอร์มใหม่ล่าสุดของ LINE แต่ก็มีจำนวนผู้ใช้เพิ่มขึ้นอย่างรวดเร็ว"*** มีความหมายว่าอย่างไร

ก. LINE JOBS เป็นแพลตฟอร์มใหม่ที่ได้รับความนิยมสูงสุดของ LINE

ข. LINE JOBS มีจำนวนผู้ใช้มากที่สุดจากแพลตฟอร์มทั้งหมดของ LINE

ค. แม้ว่า LINE JOBS เพิ่งเปิดตัวได้ไม่นานก็เป็นแพลตฟอร์มที่กำลังได้รับความนิยม

ง. การทำการตลาดอย่างต่อเนื่องของ LINE ยังไม่ช่วยให้ LINE JOBS ได้รับความนิยม

จ. แม้ว่า LINE JOBS เป็นแพลตฟอร์มที่มีความทันสมัยแต่มีผู้ใช้งานจำนวนไม่มากนัก

8.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะผู้ใช้งานของ LINE ส่วนใหญ่สนใจค้นหางานที่ LINE JOBS

ข. เพราะ LINE ใช้ช่องทางที่หลากหลายในการประชาสัมพันธ์ LINE JOBS

ค. เพราะบทอ่านกล่าวถึงผู้ใช้ LINE JOBS ที่มีจำนวนผู้ใช้เพิ่มมากขึ้นอย่างรวดเร็ว

9.1 บุคคลใดมีแนวโน้มใช้ LINE JOBS ในการค้นหางานมากที่สุด

ก. นายมานพ อายุ 18 ปี มีวุฒิการศึกษาชั้นม.6 ต้องการหางานพนักงานฝ่ายผลิตในโรงงานผลิตรถยนต์

ข. นางสาวชูใจ อายุ 20ปี มีวุฒิการศึกษา ปวส. ต้องการหางานแผนกธุรการ

ค. นางสาวสมร อายุ 23 ปี มีวุฒิการศึกษาปริญญาตรี ต้องการหางานทางด้านบัญชี

ง. นายธนัช อายุ 40 ปี มีวุฒิการศึกษาปริญญาโท ต้องการหางานหัวหน้าแผนกส่งออกนำเข้า

จ. นางอมร อายุ 50 ปี มีวุฒิการศึกษาชั้นม.3 ต้องการหางานรับจ้างทั่วไป

9.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะผู้ที่สำเร็จการศึกษาในระดับดังกล่าวเป็นกลุ่มที่มีอัตราการว่างงานสูงสุด

ข. เพราะผู้ใช้ LINE JOBS ส่วนใหญ่สำเร็จการศึกษาในระดับดังกล่าว และมีช่วงอายุในระดับนั้น

ค. เพราะจำนวนผู้สำเร็จการศึกษาในระดับดังกล่าวมีมากที่สุดในตลาดแรงงาน

10.1

> ใครที่มีภาพจำว่าผู้สูงอายุคือกลุ่มคนที่ไม่ชอบเทคโนโลยี เข้าใจเทคโนโลยียาก ไม่ชอบเรียนรู้สิ่งใหม่ๆ บอกเลยว่ายุคปัจจุบันไม่ใช่อีกแล้วนะคะ ในหมู่นักการตลาดกว่า 79% ยังใช้ **"อายุ"** เป็นตัวชี้วัดในการกำหนดตัวตนของผู้สูงอายุ แทนที่จะศึกษาจากความชอบ ไลฟ์สไตล์ หรือความต้องการที่แท้จริง
>
> 88% ของผู้บริโภคเห็นด้วยว่าแบรนด์ควรโฟกัสความต้องการ - ความสนใจของผู้บริโภค ไม่ใช่ดูกันที่อายุอย่างเดียว
>
> ปรับมาจาก Clisk (2019)

ข้อความข้างต้นมีความขัดแย้งกับรูปภาพที่ 1 ที่กล่าวถึงกลยุทธ์การสื่อสารของนักการตลาด ข้อใดเป็นประเด็นความขัดแย้งที่ผู้เขียนต้องการนำเสนอ

ก. นักการตลาดไม่ได้เน้นที่ความสนใจที่แท้จริงของผู้บริโภค แต่ใช้อายุในการกำหนดกลยุทธ์ทางการตลาด

ข. นักการตลาดสนใจผู้บริโภคในกลุ่มวัยอื่น ๆ มากกว่ากลุ่มวัยผู้สูงอายุ

ค. นักการตลาดสนใจผู้สูงอายุที่ชอบใช้เทคโนโลยี มากกว่ากลุ่มที่ไม่ชอบใช้เทคโนโลยี

ง. แม้ผู้สูงอายุจะมีความชอบที่แตกต่างกัน แต่ผู้สูงอายุจำเป็นต้องเรียนรู้การใช้โซเชียลมีเดียด้วย

จ. นักการตลาดชอบใช้โซเชียลมีเดียในการสื่อสารกับผู้สูงอายุมากกว่าช่องทางสื่อสารอื่น ๆ

10.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

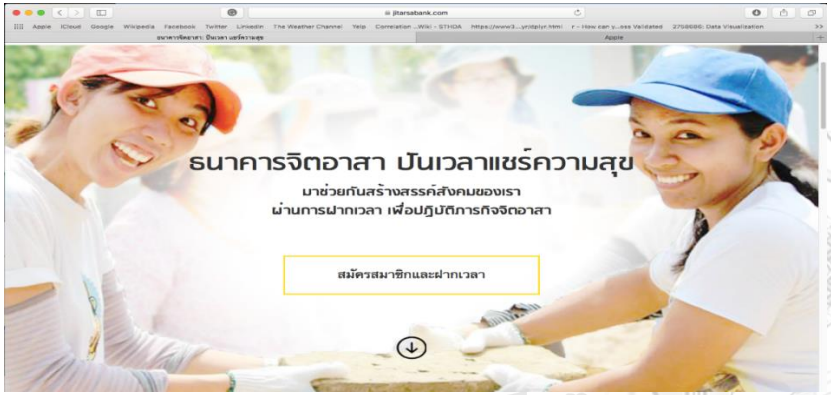
ก. เพราะกลุ่มผู้สูงอายุจะมีความชอบและความสนใจที่ใกล้เคียงกัน

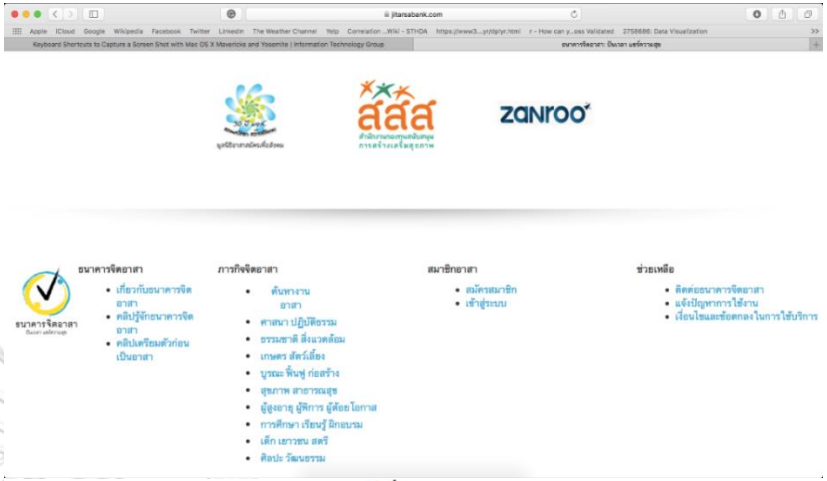
ข. เพราะการกำหนดผู้บริโภคตามกลุ่มอายุทำได้ง่ายและชัดเจนกว่าการแบ่งตามความชอบหรือความสนใจ

ค. เพราะผู้เขียนกล่าวว่านักการตลาดยังใช้อายุเป็นตัวชี้วัดผู้สูงอายุ แทนที่จะศึกษาจากความสนใจหรือความชอบ

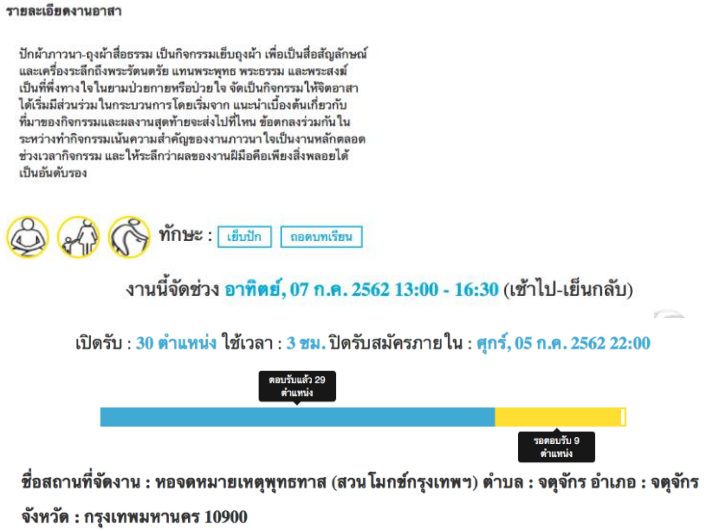
บทอ่านที่สาม: ธนาคารจิตอาสา

สถานการณ์: เด็กชายพงษ์ต้องการใช้เวลาว่างช่วงปิดเทอมให้เป็นประโยชน์ จึงปรึกษาคุณครูที่โรงเรียน คุณครูจึงแนะนำเด็กชายพงษ์ให้เข้าเว็บธนาคารจิตอาสาและลองสมัครเป็นจิตอาสาให้งานที่สนใจดู





ภาวนาวิถีเมือง ที่ชวนคุณค้นหาความสุขสงบภายในใจ ใช้เวลาที่มีอยู่อย่างคุ้มค่า ผ่านการภาวนากับการปักผ้า ฝึกสติอยู่กับปัจจุบันขณะทุกฝีเข็ม เส้นด้ายกับถุงผ้า อย่างต่อเนื่องเคลื่อนไหว ฝึกการเรียนรู้และสังเกตกายและใจระหว่างการปักผ้า รู้ เห็น สิ่งที่กระทบรอบตัว ตามความเป็นจริง เพื่อความเข้าใจชีวิต เปิดรับความสุขด้วยการปักผ้า และเป็นผู้ให้ถุงผ้าสื่อธรรมมอบเป็นขวัญกำลังใจแก่ผู้ป่วย เด็ก ผู้สูงอายุ และจิตอาสาผู้บำเพ็ญประโยชน์เพื่อสังคมต่อไป



ที่มา: ธนาคารจิตอาสา (2019)

11.1 ใครคือกลุ่มเป้าหมายหลักที่ได้รับถุงผ้าสื่อธรรม

ก. ทหาร

ข. พระบวชใหม่

ค. ผู้ป่วย

ง. นักโทษ

จ. ผู้บำบัดยาเสพติด

11.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะถุงผ้ามีคุณค่าทางจิตใจสำหรับกลุ่มผู้ป่วย

ข. เพราะกลุ่มผู้ป่วยนิยมใช้ถุงผ้าในการใส่ยารักษาโรคแทนการใช้ถุงพลาสติก

ค. เพราะในประกาศกล่าวว่าถุงผ้าจะถูกมอบเพื่อเป็นกำลังใจให้แก่กลุ่มผู้ป่วย

12.1 มีผู้สนใจเข้าร่วมโครงการอาสา “ถุงผ้าสื่อธรรม” จำนวนทั้งหมดกี่คน

ก. 9 คน

ข. 29 คน

ค. 30 คน

ง. 38 คน

จ. ไม่ปรากฏจำนวนผู้สนใจร่วมโครงการ

12.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะมีผู้สนใจเข้าร่วมกิจกรรมมากกว่าจำนวนที่เปิดรับ

ข. เพราะเป็นจำนวนตัวเลขของผู้สนใจเข้าร่วมที่เหมาะสมในการทำกิจกรรม

ค. เพราะพิจารณาจากจำนวนของผู้ที่ตอบรับโครงการ และผู้ที่รอตอบรับ

13.1 ข้อใดคือวัตถุประสงค์ที่สำคัญที่สุดของกิจกรรมโครงการอาสา “ถุงผ้าสื่อธรรม”

ก. เพื่อใช้เวลาว่างให้เกิดประโยชน์

ข. เพื่อทำถุงผ้าเป็นของที่ระลึกให้แก่ตนเอง

ค. เพื่อฝึกการเรียนรู้ในการปักผ้าและเย็บถุงผ้าที่ถูกต้อง

ง. เพื่อร่วมกิจกรรมในการฝึกสติและเจริญภาวนาในการปักผ้า

จ. เพื่อเรียนรู้การทำงานร่วมกับผู้อื่น

13.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะเป็นกิจกรรมที่เน้นความสำคัญของการภาวนาใจเป็นหลัก

ข. เพราะเป็นกิจกรรมที่เน้นการสอนหลักธรรมคำสอนทางศาสนาให้แก่ผู้ที่เข้าร่วม

ค. เพราะเป็นกิจกรรมที่ต้องการให้ผู้ที่เข้าร่วมปักผ้าเป็น

14.1 บุคคลใดเหมาะสมที่จะเข้าร่วมกิจกรรมงานอาสา “ถุงผ้าสื่อธรรม” มากที่สุด

ก. สายใจต้องการเรียนรู้วิธีการปักผ้าและเย็บถุงผ้า

ข. สายสมรต้องการใช้เวลาว่างให้เป็นประโยชน์

ค.สมใจต้องการรู้จักเพื่อนใหม่และทำกิจกรรมร่วมกับผู้อื่น

ง. สมศรีต้องการถุงผ้ามามอบให้แก่ญาติผู้ใหญ่

จ. สมฤดีต้องการฝึกเจริญสติและสมาธิด้วยการปักผ้า

14.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะกิจกรรมถุงผ้าสื่อธรรมต้องการให้ผู้เข้าร่วมได้ทำถุงผ้าที่ประณีต

ข. เพราะกิจกรรมถุงผ้าสื่อธรรมต้องการให้ผู้เข้าร่วมมีความระมัดระวังในการปักผ้า

ค. เพราะกิจกรรมถุงผ้าสื่อธรรมต้องการให้ผู้เข้าร่วมได้ฝึกนำธรรมะมาใช้ในการปักผ้า

15.1 นักเรียนคิดว่าประกาศของกิจกรรม “ถุงผ้าสื่อธรรม” มีความน่าเชื่อถือมากที่สุดในประเด็นใด

ก. เป็นกิจกรรมที่ครูที่โรงเรียนเป็นผู้แนะนำให้เข้าร่วม

ข. ประชาสัมพันธ์ทางเว็บไซค์ที่มีผู้เข้าเยี่ยมชมเป็นจำนวนมาก

ค. มีผู้สมัครเป็นจิตอาสาจำนวนมากเกินกว่าที่กำหนด

ง. ไม่มีค่าใช้จ่ายในการเข้าร่วมกิจกรรม

จ. มีรายละเอียดของวัน เวลา และสถานที่จัดกิจกรรมที่แสดงไว้อย่างชัดเจน

15.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

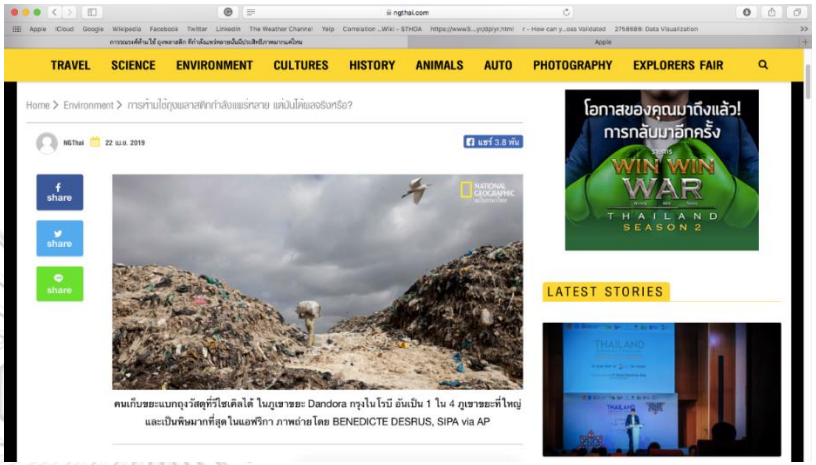
ก. เพราะจิตอาสาตัดสินใจเข้าร่วมกิจกรรมจากรายละเอียดในประกาศ

ข. เพราะความน่าเชื่อถือพิจารณาได้จากรายละเอียดของข้อมูลในประกาศ

ค. เพราะการประชาสัมพันธ์ในเว็บธนาคารจิตอาสาต้องมีการกำหนดรายละเอียดของวัน เวลา และสถานที่ที่ชัดเจน

บทอ่านที่สี่: การห้ามใช้ถุงพลาสติก

สถานการณ์: คุณครูในนักเรียนเขียนเรียงความเกี่ยวกับขยะพลาสติก เด็กหญิงน้ำใจจึงหาข้อมูลทางอินเทอร์เน็ตและพบประเด็นที่น่าสนใจเกี่ยวกับการงดใช้ถุงพลาสติก



การห้ามใช้ถุงพลาสติกกำลังแพร่หลาย แต่มันได้ผลจริงหรือ?

ในนาคูรู ประเทศเคนยา ถุงพลาสติก หูหิ้วได้หายไปจากตลาด Wakulima เนื่องจากการห้ามใช้ถุงพลาสติกในประเทศก็จริง แต่มันกลับถูกแทนที่ด้วยถุงโพลีโพรพีลีน (Polypropelene) ซึ่งเป็นถุงพลาสติกอีกประเภท แต่ James Wakaiba นักกิจกรรมผู้เป็นตัวแทนของการรณรงค์ที่นำไปสู่การห้ามใช้ถุงพลาสติก กล่าวว่าการงดใช้ที่ไม่สมบูรณ์แบบยังดีกว่าการไม่ทำอะไรเลย

“โอเค ถุงพวกนี้ทำจากโพลีโพรพิลีนก็จริง แต่พวกมันถูกนำมาใช้ใหม่ได้ และไม่ได้เป็นถุงบางๆ ที่ปลิวได้ง่าย” เขากล่าว “สหประชาชาติ (UN) บอกว่าแค่ปีเดียว คนเคนยาก็ใช้ถุงพลาสติกในซูเปอร์มาร์เก็ตตั้ง 100 ล้านใบ แสดงว่าเราก็ประหยัดถุงไปได้ 100 ล้านใบแล้ว ผมคิดว่าความสำเร็จของเรื่องนี้มีประมาณร้อยละ 80”

ถุงพลาสติกซึ่งมักถูกมองว่าเป็นทั้งสินค้าขายดีที่สุดของโลกและเป็นสิ่งที่พบได้ทุกหนทุกแห่ง กลายเป็นหนึ่งในสินค้าที่ถูกห้ามใช้มากที่สุดในโลก จากข้อมูลของ UN เมื่อเดือนกรกฎาคมปีที่แล้ว กว่า 127 ประเทศมีการห้าม

ใช้หรือเก็บภาษีถุงพลาสติก นอกจากนี้ การควบคุมการใช้ผลิตภัณฑ์ชนิดนี้ยังแพร่กระจายอย่างรวดเร็ว โดยเฉพาะในระดับท้องถิ่น แม้แต่กลุ่มก่อการร้ายที่ได้รับการสนับสนุนจากอัลกออิดะห์ยังเข้าร่วมการงดใช้ถุงพลาสติกด้วยเหตุผลว่า

“[ถุงช้อปปิ้งพลาสติก] เป็นภัยอันตรายต่อทั้งมนุษย์และสัตว์”

เมื่อเร็ว ๆ นี้ สหภาพยุโรปเริ่มมีการเคลื่อนไหวเพื่องดใช้ถุงพลาสติก โดยเป็นส่วนหนึ่งของการเก็บกวาดผลิตภัณฑ์พลาสติกที่พบเห็นได้บ่อยที่สุดบนชายหาดในยุโรป ส่วนในสหรัฐ ฮาวายห้ามใช้ถุงพลาสติกอย่างไม่เป็นทางการ เนื่องจากมีการงดใช้อยู่ในทุกเทศมณฑล (County) และนิวยอร์กเป็นรัฐที่สองที่สั่งห้ามถุงพลาสติก (ต่อจากแคลิฟอร์เนีย) นอกจากนี้ยังมีการเสนอร่างกฎหมายเกี่ยวกับพลาสติกกว่า 95 ครั้ง ในปลายปีที่ผ่านมา ซึ่งเป็นสถิติที่มากกว่าปีอื่น ๆ

กระนั้นก็ตาม ในขณะที่การห้ามใช้ถุงพลาสติกเกิดขึ้นทั่วโลก ประสิทธิผลของมันกลับยังไม่แน่ชัด แม้ว่า Wakibia จะชื่นชมการกระทำดังกล่าวของเคนยาก็ตาม การงดใช้ถุงพลาสติกก่อให้เกิดการงดใช้ผลิตภัณฑ์พลาสติกแบบใช้ครั้งเดียวทิ้ง (Single-use Plastics) ประเภทอื่น ๆ รวมถึงจาน ถ้วย ช้อนส้อม หลอด และขวด ซึ่งพลาสติกประเภทที่ว่านี้มีจำนวนร้อยละ 40 ของพลาสติกที่มีการผลิตทั่วโลก แต่ยังคงมีคำถามว่าสิ่งนี้จะสามารถช่วยลดขยะพลาสติกที่รั่วไหลลงในมหาสมุทรเป็นจำนวนเฉลี่ยกว่า 8 ล้านตันต่อปีได้มากเพียงใด โดยเฉพาะเมื่อมีการคาดการณ์ว่าการผลิตวัสดุดังกล่าวจะเพิ่มขึ้นกว่าสองเท่าในราวปี 2040 และอาจเป็นส่วนประกอบกว่าร้อยละ 20 จากการผลิตน้ำมันของโลก ในราวปี 2050

แม้แต่ผู้ที่สนับสนุนการงดใช้อย่างแข็งขันก็รับทราบถึงข้อจำกัดเหล่านี้

“แม้การห้ามใช้ถุงพลาสติกแบบใช้ครั้งเดียวทิ้งจะมีความสำคัญเพียงใดในการลดขยะพลาสติก มันก็ยังเปลี่ยนโลกไม่ได้” มาร์ค เมอเรย์ (Mark Murray) ผู้อำนวยการบริหารของ Californians Against Waste กล่าว “พูดกันตามตรง ประเด็นหลักของเรื่องนี้คือการส่งสารถึงผู้กำหนดนโยบาย สาธารณชน และอุตสาหกรรม ว่าเราจะทำอะไรบางอย่างอย่างจริงจังเพื่อลดจำนวนบรรจุภัณฑ์พลาสติก และถ้าพวกคุณไม่รู้ว่าจะทำมันยังไง พวกเราจะเริ่มงดใช้ผลิตภัณฑ์ของคุณไปทีละอย่าง”

ที่มา: National Geographic (2019)

16.1 ประเทศเคนยามีนโยบายในการงดใช้ถุงพลาสติกอย่างไร

ก. เก็บภาษีถุงพลาสติกแทน

ข. ใช้ถุงโพลีโพรพีลีนแทน

ค. เก็บภาษีผลิตภัณฑ์พลาสติกประเภทอื่น ๆ แทน

ง. ลดการใช้ถุงพลาสติกแบบใช้ครั้งเดียวทิ้ง

จ. ลดการใช้ผลิตภัณฑ์พลาสติกประเภทอื่น ๆ

16.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะประเทศเคนยาสามารถผลิตถุงโพลีโพรพีลีนได้จำนวนมาก

ข. เพราะถุงโพลีโพรพีลีนสามารถย่อยสลายได้ง่าย ไม่ทำลายสิ่งแวดล้อม

ค. เพราะถุงโพลีโพรพีลีนสามารถนำมาใช้ใหม่ได้ และมีความทนทานมากกว่าถุงพลาสติก

17.1 จากข้อความ ***“แม้การห้ามใช้ถุงพลาสติกแบบใช้ครั้งเดียวทิ้งจะมีความสำคัญเพียงใดในการลดขยะพลาสติก มันก็ยังเปลี่ยนโลกไม่ได้”*** หมายความว่าอย่างไร

ก. ทุกคนควรร่วมมือกันลดการใช้ถุงพลาสติก

ข. หน่วยงานทั้งภาครัฐและเอกชนควรมีนโยบายหรือแผนงานอย่างจริงจังในการลดขยะพลาสติก

ค. ขยะพลาสติกมีมากจนไม่สามารถทำให้ลดลงได้

ง. ขยะพลาสติกไม่ได้มีแค่ถุงพลาสติก แต่ยังมีผลิตภัณฑ์พลาสติกอื่น ๆ อีกที่ยังไม่ได้มีการควบคุม

จ. แม้ว่าจะมีมาตรการควบคุมการใช้ถุงพลาสติกแล้ว แต่ก็ยังบังคับใช้ไม่ได้เต็มที่

17.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะผู้เขียนมีความเห็นว่าผลิตภัณฑ์จากพลาสติกยังไม่ได้มีการควบคุมทั้งหมด แต่จะมีจำนวนเพิ่มขึ้นอีก ในอนาคต

ข. เพราะผู้เขียนเห็นว่าขยะพลาสติกมีจำนวนมหาศาล ทำให้ไม่สามารถกำจัดหมดไปได้ภายในเวลาไม่กี่ปี

ค. เพราะผู้เขียนต้องการให้หน่วยงานภาครัฐเข้ามามีส่วนร่วมในการออกมาตรการเพื่อลดขยะพลาสติก

18.1 นาย James Wakaiba มีความคิดเห็นเกี่ยวกับการงดใช้ถุงพลาสติกของประเทศเคนยาอย่างไร

ก. เป็นนโยบายที่เป็นต้นแบบให้ประเทศอื่น ๆ ในการงดใช้ถุงพลาสติก

ข. เป็นนโยบายที่ดีในการลดการใช้ถุงพลาสติกในประเทศเคนยา

ค. เป็นนโยบายที่ทำให้เกิดการเปลี่ยนแปลงอย่างยั่งยืน

ง. เป็นนโยบายที่ไม่เกิดประโยชน์ต่อประเทศเคนยา

จ. เป็นนโยบายที่ส่งผลเสียต่อสิ่งแวดล้อม

18.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะทั่วโลกเริ่มมีการตื่นตัวมากขึ้นในการควบคุมการใช้ถุงพลาสติก

ข. เพราะมีการเสนอร่างกฎหมายเกี่ยวกับการควบคุมการใช้ถุงพลาสติกและผลิตภัณฑ์จากพลาสติก

ค. เพราะนโยบายนี้ช่วยประหยัดถุงพลาสติกที่ใช้ในซูเปอร์มาร์เก็ตไปได้ 100 ล้านใบ ซึ่งถือว่าเป็นความสำเร็จประมาณร้อยละ 80

19.1 “**แต่ยังคงมีคำถามว่าสิ่งนี้จะสามารถช่วยลด*ขยะพลาสติกที่รั่วไหลลงในมหาสมุทรเป็นจำนวนเฉลี่ยกว่า 8 ล้านตันต่อปีได้มากเพียงใด*** ” จากข้อความข้างต้น ผู้เขียนมีจุดมุ่งหมายในการเขียนอย่างไร

ก. เพื่อสื่อสารว่ายังไม่ได้คำตอบว่าการงดใช้พลาสติกนำไปสู่การลดขยะพลาสติกในมหาสมุทรได้

ข.เพื่อช่วยหาวิธีการลดขยะพลาสติกในมหาสมุทร

ค. เพื่อเสนอมาตรการที่สำคัญในการลดขยะพลาสติกในมหาสมุทร

ง. เพื่อเรียกร้องให้หน่วยงานภาครัฐหันมาสนใจปัญหาขยะพลาสติกในมหาสมุทรอย่างจริงจัง

จ. เพื่อรณรงค์ให้ผู้อ่านลดการทิ้งขยะพลาสติกในมหาสมุทร

19.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะผู้เขียนได้ตั้งคำถามเกี่ยวกับการงดใช้พลาสติกว่ามีส่วนช่วยลดขยะพลาสติกที่รั่วไหลลงในมหาสมุทรได้มากน้อยเพียงใด

ข. เพราะผู้เขียนได้สื่อถึงความต้องการที่จะรณรงค์ให้ลดการทิ้งขยะพลาสติกลงในมหาสมุทร

ค. เพราะผู้เขียนได้เขียนเกี่ยวกับปัญหาขยะพลาสติกในมหาสมุทรที่ส่งผลต่อระบบนิเวศทางทะเล

20.1 จากบทความเรื่องนี้ ผู้เขียนแสดง**<u>ความคิดเห็นขัดแย้ง</u>**ที่เกี่ยวข้องกับปัญหาพลาสติกในประเด็นใดมากที่สุด

ก. การงดใช้ถุงพลาสติกแต่ผลิตภัณฑ์พลาสติกอื่น ๆ จะเพิ่มกำลังการผลิต

ข. การห้ามใช้ถุงพลาสติกแต่มีการทดแทนด้วยการใช้ถุงโพลีโพรพีลีน

ค. ถุงพลาสติกเคยเป็นสินค้าขายดีแต่ถูกเปลี่ยนเป็นสินค้าที่ถูกห้ามใช้

ง. นโยบายการงดใช้ถุงพลาสติกระหว่างการเก็บภาษีถุงพลาสติกหรือการใช้ถุงโพลีโพรพีลีน

จ. การใช้ถุงโพลีโพรพีลีนมีผลดีหรือผลเสียต่อสิ่งแวดล้อม

20.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะเนื้อเรื่องกล่าวถึงการรณรงค์งดใช้ผลิตภัณฑ์พลาสติกอื่น ๆ นอกเหนือจากถุงพลาสติก

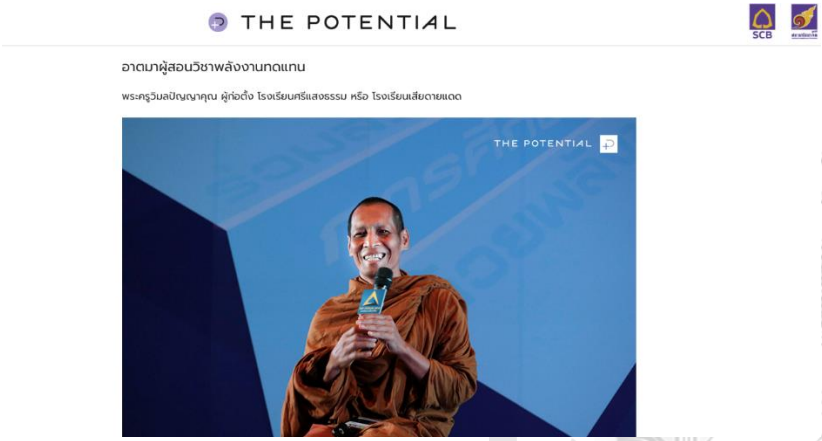
ข. เพราะเนื้อเรื่องให้ความสำคัญถึงปัญหาขยะพลาสติกในมหาสมุทรที่เกิดจากการผลิตวัสดุพลาสติกประเภทใช้ครั้งเดียวทิ้ง

ค. เพราะเนื้อเรื่องของบทความในตอนต้นกล่าวถึงการงดใช้ถุงพลาสติก แต่ตอนท้ายเรื่องกล่าวถึงปัญหาของการผลิตวัสดุจากพลาสติกที่เพิ่มขึ้น

ตัวอย่างข้อสอบการรู้เรื่องการอ่าน ฉบับที่สอง

บทอ่านที่ห้า: พระผู้สอนวิชาพลังงานทดแทน

สถานการณ์: เด็กชายน้อยได้อ่านเรื่องราวของพระท่านหนึ่งที่เปิดโรงเรียนศรีแสงธรรมที่เป็นโรงเรียนพิเศษไม่เหมือนโรงเรียนทั่วไป



ย้อนไปปี 2553 ตอนที่อาตมาเปิดโรงเรียนแรก ๆ ท่ามกลางกับเสียงวิพากษ์วิจารณ์ เขาหาว่าเราเป็นโรงเรียนเถื่อน เป็นพระที่ไม่มีความรู้แล้วจะมาจัดการศึกษาอย่างไร ไม่มีใครอยากเข้ามาเรียน ช่วงแรกมีเด็ก 96 คน ระดับชั้น ม.1 – ม.4 แต่เมื่อสอน ๆ ไป ปีถัดมาเด็กก็เข้ามาเรียนเพิ่มขึ้น แต่อาตมาไม่มีพื้นที่เรียน ต้องมาปั้นบ้านดินเพื่อสร้างอาคารเพิ่มเติม เพราะไม่มีใครให้การสนับสนุน

เมื่อเปิดสอนจริงจังอาตมาก็คิดไม่ออก ไม่รู้ว่าจะต้องทำอย่างไร ก็เลยไปเปิดข้อมูลอ่านว่า การจะพัฒนากำลังคนในประเทศชาติต้องทำอย่างไร และดึงปัญหาออกมาช่วยชาติได้อย่างไร จึงพบข้อมูลมากมาย ทั้งเรื่องความยากจน ความเหลื่อมล้ำ ปัญหายาเสพติด ท้องก่อนวัย คุณภาพการศึกษาตกต่ำ ขาดศีลธรรม ดังนั้นถ้าเราคิดจะทำโรงเรียนทั้งที ก็น่าจะไปตอบโจทย์และแก้ไขปัญหาเหล่านี้ให้ได้ จากตอนนั้นจนถึงตอนนี้อาตมาทำมาตลอด 8 ปี ไม่เคยหยุด ปัญหาที่อาตมาพบคือ วัยแรงงานเริ่มไม่ค่อยมีทักษะและฝีมือในการทำงาน ด้วยความที่อาตมาเป็นช่าง เป็นนักประดิษฐ์ ก็เลยไปขอแผ่นโซลาร์เซลล์ที่แตก ชำรุดแล้วมาใช้เป็นสื่อการเรียนการสอน สำหรับสอนวิชาต่อวงจรไฟฟ้าให้นักเรียน โดยไม่ต้องเสียเงินซื้อ รวมถึงอาตมาก็สอนทฤษฎีต่าง ๆ ไปด้วย เช่น สอนเรื่องกฎของโอห์ม เรื่องการหาความเอียง ความชัน ทำให้เด็กเห็นว่าถ้าทำแบบนี้ เสาข้างหน้าสูง 1 มือ เอียง 15 องศา เสาข้างหลังจะเอียงเท่าไร ประยุกต์ความรู้ไปให้เด็กรู้สึกว่าเรียนแล้วได้ใช้งานจริง กระทั่งมีคนมาศึกษาดูงานที่โรงเรียนแล้วบอกว่า

‘โรงเรียนศรีแสงธรรมมีลักษณะเป็น STEM โดยธรรมชาติ’

อาตมาก็ไม่รู้ว่า STEM คืออะไร อาตมาก็แค่พาเด็กทำ พาเด็กสร้างนวัตกรรม ทำโซลาร์เซลล์ และจำลองระบบโซลาร์เซลล์ ทำทุกอย่างไว้ในโรงเรียน ทุกอาคารในโรงเรียนจะมีการติดตั้งแผ่นโซลาร์เซลล์ทั้งหมด

ถ้ามองการแก้ปัญหาระดับชาติ ระดับชุมชน ผู้ปกครองบางคนเองก็ยังไม่อยากให้เด็กมาเรียน แต่อาตมาบอกเลยว่าถ้าไม่อยากมาเรียนที่นี่ก็อย่ามา เพราะที่นี่มีที่เพียง 25 ที่ รับเด็กได้เพียง 25 คนเท่านั้น ถ้าจะเข้ามาเรียนต้องตอบอาตมาให้ได้ก่อนว่าอยากจะเรียนคณะอะไร รวมถึงปัญหาการไม่มีเงินอุดหนุนมาก เงินอุดหนุนให้มาปีละ 3 แสน แต่อาตมาต้องจ่ายเดือนละ 3.2 แสน ถ้าไม่มีผ้าป่า โรงเรียนนี้ติดลบไปนานแล้ว เพราะเงินอุดหนุนกับงบมันไม่ไปด้วยกัน เราอยู่ได้ด้วยเงินบริจาค แต่ถ้ารอแต่คนมาบริจาค แล้วอาตมาตายไป ใครจะมาบริจาค จึงคิดว่าจะทำอย่างไรให้ยั่งยืน อาตมาจึงให้เด็ก ๆ ออกไปทำงาน รับงานติดตั้งโซลาร์เซลล์ตามสถานที่ต่าง ๆ เพราะเด็กที่มาเรียนกับอาตมาอยู่ฟรี กินฟรี มีรถรับส่งฟรีอยู่แล้ว ดังนั้นต้องออกไปทำงานช่วงปิดเทอม หรือไม่ก็ช่วงเสาร์อาทิตย์ จะได้มีเงินซื้ออาหารกลางวัน แม้ปกติก็ทำนาอยู่ แต่ก็ไม่เพียงพอเพราะเราก็ต้องซื้อหมูซื้อไก่กิน โดยที่เด็ก ม.1-3 อาตมาจะให้เขาเป็นผู้ช่วยอบรม ส่วนพี่ ๆ ม.4-5 ก็จะออกทำงานติดตั้ง เคยไปทำไกลสุดที่ภูเก็ต กรุงเทพฯ ก็เคยไป ส่วนสถานที่สาธารณะก็เคยไปติดให้ เช่น โรงพยาบาลต่าง ๆ

ที่มา: The Potential (2019)

21.1 ข้อใดคือสิ่งที่พระอาจารย์**ไม่ได้**สอนที่โรงเรียนศรีแสงธรรม

ก. วิชาต่อวงจรไฟฟ้า

ข. กฎของโอห์ม

ค. การหาความชันความเอียง

ง. การทำโซลาร์เซลล์

จ. การสร้างบ้านดิน

21.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะพระอาจารย์สร้างบ้านดินเป็นอาคารเรียน

ข. เพราะการสร้างบ้านดินไม่ใช่วิชาที่จะสร้างอาชีพเสริมให้แก่เด็กนักเรียนได้

ค. เพราะนักเรียนสนใจเรียนวิชาวิทยาศาสตร์และนวัตกรรมมากกว่า

22.1 ***"โรงเรียนศรีแสงธรรมมีลักษณะเป็น STEM โดยธรรมชาติ"*** ข้อความนี้หมายความว่าอย่างไร

ก. โรงเรียนสอนให้เรียนรู้ STEM ผ่านการลงมือปฏิบัติจริงในโรงเรียน

ข. โรงเรียนเปิดหลักสูตรเพื่อจัดการเรียนการสอนวิชา STEM เป็นหลัก

ค. โรงเรียนเปิดโอกาสให้นักเรียนสามารถเลือกเรียนวิชา STEM ได้ด้วยตนเอง

ง. โรงเรียนสนับสนุนนักเรียนให้มีความรู้เกี่ยวกับ STEM

จ. โรงเรียนใช้ STEM เป็นสื่อการเรียนการสอน

22.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะผู้พูดต้องการให้นักเรียนออกไปทำงานเพื่อหารายได้ช่วงปิดภาคเรียน

ข. เพราะผู้พูดเคยเป็นช่างและเป็นนักประดิษฐ์มาก่อน

ค. เพราะผู้พูดได้พานักเรียนลงมือทำ สร้างนวัตกรรมและทำโซลาร์เซลล์ใช้เองในโรงเรียน

23.1 ข้อใดคือหลักปฏิบัติที่สำคัญที่สุดของพระอาจารย์ในบทความเรื่องนี้

ก. ความเรียบง่าย

ข. ความอดทน

ค. ความซื่อสัตย์

ง. การพึ่งตนเอง

จ. การประหยัด

23.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะผู้พูดบอกว่าโรงเรียนได้รับเงินอุดหนุนไม่เพียงพอ

ข. เพราะผู้พูดพัฒนาโรงเรียนด้วยตนเอง รวมถึงการติดตั้งโซลาร์เซลล์ใช้ในโรงเรียนด้วย

ค. เพราะผู้พูดพยายามที่จะแก้ปัญหาของเยาวชนและพัฒนากำลังคนในประเทศชาติ

24.1 เพราะเหตุใดพระอาจารย์จึงกล่าวถึง **ความยั่งยืน** ในข้อความที่ว่า ***"เราอยู่ได้ด้วยเงินบริจาค แต่ถ้ารอคนมาบริจาค แล้วอาตมาตายไป ใครจะมาบริจาค จึงคิดว่าจะทำอย่างไรให้ยั่งยืน อาตมาจึงให้เด็ก ๆ ออกไปทำงาน"***

ก. เพื่อประชาสัมพันธ์โรงเรียนให้แก่ผู้ปกครองหรือบุคคลที่สนใจทราบ

ข. เพื่ออธิบายถึงปัญหาของการไม่มีเงินอุดหนุนโรงเรียน

ค. เพื่อเปรียบเทียบวิธีการที่พระอาจารย์ใช้ในการหารายได้เข้าโรงเรียน

ง. เพื่อเสนอแนะแนวทางที่ใช้ในการหารายได้เข้าโรงเรียน

จ. เพื่ออธิบายถึงวิธีการในการหารายได้เพื่อให้โรงเรียนสามารถพึ่งตนเองได้

24.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะผู้พูดได้กล่าวถึงปัญหาเงินอุดหนุนไม่เพียงพอ จึงคิดหารายได้เข้าโรงเรียนนอกจากการรับเงินบริจาคเท่านั้น

ข. เพราะผู้พูดได้กล่าวว่าต้องการให้นักเรียนได้ใช้เวลาว่างในช่วงปิดภาคเรียนให้เกิดประโยชน์

ค. เพราะผู้พูดได้กล่าวว่าต้องการให้นักเรียนมีเงินที่เพียงพอในการซื้ออาหารกลางวัน

25.1 นักเรียนคิดว่าบทความเรื่องนี้มีประเด็นใดที่น่าเชื่อถือมากที่สุด

ก. มีข้อมูลที่ได้จากผู้พูดเอง ไม่ได้มีถูกปรับแต่งหรือเพิ่มเติมโดยบุคคลอื่น ๆ

ข. มีการเผยแพร่ในเว็บไซต์ที่ได้รับการสนับสนุนจากหน่วยงานที่มีความน่าเชื่อถือ

ค. มีการระบุชื่อของเว็บไซต์ที่เผยแพร่บทความ

ง. มีการระบุวันเดือนปีที่เผยแพร่บทความ

จ. มีการให้ที่อยู่อีเมลที่ผู้อ่านสามารถติดต่อผู้ดูแลเว็บไซต์ได้

25.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะเป็นข้อเท็จจริงที่เกิดจากประสบการณ์ของผู้พูดเอง

ข. เพราะผู้พูดเป็นบุคคลต้นแบบในการสอน

ค. เพราะผู้พูดมีความจริงใจในการถ่ายทอดเรื่องราว

CHULALONGKORN UNIVERSITY จุฬาลงกรณ์มหาวิทยาลัย

บทอ่านที่หก: การรณรงค์ลดการใช้ขยะ

สถานการณ์: เด็กหญิงแจนได้เห็นข่าวการลดใช้ขยะทั้งโครงการ zero waste ของจุฬาลงกรณ์มหาวิทยาลัย และหลัก 7R ช่วยลดปริมาณขยะของ greenery

ส่งขยะพลาสติกแบบใช้ครั้งเดียวทิ้งมาให้เรา Chula zero waste ขอรับไว้เอง

หลอดใช้เสร็จแล้วไม่อยากทิ้งถังขยะ ฝาครอบแก้วหรือถุงหิ้วที่เผลอรับมาก็ยังสะอาดอยู่ ถ้าไม่รู้ว่าจะเอาไปเก็บไหนดี จะทิ้งลงถังขยะก็รีไซเคิลได้ยาก CHULA zero waste ขอรับพลาสติกใช้แล้วไว้เอง ทั้ง 4 ประเภทได้แก่...หลอดพลาสติก ฝาครอบแก้วแก้วพลาสติก และถุงหิ้วพลาสติก ที่แห้งและสะอาดเพื่อนำมาทำประติมากรรมจากขยะแต่จะทำเป็นรูปอะไรนั้นขออุบไว้ก่อนนะ เปิดรับตั้งแต่วันนี้ถึงวันที่ 15 มกราคม 2562 เท่านั้น

สามารถนำไปให้ได้ที่จุดรับบริจาคในร้านค้าสหกรณ์จุฬาฯ สาขาศาลาพระเกี้ยว สาขาหอใน และที่อาคารจามจุรี 5 หรือส่งเป็นพัสดุมาได้ที่โครงการ Chula zero waste ชั้น 2 อาคารสถาบัน 2 สถาบันวิจัยสภาวะแวดล้อม จุฬาลงกรณ์มหาวิทยาลัย ถนนพญาไท แขวงวังใหม่ เขตปทุมวัน กรุงเทพมหานคร 10330 ได้เลย

รูปที่ 1 ขอรับบริจาคขยะพลาสติกแบบใช้ครั้งเดียวทิ้ง

ที่มา: Facebook CHULA Zero Waste (2019)

เจ็ดแล้วจำ: หลัก 7R ช่วยลดปริมาณขยะ

นอกจากหลัก 3R ที่หลายคนจำได้ขึ้นใจแล้ว เรามาแอนวานซ์ขึ้นไปอีกขั้นด้วยหลัก 7R ที่ชวนเรา ลด ละ และเลิกพฤติกรรมที่จะสร้างขยะเกินความจำเป็น ลองจำให้ขึ้นใจ แล้วนำไปปรับใช้จริง เราก็จะกลายเป็นหนึ่งหน่วยเล็ก ๆ ที่กำลังพยายามช่วยลดภาระให้โลกใบนี้ เหมือนอีกหลาย ๆ หน่วยที่จะรวมกันเป็นพลัง

รูปที่ 2 หลัก 7R ช่วยลดปริมาณขยะ

ที่มา: Greenery (2017)

26.1 จากรูปภาพที่ 1 ข้อใด**ไม่ใช่**ขยะพลาสติกแบบใช้ครั้งเดียวทิ้งที่จุฬาลงกรณ์มหาวิทยาลัยขอรับบริจาค

ก. ฝาครอบแก้ว

ข. ถุงพลาสติกแบบหูหิ้ว

ค. หลอดพลาสติก

ง. แก้วน้ำพลาสติก

จ. ฝาขวดพลาสติก

26.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะตัวเลือกดังกล่าวนำมาสร้างเป็นประติมากรรมได้ยาก

ข. เพราะตัวเลือกดังกล่าวไม่มีรูปปรากฏอยู่ในขยะพลาสติกที่ขอรับบริจาค

ค. เพราะตัวเลือกดังกล่าวจะรับบริจาคพร้อมกับขวดน้ำพลาสติกในภายหลัง

27.1 ผู้อ่านสามารถบริจาคขยะพลาสติกได้ที่ใด

ก. กล่องรับบริจาคที่หน่วยงานไปรษณีย์ใกล้บ้าน

ข. ร้านสหกรณ์ จุฬาลงกรณ์มหาวิทยาลัย

ค. กล่องรับบริจาคตามห้างร้านต่าง ๆ

ง. สำนักกองทุนสนับสนุนการสร้างเสริมสุขภาพ

จ. กล่องรับบริจาคภายในจุฬาลงกรณ์มหาวิทยาลัย

27.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะเป็นสถานที่ที่เข้าถึงได้ง่าย สามารถเดินทางไปบริจาคได้สะดวก

ข. เพราะเป็นสถานที่ที่มีความสะดวกในการเก็บรักษาขยะพลาสติกที่ได้จากการบริจาค

ค. เพราะจากข่าวประกาศจะพบว่าจุดรับบริจาคคือร้านสหกรณ์จุฬาลงกรณ์มหาวิทยาลัยและอาคารจามจุรี 5หรือส่งเป็นพัสดุมาที่โครงการ Chula zero waste จุฬาลงกรณ์มหาวิทยาลัย

28.1 ข้อความ **“ส่งขยะพลาสติกแบบใช้ครั้งเดียวทิ้งมาให้เรา Chula zero waste ขอรับไว้เอง”** มีความหมายว่าอย่างไร?

ก. Chula zero waste เป็นหน่วยงานที่คัดแยกขยะพลาสติก

ข. Chula zero waste รวบรวมขยะพลาสติกเพื่อนำไปกำจัดทิ้ง

ค. Chula zero waste เป็นโครงการที่เปลี่ยนขยะให้เป็นเงิน

ง. Chula zero waste ต้องการให้นิสิตคัดแยกขยะพลาสติกเป็น

จ. Chula zero waste ต้องการขยะพลาสติกแบบใช้ครั้งเดียวทิ้งมารีไซเคิล

28.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. ตามรูปภาพที่ 1 Chula zero waste ขอรับบริจาคขยะเพื่อมาทำผลประติมากรรมจากขยะ

ข. Chula zero waste เป็นโครงการที่ช่วยอนุรักษ์สิ่งแวดล้อม

ค. โครงการนี้ช่วยรณรงค์การลดใช้ขยะพลาสติก

29.1 ผู้ใดไม่ได้ปฏิบัติตามหลักของ 7R

ก. เด็กชายแต้มจะแยกแก้วน้ำ หลอด และฝาครอบแก้วก่อนทิ้งเสมอ

ข. เด็กชายโต้ใช้ถุงผ้าในการซื้อของในร้านสะดวกซื้อ

ค. เด็กชายต้นชอบซื้อครีมอาบน้ำและน้ำยาล้างจานแบบถุงเติม

ง. เด็กชายตั้มไม่รับกล่องโฟมแต่ใช้ปิ่นโตหรือกล่องข้าวตอนซื้อกับข้าว

จ. เด็กชายต๋อยจะซื้อคอมพิวเตอร์และมือถือเครื่องใหม่ใช้ ไม่ชอบซ่อมของเสีย

29.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะการซ่อมคอมพิวเตอร์และมือถือมีค่าใช้จ่ายที่สูงกว่าซื้อใหม่

ข. เพราะเด็กชายต๋อยไม่ได้ปฏิบัติตามหลักของ Repair โดยการซ่อมแซมของเสียก่อน

ค. เพราะคอมพิวเตอร์และมือถือมีความเก่าและล้าสมัยแล้ว จึงไม่คุ้มค่าในการซ่อม

นานเกินไปแล้วที่บริษัทยักษ์ใหญ่กอบโกยรายได้มหาศาล โดยยัดเยียดบรรจุภัณฑ์พลาสติกให้ผู้บริโภคโดยไม่มีทางเลือก แม้บริษัททั้งหลายจะอ้างว่าไม่รู้ปลายทางของบรรจุภัณฑ์พลาสติกของตนจบลง ณ ที่ใด ทางแก้ปัญหากลับเป็นเพียงการเน้นไปที่การเพิ่มความสามารถรีไซเคิลเท่านั้น ซึ่งการรีไซเคิลไม่ใช่ทางออกของปัญหา เพราะ 90% ของพลาสติกเหล่านี้ไม่ได้รับการรีไซเคิล แต่ถูกนำไปที่หลุมฝังกลบหรือถูกนำไปเผาและปล่อยมลพิษกลับเข้าสู่สิ่งแวดล้อม เราไม่อาจกู้วิกฤตมลพิษพลาสติกด้วยเพียงแค่รีไซเคิล ถึงเวลาแล้วที่บริษัททั้งหลายต้องบอกลาพลาสติกที่ใช้ครั้งเดียวไปพร้อมกัน

ปรับมาจาก Greenpeace (2019)

30.1 จากข้อความข้างต้นที่มีความขัดแย้งกับรูปภาพที่ 1 ที่กล่าวถึงการขอรับบริจาคขยะพลาสติกเพื่อนำไปรีไซเคิล

ข้อใดเป็นประเด็นความขัดแย้งที่ผู้เขียนต้องการนำเสนอ

ก. บริษัทผู้ผลิตพยายามส่งเสริมการรีไซเคิลขยะพลาสติกที่ถูกวิธีให้แก่สังคม แต่ขยะพลาสติกบางประเภทนำมารีไซเคิลไม่ได้

ข. ทางแก้ปัญหาขยะพลาสติกอยู่ที่บริษัทผู้ผลิตต้องลดจำนวนการผลิตบรรจุภัณฑ์พลาสติก ไม่ใช่แค่การรีไซเคิลที่ แก้ปัญหาได้เพียงบางส่วน

ค. ทางออกของการแก้ปัญหาขยะพลาสติกในอดีตที่แตกต่างจากในปัจจุบัน

ง. ความคุ้มค่าของการจัดการขยะพลาสติกด้วยวิธีการฝังกลบและการเผา เมื่อเปรียบเทียบกับการเพิ่มความสามารถในการรีไซเคิล

จ. มาตรการของภาครัฐและนโยบายของบริษัทผู้ผลิตเกี่ยวกับการรีไซเคิลขยะพลาสติกไม่สอดคล้องกัน

30.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

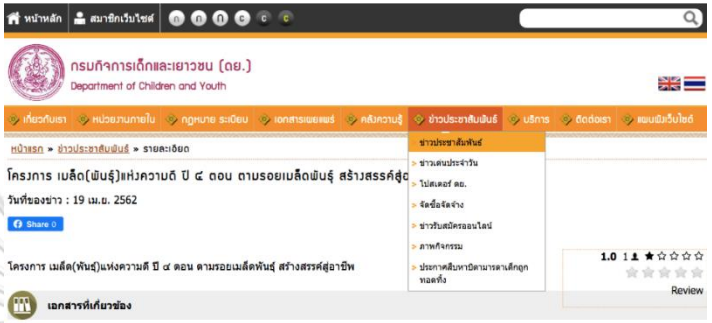
ก. เพราะคนไทยยังมีข้อจำกัดทางด้านการใช้เทคโนโลยีในการรีไซเคิลขยะพลาสติก

ข. เพราะประเทศไทยยังไม่มีกฎหมายควบคุมการผลิตบรรจุภัณฑ์พลาสติกของบริษัทเอกชนที่เข้มงวดมากพอ

ค. เพราะผู้เขียนได้กล่าวถึงการรีไซเคิลว่าไม่ใช่ทางออกของปัญหา และถึงเวลาแล้วที่บริษัทต้องบอกลาพลาสติกที่ ใช้ครั้งเดียวทิ้ง

บทอ่านที่เจ็ด: กองทุนเยาวชนพัฒนา

สถานการณ์: เด็กชายบอลได้ทราบข่าวการประกวดโครงการเมล็ด(พันธุ์)แห่งความดีจากคุณครู จึงเข้าอินเทอร์เน็ตเพื่อหาข้อมูลเกี่ยวกับโครงการดังกล่าวเพิ่มเติมจากเว็บไซต์ของกรมกิจการเด็กและเยาวชน



หลักเกณฑ์การพิจารณาเงินสนับสนุนโครงการเมล็ด(พันธุ์)แห่งความดี ปี ๔ ตอน... “ตามรอยเมล็ดพันธุ์ สร้างสรรค์สู่อาชีพ”

๑. **เงื่อนไขการพิจารณาเงินสนับสนุนโครงการ**

๑.๑ เป็นโครงการที่ดำเนินงานเพื่อเด็กเยาวชนและชุมชน

๑.๒ เป็นโครงการที่เล็งเห็นถึงความสำคัญในการส่งเสริมสนับสนุนและพัฒนาเด็กและเยาวชนในพื้นที่

๑.๓ เป็นโครงการที่สามารถเป็นต้นแบบแก่หน่วยงาน/ ภาครัฐ/ เอกชนหรือเครือข่ายอื่น ๆ ได้

๑.๔ ไม่เป็นโครงการที่อยู่ในระหว่างการดำเนินการหรือรับเงินอุดหนุนใดจากกระทรวงการพัฒนาสังคมและความมั่นคงของมนุษย์

๑.๕ เป็นโครงการที่ดีมีประโยชน์ตรงเป้าหมายและเป็นรูปธรรมชัดเจนหรือเป็นโครงการที่สร้างนวัตกรรมใหม่ ๆ ในการพัฒนาเด็กและเยาวชนหรือตามเป้าหมายพัฒนาที่ยั่งยืน (SDGs)

๑.๖ เป็นโครงการที่สอดคล้องกับแผนพัฒนาเด็กและเยาวชนแห่งชาติ โดยการพัฒนาศักยภาพและสร้างภูมิคุ้มกันเด็กและเยาวชน ในเรื่องของการพัฒนาทักษะส่งเสริมสร้างสรรค์ เรื่อง อาชีพ

๑.๗ เป็นโครงการที่ปลูกจิตสำนึกในการรับผิดชอบต่อสังคม

๑.๘ สภาเด็กและเยาวชนแห่งประเทศไทยมีสิทธิ์เผยแพร่จัดพิมพ์โครงการทุกโครงการที่เข้าร่วมประกวดโดยปราศจากข้อผูกมัดใด ๆ

๑.๙ การตัดสินการประกวดจะพิจารณาโดยคณะกรรมการผู้ทรงคุณวุฒิจากสภาเด็กและเยาวชนแห่งประเทศไทย โดยให้ถือเป็นที่สิ้นสุด

๑.๑๐ โครงการที่ได้รับการพิจารณาคัดเลือกต้องดำเนินงานให้เสร็จภายในเดือนสิงหาคม ๒๕๖๒

ที่มา: กรมกิจการเด็กและเยาวชน (2019)

31.1 ถ้าต้องการทราบข้อมูลเพิ่มเติมเกี่ยวกับการประกวดโครงการเมล็ด(พันธุ์)แห่งความดี ต้องสืบค้นจากหัวข้อใดบนเว็บไซต์

ก. ภาพกิจกรรม

ข. ข่าวเด่นประจำวัน

ค. ข่าวประชาสัมพันธ์

ง. จัดซื้อจัดจ้าง

จ. โปสเตอร์ ดย.

31.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะหัวข้อดังกล่าวเป็นข่าวที่บุคคลทั่วไปสามารถเข้าถึงได้

ข. เพราะตามสถิติเว็บไซต์ จำนวนของคนที่เข้ามาดูในส่วนของหัวข้อดังกล่าวมีมากกว่าส่วนอื่น ๆ

ค. เพราะโครงการที่ประชาสัมพันธ์จากหน่วยงานที่เกี่ยวข้องและเปิดรับสมัครจะอยู่ในส่วนของหัวข้อดังกล่าว

32.1 โครงการเมล็ด(พันธุ์)แห่งความดี จัดขึ้นในปี 2562 เป็นครั้งที่เท่าไหร่

ก. ครั้งที่ 1

ข. ครั้งที่ 2

ค. ครั้งที่ 3

ง. ครั้งที่ 4

จ. ไม่มีข้อมูลปรากฏ

32.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะโครงการเมล็ด(พันธุ์)แห่งความดี จัดขึ้นเป็นปีที่ 4 ในปี 2562

ข. เพราะโครงการได้งบจากกระทรวงพัฒนาสังคมและความมั่นคงของมนุษย์

ค. เพราะแผนพัฒนาเด็กและเยาวชนแห่งชาติกำหนดไว้ให้จัดโครงการส่งเสริมเยาวชน

33.1 ข้อใดคือวัตถุประสงค์หลักของโครงการเมล็ด(พันธุ์)แห่งความดี

ก. เพื่อสนับสนุนเด็กและเยาวชนในการสร้างอาชีพสร้างรายได้ในชุมชน

ข. เพื่อสนับสนุนเด็กและเยาวชนในการอนุรักษ์สิ่งแวดล้อมในชุมชน

ค. เพื่อสนับสนุนเด็กและเยาวชนให้ได้รับทุนการศึกษา

ง. เพื่อสนับสนุนเด็กและเยาวชนในการทำความดีตอบแทนผู้ปกครองและครู

จ. เพื่อสนับสนุนเด็กและเยาวชนให้มีอาชีพเสริมที่โรงเรียน

33.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะโครงการนี้สนับสนุนงบเพื่อพัฒนาเด็กและเยาวชน ตามเงื่อนไขการพิจารณาเงินสนับสนุนโครงการ

ข. เพราะโครงการนี้ต้องการให้เยาวชนเติบโตไปเป็นผู้ใหญ่ที่มีความรู้ความสามารถในอนาคต

ค. เพราะโครงการนี้ต้องการส่งเสริมเยาวชนซึ่งเป็นวัยที่กล้าแสดงออก กล้าคิด กล้าทำ

34.1 ใครเป็นผู้ที่มีคุณสมบัติเหมาะสมในการเข้าร่วมโครงการเมล็ด(พันธุ์)แห่งความดีมากที่สุด

ก. เด็กชายเก่งต้องการเงินทุนเพื่อมาประกอบอาชีพ

ข. เด็กชายหนึ่งสนใจเข้าร่วมโครงการเพื่อเพิ่มทักษะการทำงานเป็นทีม

ค. เด็กหญิงใหม่ต้องการทำกิจกรรมเพื่อพัฒนาโรงเรียน

ง. เด็กหญิงพลอยสนใจเข้าร่วมโครงการเพื่อพัฒนาชุมชนของตนเอง

จ. เด็กชายเพชรต้องการรางวัลและโล่ประกาศเกียรติคุณจากรัฐมนตรี

34.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะการให้ทุนช่วยสร้างประโยชน์ในระยะยาวให้แก่ชุมชน

ข. เพราะเงื่อนไขทุนมีการส่งเสริมโครงการที่ปลูกจิตสำนึกในการรับผิดชอบสังคม

ค. เพราะเงินทุนของโครงการเป็นปัจจัยที่ก่อให้เกิดรายได้และเกิดการพัฒนาชุมชน

35.1 ประกาศโครงการ “เมล็ด(พันธุ์)แห่งความดี” มีความน่าเชื่อถือ ยกเว้นข้อใดต่อไปนี้

ก. ไม่มีค่าใช้จ่ายในการสมัคร

ข. มีการบอกรายละเอียดข้อมูลการรับสมัครไว้อย่างชัดเจน

ค. มีรายละเอียดของหมายเลขโทรศัพท์และที่อยู่ที่สามารถติดต่อได้เพื่อใช้ในการสอบถามข้อมูลเพิ่มเติม

ง. มีการประชาสัมพันธ์ในเว็บไซต์ของหน่วยงานภาครัฐ

จ. มีการระบุชื่อหน่วยงานที่จัดไว้อย่างชัดเจน

35.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะการมีหรือไม่มีค่าใช้จ่ายเป็นรายละเอียดของการรับสมัครไม่เกี่ยวกับความน่าเชื่อถือของประกาศ

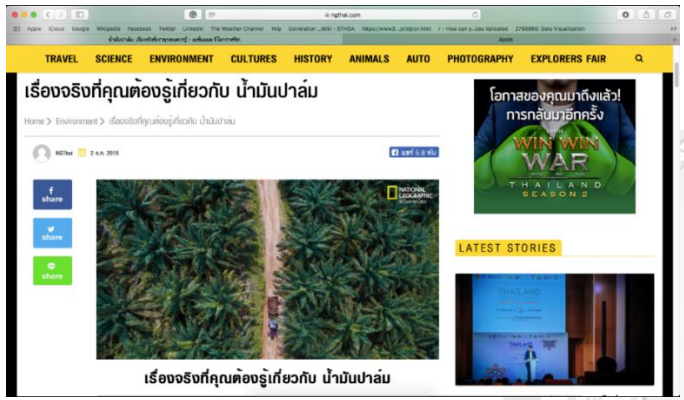
ข. เพราะโครงการเมล็ด(พันธุ์)แห่งความดีเปิดรับสมัครเฉพาะเด็กและเยาวชนจึงไม่คิดค่าใช้จ่ายในการสมัคร

ค. เพราะโครงการเมล็ด(พันธุ์)แห่งความดีได้รับการสนับสนุนเงินจากสภาเด็กและเยาวชนแห่งประเทศไทย

จุฬาลงกรณ์มหาวิทยาลัย CHULALONGKORN UNIVERSITY

บทอ่านที่แปด: น้ำมันปาล์ม

สถานการณ์: คุณครูได้ให้เด็กชายเฟิร์สหาข่าวเกี่ยวกับสถานการณ์ในปัจจุบันเพื่อนำมาเล่าให้เพื่อนฟังในชั้นเรียน เด็กชายเฟิร์สได้เจอบทความที่เกี่ยวกับน้ำมันปาล์ม



เรื่องจริงที่คุณต้องรู้เกี่ยวกับ น้ำมันปาล์ม

ปัจจุบัน น้ำมันปาล์มคือน้ำมันพืชที่ใช้กันมากที่สุดในโลก หรือคิดเป็นหนึ่งในสามของการบริโภคน้ำมันพืชทั่วโลก โดยเป็นน้ำมันปรุงอาหารหลักในหลายประเทศ น้ำมันปาล์มเป็นส่วนผสมที่แทบจะหลีกเลี่ยงไม่ได้ในสินค้าสารพัด ตั้งแต่ผลิตภัณฑ์ที่วางขายในซูเปอร์มาร์เก็ต ไปจนถึงน้ำมันไบโอดีเซลที่เชื่อกันว่าเป็นมิตรต่อสิ่งแวดล้อม

ความต้องการน้ำมันปาล์มทั่วโลกยังคงสูงขึ้น อินเดียบริโภคน้ำมันปาล์มมากที่สุด หรือคิดเป็นร้อยละ 17 ของโลก ตามมาด้วยอินโดนีเซีย สหภาพยุโรป และจีน ตามลำดับ ส่วนสหรัฐฯ รั้งอันดับแปด ในปี 2018 คาดว่าการบริโภคน้ำมันปาล์มของโลกจะสูงถึง 65.5 ล้านตัน หรือเทียบเท่ากับน้ำมันปาล์มราว 9 กิโลกรัมต่อคน

การตอบสนองความต้องการดังกล่าวส่งผลเสียใหญ่หลวง นับตั้งแต่ปี 1973 ป่าดิบชื้น 41,000 ตารางกิโลเมตรในบอร์เนียว ซึ่งเป็นเกาะที่มาเลเซียและอินโดนีเซียถือครองร่วมกัน ถูกแผ้วถางเพื่อทำไม้ เผา และไถปราบเพื่อปลูกปาล์มน้ำมัน คิดเป็นหนึ่งในห้าของการตัดไม้ทำลายป่าบนเกาะบอร์เนียวตั้งแต่ปี 1973 และร้อยละ 47 ตั้งแต่ปี 2000

การตัดไม้ทำลายป่าทั้งหมดนั้นส่งผลกระทบหนักหนาสาหัสต่อสัตว์ป่า อุรังอุตังเกาะบอร์เนียวซึ่งอยู่ในสถานะใกล้สูญพันธุ์อย่างยิ่งล้มตายไปเกือบ 150,000 ตัวระหว่างปี 1999 ถึง 2015 และถึงแม้สาเหตุหลักจะเป็นการตัดไม้และการล่าสัตว์ แต่น้ำมันปาล์มก็เป็นปัจจัยสำคัญ ทั้งยังเร่งการเปลี่ยนแปลงสภาพภูมิอากาศด้วย การปล่อยก๊าซเรือนกระจกของอินโดนีเซียเกือบครึ่งหนึ่งมาจากการตัดไม้ทำลายป่า และการเปลี่ยนแปลงการใช้ประโยชน์ในที่ดินอื่น ๆ ตลอดจนมลพิษทางอากาศขั้นรุนแรง

ผู้คนที่อาศัยในที่ดินทำสวนปาล์มยังทนทุกข์กับปัญหาอื่น ๆ การละเมิดสิทธิมนุษยชน เช่น การใช้แรงงานเด็ก และการบังคับขับไล่คือปัญหาที่มีรายงานชัดเจน บางครั้งบริษัทน้ำมันปาล์มหลายแห่งบนเกาะสุมาตราของอินโดนีเซียใช้แทรกเตอร์ไถหมู่บ้านชนพื้นเมืองราบเป็นหน้ากลองทั้งหมู่บ้าน ทำให้ชาวบ้านไร้ที่อยู่อาศัยและต้องพึ่งพาความช่วยเหลือจากรัฐบาล

การทำลายระบบนิเวศเพื่อประโยชน์ระยะสั้นเช่นนั้น คือสิ่งที่ประเทศกาบองพยายามหลีกเลี่ยง สวนสวรรค์ที่ฉันไปเยือนจะไม่ถูกทำลาย โอลัมปกป้องป่าผืนนั้นไว้ตามข้อตกลงเพื่อแลกกับการที่รัฐบาลยอมให้บริษัทปลูกปาล์มน้ำมันที่อื่นใดก็ได้ในที่ดินสัมปทาน

“สิ่งที่เราพยายามทำในกาบอง คือการหาวิธีพัฒนาใหม่ที่ไม่ต้องตัดต้นไม้ทั้งป่าทิ้ง แต่รักษาสมดุลระหว่างสวนปาล์มน้ำมัน การเกษตร และการอนุรักษ์ป่าไว้ได้ครับ” ลี ไวต์ นักชีววิทยาเชิงอนุรักษ์ อธิบดีกรมอุทยานของกาบอง บอก เมื่อประเทศที่มีประชากรไม่ถึงสองล้านคนเริ่มทำการเกษตรระดับอุตสาหกรรม รัฐบาลก็ใช้การประเมินทางวิทยาศาสตร์เพื่อตัดสินว่าผืนป่าอันกว้างใหญ่ส่วนใดมีคุณค่าเชิงอนุรักษ์สูง และส่วนใดอาจเปิดพื้นที่ให้ปลูกปาล์มน้ำมันได้

ปาล์มน้ำมันจะอยู่ต่อไปในแอฟริกา เช่นเดียวกับในเอเชียตะวันออกเฉียงใต้ ประเทศผู้ผลิตพึ่งพารายได้จากพืชเศรษฐกิจชนิดนี้ การคว่ำบาตรน้ำมันปาล์มจึงไม่ใช่มาตรการที่ฉลาดนัก เพราะพืชน้ำมันทางเลือกอื่น ๆ อาจใช้ที่ดินมากกว่าเสียอีก มิหนำซ้ำยังอาจไร้ประโยชน์อีกด้วย เนื่องจากน้ำมันปาล์มแพร่หลายอย่างกว้างขวางและมักแปรรูปเป็นส่วนประกอบอื่น ๆ เช่น โซเดียมลอริลซัลเฟตและกรดสเตียริกซึ่งผู้บริโภคมักไม่รู้ที่มา การจะหักดิบจากการบริโภคน้ำมันปาล์มจึงเป็นเรื่องยาก วิธีเดียวที่ทำได้คือทำให้การเพาะปลูกและการผลิตส่งผลกระทบเลวร้ายน้อยที่สุด

ที่มา: National Geographic (2019)

36.1 ประเทศหรือกลุ่มประเทศใดที่บริโภคน้ำมันปาล์มมากที่สุดเป็นอันดับสี่ของโลก

ก. ประเทศอินเดีย

ข. ประเทศสหรัฐอเมริกา

ค. ประเทศจีน

ง. ประเทศอินโดนีเซีย

จ. สหภาพยุโรป

36.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะเป็นประเทศที่มีพื้นที่มาก สามารถปลูกต้นปาล์มได้มาก

ข. เพราะบทความได้บอกลำดับของประเทศที่บริโภคน้ำมันปาล์ม

ค. เพราะเป็นประเทศที่มีประชากรมากที่สุดเป็นอันดับต้น ๆ ของโลก

37.1 นักเรียนคิดว่าข้อความนี้มีความหมายว่าอย่างไร

"การทำลายระบบนิเวศเพื่อประโยชน์ระยะสั้นเช่นนั้น คือสิ่งที่ประเทศกาบองพยายามหลีกเลี่ยง"

ก. ประเทศกาบองจะลดการทำลายระบบนิเวศโดยรักษาสมดุลระหว่างการปลูกปาล์มและการอนุรักษ์ป่า

ข. ประเทศกาบองไม่สนับสนุนการปลูกปาล์มจากบริษัทปลูกปาล์มน้ำมัน

ค. ประเทศกาบองมีการปลูกปาล์มโดยไม่คำนึงถึงปัญหาทางสิ่งแวดล้อม

ง. ประเทศกาบองมีแนวคิดในการปลูกพืชน้ำมันทางเลือกเพื่อทดแทนการตัดไม้ทำลายป่า

จ. ประเทศกาบองมีนโยบายลดการปลูกปาล์มน้ำมัน

37.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะผู้เขียนกล่าวถึงวิธีที่ประเทศกาบองใช้เพื่อปลูกปาล์ม โดยคำนึงถึงการอนุรักษ์ป่า

ข. เพราะผู้เขียนกล่าวว่าคนในประเทศกาบองต่อต้านการตัดไม้ทำลายป่า

ค. เพราะผู้เขียนกล่าวว่าพืชน้ำมันทางเลือกอื่น ๆ จะใช้ที่ดินมากกว่าการปลูกปาล์ม

38.1 ผู้เขียนกล่าวถึงความต้องการน้ำมันปาล์มที่สูงขึ้นมีผลต่อคนในพื้นที่อย่างไร

ก. สร้างงานสร้างอาชีพให้คนในพื้นที่

ข. พัฒนาคุณภาพชีวิตแรงงาน

ค. นำรายได้มาพัฒนาท้องถิ่น

ง. กดดันให้ย้ายถิ่นฐาน

จ. มีการเพิ่มขึ้นของประชากรในพื้นที่

38.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะการปลูกปาล์มมีการบุกรุกพื้นที่คนท้องถิ่น

ข. เพราะการปลูกปาล์มต้องตัดไม้ทำลายป่า

ค. เพราะการปลูกปาล์มทำให้เกิดมลพิษทางสิ่งแวดล้อม

39.1 ***"การคว่ำบาตรปาล์มน้ำมันจึงไม่ใช่มาตรการที่ฉลาดนัก เพราะพืชน้ำมันทางเลือกอื่น ๆ อาจใช้ที่ดินมากกว่าเสียอีก"***จากข้อความข้างต้นสะท้อนถึงจุดมุ่งหมายของผู้เขียนอย่างไร

ก. เพื่อรณรงค์ให้ผู้อ่านเห็นด้วยกับการคว่ำบาตรปาล์มน้ำมัน

ข. เพื่อเรียกร้องให้หน่วยงานภาครัฐออกนโยบายในการคว่ำบาตรปาล์มน้ำมัน

ค. เพื่อสื่อสารว่าการคว่ำบาตรปาล์มน้ำมันไม่ใช่ทางออกที่ดี

ง. เพื่อเสนอวิธีการให้หน่วยงานที่เกี่ยวข้องในการควบคุมปัญหาเรื่องปาล์มน้ำมัน

จ. เพื่อยกตัวอย่างถึงวิธีการที่ใช้ในการแก้ปัญหาปาล์มน้ำมัน

39.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะผู้เขียนได้ชี้ให้เห็นว่าพืชทางเลือกอื่น ๆ อาจมีผลเสียมากกว่าการปลูกปาล์ม

ข. เพราะประเทศทั่วโลกใช้พื้นที่ในการปลูกปาล์มน้ำมันมากกว่าปลูกพืชน้ำมันชนิดอื่น ๆ

ค. เพราะปาล์มน้ำมันเป็นน้ำมันพืชที่ใช้กันมากที่สุดในโลก

40.1 ผู้เขียนบทความเรื่องนี้แสดง**ความคิดเห็นที่ขัดแย้ง**ในเรื่องของปัญหาปาล์มน้ำมันตามข้อใด

ก. แม้การปลูกปาล์มน้ำมันจะทำลายระบบนิเวศ แต่ก็มีการเปลี่ยนมาปลูกปาล์มร่วมกับการอนุรักษ์ป่าไม้

ข. แม้จะมีการปลูกพืชน้ำมันอื่น ๆ เพื่อเป็นทางเลือกทดแทนเพิ่มมากขึ้น แต่ปริมาณการบริโภคน้ำมันปาล์มยังสูงขึ้นต่อเนื่อง

ค. แม้จะมีความต้องการปาล์มน้ำมันที่เพิ่มขึ้นทั่วโลก แต่จำนวนการปลูกต้นปาล์มลดลง

ง. แม้รายได้จากการส่งออกปาล์มเพิ่มขึ้น แต่ปัญหาการละเมิดสิทธิมนุษยชนก็เพิ่มมากขึ้นด้วย

จ. แม้รัฐบาลมีนโยบายปกป้องพื้นป่าที่จริงจัง แต่ก็มีความขัดแย้งกับเอกชนที่ต้องการปลูกปาล์มน้ำมัน

40.2 เหตุใดนักเรียนจึงเลือกตัวเลือกนี้

ก. เพราะเนื้อเรื่องได้กล่าวถึงปัญหาการตัดไม้ทำลายป่าเพื่อปลูกปาล์มและการรักษาสมดุลระหว่างป่าและการปลูกปาล์ม

ข. เพราะเนื้อเรื่องได้กล่าวถึงการปลูกปาล์มที่ส่งผลกระทบต่อระบบนิเวศป่าไม่ สัตว์ป่าและที่อยู่อาศัยของชาวบ้าน

ค. เพราะเนื้อเรื่องได้กล่าวถึงความสำคัญของน้ำมันปาล์มและการใช้พืชน้ำมันทดแทนไม่ใช่ทางเลือกที่ดี

VITA

NAME	Miss Yanika Lunrasri
DATE OF BIRTH	18 March 1990
PLACE OF BIRTH	Samutprakan
INSTITUTIONS ATTENDED	Chulalongkorn University, Bangkok (Class of 2015) Master of Education: Faculty of Education Major in Teaching English as a Foreign Language (GPA =3.95) Khon Kaen University (KKU), Khon Kaen (Class of 2012) Bachelor of Arts: Faculty of Humanities and Social Sciences Major in Business English (GPA = 3.87)
PUBLICATION	1) Lunrasri, Y., Tangdhanakanond, K., & Pasiphol, S. (in press). Effects of prompting type and learning achievement on reading literacy of ninth graders. Kasetsart Journal of Social Sciences. 2) Lunrasri, Y., Tangdhanakanond, K., & Pasiphol, S. (in press). Item analysis of multiple-choice reading literacy instruments using item response theory. Journal of Education Naresuan University. 3) Lunrasri, Y., Wechviriyakul, W., & Lawthong, N. (in press). Validation of the Career Adapt-Abilities Scale-short form for preservice teachers. Journal of Education Naresuan University. [in Thai] 4) Lunrasri, Y., & Gajaseni, C. (2014). Washback effects of the Ordinary National Educational Test on English language learning as perceived by Grade 9 students. An Online Journal of Education, 9(4), 226-240.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University