



การพัฒนาแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อ
ความเข้าใจภาษาอังกฤษ



วิไลภรณ์ คำมัน

วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้

ปีการศึกษา 2566

ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

การพัฒนาแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อ
ความเข้าใจภาษาอังกฤษ



วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนครสวรรค์
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้
ปีการศึกษา 2566
ลิขสิทธิ์เป็นของมหาวิทยาลัยนครสวรรค์

วิทยานิพนธ์ เรื่อง "การพัฒนาแบบสอบถามแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการ
อ่านเพื่อความเข้าใจภาษาอังกฤษ"
ของ วิไลภรณ์ คำมั่น
ได้รับการพิจารณาให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการสอบวิทยานิพนธ์
(รองศาสตราจารย์ ดร.พัชรี จันทรเพ็ญ)

..... ประธานที่ปรึกษาวิทยานิพนธ์
(รองศาสตราจารย์ ดร.สำราญ มีแจ้ง)

..... กรรมการที่ปรึกษาวิทยานิพนธ์
(รองศาสตราจารย์ ดร.น้ำทิพย์ งามอาภาภิรักษ์)

..... กรรมการที่ปรึกษาวิทยานิพนธ์
(ศาสตราจารย์ ดร.ปกรณ์ ประจันบาน)

..... กรรมการผู้ทรงคุณวุฒิภายใน
(รองศาสตราจารย์ ดร.กฤษฎากาญจน์ โต้พิทักษ์)

อนุมัติ

.....
(รองศาสตราจารย์ ดร.กรรองกาญจน์ ชูทิพย์)
คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง	การพัฒนาแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญา ความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ
ผู้วิจัย	วิไลภรณ์ คำมัน
ประธานที่ปรึกษา	รองศาสตราจารย์ ดร.สำราญ มีแจ่ม
กรรมการที่ปรึกษา	รองศาสตราจารย์ ดร.น้ำทิพย์ ่องอาจวานิชย์ ศาสตราจารย์ ดร.ปกรณ์ ประจันบาน
ประเภทสารนิพนธ์	วิทยานิพนธ์ ปริญญาโท. นวัตกรรมทางการวัดผลการเรียนรู้, มหาวิทยาลัย นเรศวร, 2566
คำสำคัญ	แบบสอบแบบทดสอบเพื่อวินิจฉัย, การอ่านเพื่อความเข้าใจ, การวินิจฉัยทาง ปัญญา, โมเดล G-DINA

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์ 2 ข้อ คือ 1) เพื่อสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ 2) เพื่อวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ ตัวอย่างวิจัยกลุ่มใช้วินิจฉัยคือนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 396 คน ได้มาจากการสุ่มแบบหลายขั้นตอน เครื่องมือคือแบบสอบแบบทดสอบเพื่อวินิจฉัยทางปัญญาความสามารถด้านการอ่านเพื่อความเข้าใจภาษาอังกฤษ วิเคราะห์ข้อมูลด้วยค่าสถิติเชิงบรรยายคือ ความถี่ ร้อยละ และการวิเคราะห์ G-DINA MODEL

ผลการวิจัย สรุปได้ดังนี้ 1) แบบสอบแบบทดสอบเพื่อวินิจฉัยทางปัญญาความสามารถด้านการอ่านเพื่อความเข้าใจภาษาอังกฤษที่พัฒนาขึ้นประกอบด้วย 6 คุณลักษณะ มีลักษณะเป็นชุดข้อสอบทดสอบจำนวน 4 ทดสอบแบบเลือกตอบ 4 ตัวเลือก โดยมีเนื้อหาตามกรอบมาตรฐานทางภาษา CEFR ระดับ A2 (Beginner) แบบสอบฉบับร่างจำนวน 42 ข้อ มีความตรงเชิงเนื้อหา (IOC) อยู่ระหว่าง 0.60 ถึง 1.00 เมื่อตรวจสอบคุณภาพด้วยทฤษฎีทดสอบดั้งเดิมผ่านการตรวจสอบคุณภาพจำนวน 33 ข้อ โดยมีค่าความยากอยู่ระหว่าง (p) 0.25 ถึง 0.70 ค่าอำนาจจำแนก (r) อยู่ระหว่าง 0.26 ถึง 0.87 ความเที่ยงผ่านเกณฑ์ (KR-20 = 0.9014) ค่าอิทธิพลทดสอบมีขนาดเล็กและไม่มีนัยสำคัญ เมื่อตรวจสอบคุณภาพด้วยทฤษฎีตอบสนองข้อสอบ ผ่านการตรวจสอบคุณภาพจำนวน 18 ข้อ มีค่าพารามิเตอร์ความยากอยู่ระหว่าง -0.29 ถึง 1.19 ค่าพารามิเตอร์อำนาจ

จำแนกอยู่ระหว่าง 0.82 ถึง 2.50 และค่าสัมประสิทธิ์การเดาอยู่ระหว่าง 0.12 ถึง 0.28 2) โมเดล G-DINA มีความสอดคล้องเชิงสัมพัทธ์ เมื่อเปรียบเทียบกับค่า-2LL, AIC, BIC กับโมเดล CDM พบว่าต่ำกว่า มีความถูกต้องของการวินิจฉัยผ่านเกณฑ์ทุกด้านและทุกคุณลักษณะ มีค่าพารามิเตอร์การเดา และค่าพารามิเตอร์ความสะเพร่าผ่านเกณฑ์เกือบทุกข้อ ค่าดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ (RMSEA) ส่วนใหญ่อยู่ในระดับต่ำ(ค่าเฉลี่ย RMSEA ผ่านเกณฑ์) ค่าดัชนีอำนาจจำแนกของข้อสอบรายข้อ (IDI) ผ่านเกณฑ์ทุกข้อ ผลการวิเคราะห์สามารถแบ่งกลุ่มแฝงของผู้สอบที่มีความสามารถต่างกันได้ ผลการออกแบบ PROFILE พบว่า “Reading Comprehension Profile” มีลักษณะเป็นสารสนเทศผลการวินิจฉัยย้อนกลับความสามารถด้านการอ่านเพื่อความเข้าใจภาษาอังกฤษหลังจากทำการทดสอบประกอบด้วย 6 ส่วน คือ 1.คะแนนดิบจากการทดสอบแบบข้อต่อข้อ 2.รูปแบบความรู้ของผู้สอบ 3. สารสนเทศผลการวินิจฉัยย้อนกลับเป็นค่าความน่าจะเป็น (Probability) ของการรอบรู้คุณลักษณะนั้น 4. ชื่อ PROFILE 5. สารสนเทศเชิงบรรยายจุดแข็ง และ 6. สารสนเทศข้อเสนอแนะ



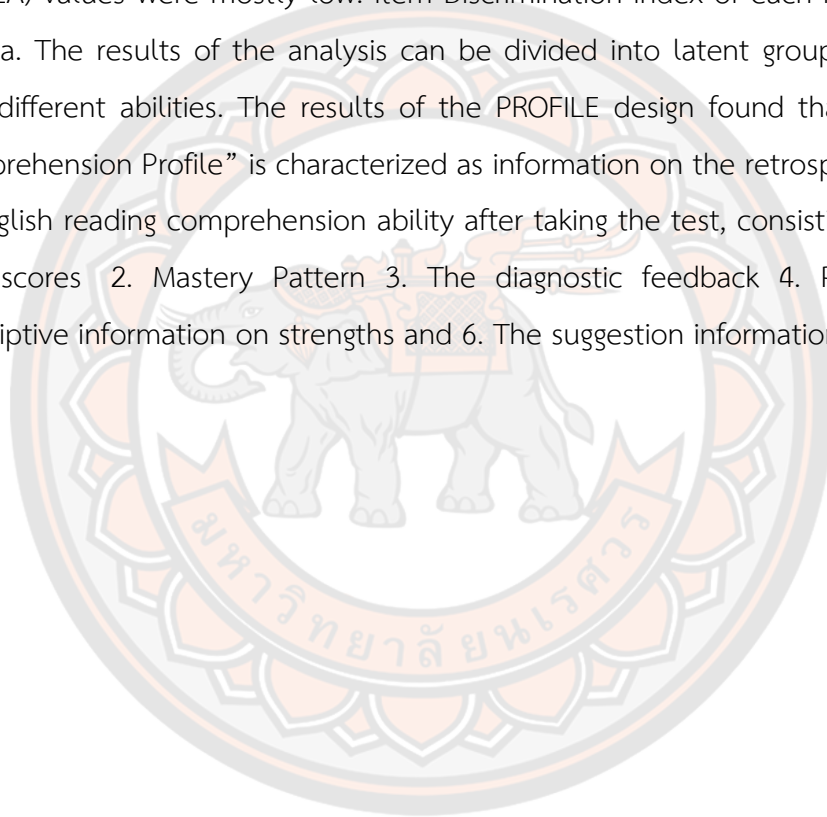
Title	THE DEVELOPMENT OF TESTLET-TEST FOR COGNITIVE DIAGNOSIS ON ENGLISH READING COMPREHENSION
Author	Wilaiporn Khamman
Advisor	Associate Professor Samran Mejang, Ed.D
Co-Advisor	Associate Professor Namthip Ongardwanich, Ed.D. Professor Pakorn Prachanban, Ed.D
Academic Paper	Ph.D. Dissertation in Innovation of Learning Measurement - (Type 2.1), Naresuan University, 2023
Keywords	Testlet Test, Reading Comprehension, Cognitive Diagnosis, G-DINA Model

ABSTRACT

This research has two objectives: 1) to create and examine the quality of a testlet test to diagnose the cognitive abilities of reading comprehension in English of Mathayom 3 students, and 2) to diagnose the abilities of Reading for English comprehension of Mathayom 3 students and designing an information profile from retrospective diagnosis. The diagnostic research sample consisted of 396 Mathayom 3 students, obtained from multi-stage random sampling. The instrument is a testlet test for the cognitive diagnosis of English reading comprehension ability. Data were analyzed using descriptive statistics such as frequency, percentage, and G-DINA MODEL analysis.

The results of the research can be summarized as follows: 1) The developed testlet test for intellectual diagnosis of reading comprehension ability in English consists of 6 attributes. The diagnostic test is the set of four testlets with 4 multiple choice. The content is based on the CEFR, level A2 (Beginner). The draft test has 42 items and the content validity (IOC) is between 0.60 and 1.00. When quality checked with the original test theory, 33 questions pass the quality check, with difficulty values between (p) 0.25 and 0.70. The discrimination index (r) ranged from 0.26 to 0.87. Reliability passed the criterion (KR-20 = 0.9014). Testlet effect is small

and insignificant. Checking quality with Items Response Theory, 18 items passed. The difficulty parameter ranged from -0.29 to 1.19, the discrimination parameter ranged from 0.82 to 2.50, and the guessing coefficient ranged from 0.12 to 0.28. 2) The G-DINA model has relative consistency. Comparing the values of -2LL, AIC, BIC with the CDM model, it was found to be lower, with diagnostic accuracy passing every criterion and every feature. There are guess parameters. And the carelessness parameters passed almost every criteria. The parameter estimation error index (RMSEA) values were mostly low. Item Discrimination Index of each items passed all criteria. The results of the analysis can be divided into latent groups of test takers with different abilities. The results of the PROFILE design found that the “Reading Comprehension Profile” is characterized as information on the retrospective diagnosis of English reading comprehension ability after taking the test, consisting of 6 parts: 1. Raw scores 2. Mastery Pattern 3. The diagnostic feedback 4. Profile name 5. Descriptive information on strengths and 6. The suggestion information



ประกาศคุณูปการ

ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงในความอดทนและเมตตาของรองศาสตราจารย์ ดร. สำราญ มีแจ้งประธานที่ปรึกษาวิทยานิพนธ์ ที่เมตตาติดตาม ให้ความช่วยเหลือ และให้คำแนะนำตลอดระยะเวลาในการทำวิทยานิพนธ์ฉบับนี้ ขอกราบขอบพระคุณกรรมการที่ปรึกษาวิทยานิพนธ์รวมทั้ง 2 ท่าน ศาสตราจารย์ ดร.ปกรณ์ ประจันบาน ที่จุดประกายความรู้ในประเด็นวิจัยและกรุณาสละเวลาให้คำแนะนำผู้วิจัยเสมอมา และรองศาสตราจารย์ ดร.น้ำทิพย์ งามอาภาวิชย์ ที่ให้ความช่วยเหลือสนับสนุนด้านความรู้ในการใช้โปรแกรมวิเคราะห์ทางสถิติในงานวิจัยฉบับนี้ และขอกราบขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์อันประกอบไปด้วย ประธานกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร.พัชรี จันท์เพ็ง กรรมการผู้ทรงคุณวุฒิภายใน รองศาสตราจารย์ ดร.กฤษยาภาณุจัน โตพิทักษ์ รองศาสตราจารย์ ดร.สำราญ มีแจ้ง ศาสตราจารย์ ดร. ปกรณ์ ประจันบาน และรองศาสตราจารย์ ดร.น้ำทิพย์ งามอาภาวิชย์ ที่ได้กรุณาให้คำแนะนำ ตลอดจนเติมเต็ม และแก้ไขจุดอ่อนของวิทยานิพนธ์ด้วยความเอาใจใส่และเมตตา จนทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วง

กราบขอบพระคุณ ผู้เชี่ยวชาญตรวจสอบคุณภาพเครื่องมือวิจัย ที่กรุณาสละเวลาประเมินเครื่องมือวิจัยด้วยความตั้งใจและอดทน ทำให้ผู้วิจัยได้ข้อมูลนำไปสร้างเครื่องมือที่มีประสิทธิภาพ ผู้วิจัยขอขอบพระคุณท่านผู้อำนวยการและคณะครูสังกัดสำนักงานเขตพื้นที่มัธยมศึกษาเพชรบูรณ์ ที่กรุณาสละเวลาให้ความอนุเคราะห์ในการจัดสอบเป็นอย่างดี และขอบคุณนักเรียนในโรงเรียนกลุ่มตัวอย่างทุกคนที่ให้ความร่วมมือ

กราบขอบพระคุณคณาจารย์สาขาวิชานวัตกรรมทางการวัดผลการเรียนรู้ คณะศึกษาศาสตร์ มหาวิทยาลัยนครสวรรค์ทุกท่าน ที่ประสิทธิ์ประสาทวิชาความรู้ และขอขอบคุณ คุณณัฐธิดา แจ๊ดล้อม ที่กรุณาให้คำแนะนำช่วยเหลือผู้วิจัยตลอดระยะเวลาการศึกษา และเจ้าหน้าที่ของสาขาคุณสุกัญญา สร้อยสุวรรณ ที่ให้ความช่วยเหลือเหนือสิ่งอื่นใดขอกราบขอบพระคุณบิดามารดา และครอบครัวที่เป็นกำลัง และให้การสนับสนุนในทุกด้านอย่างดีที่สุดเสมอมา

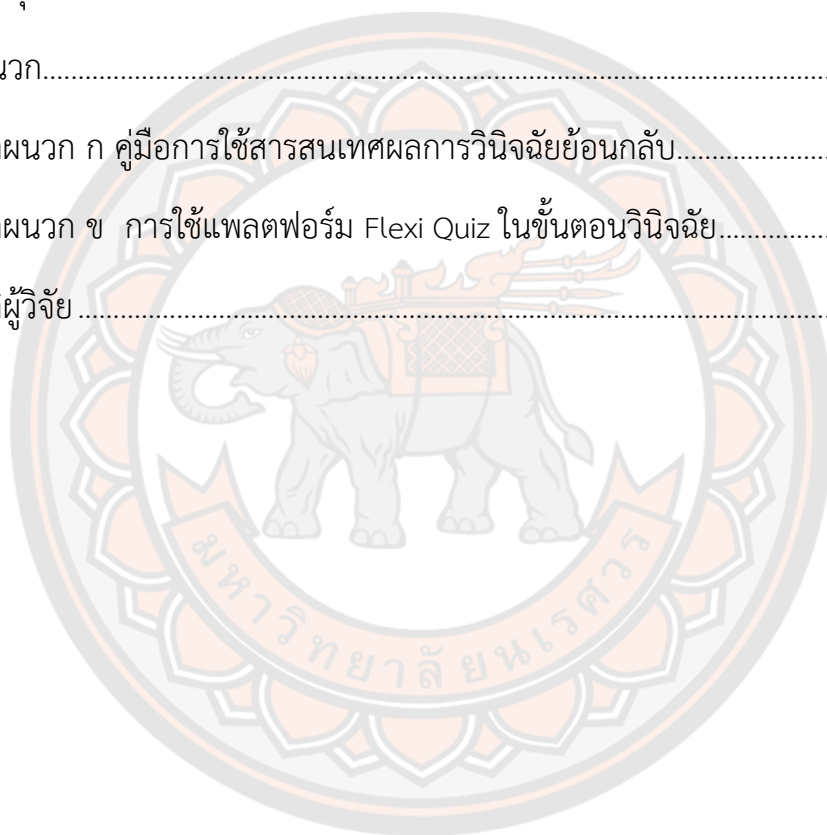
คุณค่าและคุณประโยชน์อันพึงจะมีจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบและอุทิศแต่ผู้ที่มีพระคุณทุกท่าน ผู้วิจัยหวังเป็นอย่างยิ่งว่า งานวิจัยนี้จะเป็นประโยชน์ต่อบุคลากรทางการศึกษา หรือผู้ที่สนใจบ้าง ไม่มากก็น้อย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ค
บทคัดย่อภาษาอังกฤษ.....	จ
ประกาศคุุณูปการ.....	ช
สารบัญ.....	ซ
สารบัญตาราง.....	ฅ
สารบัญภาพ.....	๗
บทที่ 1 บทนำ.....	17
ความเป็นมาและความสำคัญของปัญหา.....	17
คำถามการวิจัย.....	28
วัตถุประสงค์การวิจัย.....	28
ความสำคัญของการวิจัย.....	28
ขอบเขตของการวิจัย.....	29
1. ขอบเขตด้านแหล่งข้อมูล.....	29
2. ขอบเขตด้านเนื้อหา.....	29
3. ขอบเขตด้านตัวแปร.....	31
นิยามศัพท์เฉพาะ.....	33
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	36
ตอนที่ 1 มโนทัศน์เกี่ยวกับความสามารถการอ่านเพื่อความเข้าใจ (Reading Comprehension).....	37

ตอนที่ 2 มโนทัศน์เกี่ยวกับการประเมินเชิงวินิจัยทางพุทธิปัญญา	93
ตอนที่ 3 มโนทัศน์เกี่ยวกับข้อสอบแบบทดสอบ.....	165
ตอนที่ 4 มโนทัศน์เกี่ยวกับการให้ข้อมูลย้อนกลับ	190
ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง.....	207
ตอนที่ 6 กรอบแนวคิดการวิจัย.....	217
บทที่ 3 วิธีดำเนินการวิจัย.....	220
ระยะที่ 1 การสร้างและการตรวจสอบคุณภาพของแบบสอบแบบทดสอบเพื่อวินิจัย เชิงพุทธิปัญญา ความสามารถในการอ่านเพื่อความเข้าใจของนักเรียนชั้นมัธยมศึกษา ปีที่ 3	221
ระยะที่ 2 การทดลองและวินิจัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และการพัฒนา Skill profile สารสนเทศจากการ วินิจัยแบบย้อนกลับ	239
บทที่ 4 ผลการวิจัย.....	248
ตอนที่ 1 ผลการสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจัยทาง พุทธิปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้น มัธยมศึกษาปีที่ 3	248
ตอนที่ 2 ผลการทดลองและวินิจัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และผลการออกแบบ Profile สารสนเทศจากการ วินิจัยแบบย้อนกลับ	269
บทที่ 5 สรุปผลการวิจัย	1
สรุปผลการวิจัย.....	1
ตอนที่ 1 ผลการสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจัย ทางพุทธิปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของ นักเรียนชั้นมัธยมศึกษาปีที่ 3	1

ตอนที่ 2 ผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของ นักเรียนชั้นมัธยมศึกษาปีที่ 3 และผลการออกแบบ Profile สารสนเทศจาก การวินิจฉัยแบบย้อนกลับ	4
อภิปรายผลการวิจัย	8
ข้อเสนอแนะจากการวิจัย	11
บรรณานุกรม	13
ภาคผนวก	19
ภาคผนวก ก คู่มือการใช้สารสนเทศผลการวินิจฉัยย้อนกลับ	20
ภาคผนวก ข การใช้แพลตฟอร์ม Flexi Quiz ในขั้นตอนวินิจฉัย	27
ประวัติผู้วิจัย	32



สารบัญตาราง

หน้า

ตาราง 1 ผลการสังเคราะห์คุณลักษณะความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ	75
ตาราง 2 สรุปผลการสังเคราะห์คุณลักษณะความสามารถการอ่านเพื่อความเข้าใจ	83
ตาราง 3 แสดงเป้าหมายการพัฒนาระดับความสามารถทางภาษาของผู้เรียน	87
ตาราง 4 แสดงกรอบเป้าหมายระดับความสามารถในการใช้ภาษา CEFR.....	88
ตาราง 5 แสดงคำอธิบายความสามารถด้านการอ่านภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3.....	91
ตาราง 6 ความสอดคล้องระหว่าง CEFR กับหลักสูตรฯ กลุ่มสาระการเรียนรู้ภาษาต่างประเทศชั้น ม.3.....	92
ตาราง 7 ชื่อเรียกโมเดลวินิจฉัยทางพุทธิปัญญาจำแนกตามความหมายแฝง	99
ตาราง 8 แสดงชั้นแฝงที่ผู้สอบจะถูกจำแนกในกรณีมี 3 คุณลักษณะที่ต้องใช้ในการตอบแบบสอบ	112
ตาราง 9 การจัดกลุ่มของโมเดลการวินิจฉัยเชิงจำแนก (DCMs) (Rupp and Templin, 2007, p.56)	113
ตาราง 10 ประเภทของโมเดลเชิงวินิจฉัยทางพุทธิปัญญาแบบจำแนกเป็น 3 กลุ่ม	116
ตาราง 11 การเปรียบเทียบโมเดลวินิจฉัยทางพุทธิปัญญาเชิงจำแนก	121
ตาราง 12 แสดงตัวอย่างผลการตรวจสอบคุณภาพของเมตริกซ์ควิตั้งสองครั้ง	136
ตาราง 13 แสดงตัวอย่างคำสั่งขอ output ในโปรแกรม R	140
ตาราง 14 ตัวอย่างการพิจารณาค่าอำนาจจำแนกของข้อสอบ	148
ตาราง 15 แสดงขนาดของอิทธิพลทดสอบ Confirmatory DIMTEST Analyses	176

ตาราง 16 แสดงขนาดของเทสต์เลทในการศึกษาเกี่ยวกับความสามารถการอ่านเพื่อความเข้าใจ.....	177
ตาราง 17 แสดงความยาวของบทอ่านภาษาอังกฤษที่พบในงานวิจัย 3 เรื่อง.....	180
ตาราง 18 แสดงจำนวนประชากรซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำแนกตามขนาดโรงเรียน.....	222
ตาราง 19 แสดงกลุ่มตัวอย่างที่ใช้ในการตรวจสอบคุณภาพของแบบสอบวินิจฉัยฉบับร่าง.....	224
ตาราง 20 แสดงผลการสุ่มตัวอย่างที่ใช้ในการตรวจสอบคุณภาพแบบสอบวินิจฉัยฉบับร่าง 2.....	226
ตาราง 21 แสดงจำนวนตัวอย่างที่ใช้ในการวินิจฉัยและจัดอันดับความสามารถการอ่านเพื่อความเข้าใจ.....	240
ตาราง 22 โครงสร้างแบบสอบวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ.....	251
ตาราง 23 ผลการตรวจสอบความสอดคล้องระหว่างข้อคำถามและคุณลักษณะของความสามารถการอ่าน.....	252
ตาราง 24 แสดงผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยทฤษฎีการทดสอบดั้งเดิม.....	255
ตาราง 25 แสดงผลการตรวจสอบคุณภาพข้อคำถามรายข้อด้วยทฤษฎีการตอบสนองข้อสอบ (IRT 3PL).....	258
ตาราง 26 แสดงจำนวนเทสต์เลท เนื้อหาตามกรอบ CEFR ชื่อเรื่อง และจำนวนข้อคำถามของแบบสอบ.....	259
ตาราง 27 ผลการสร้างแบบสอบแบบเทสต์เลทเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3.....	261

ตาราง 28 แสดงผลการตอบแบบสอบวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจ (คะแนนเต็ม 18)	270
ตาราง 29 แสดงเมตริกซ์ควิแนะนำ (Suggested Q-matrix) โดยวิธี Stepwise Wald Test	271
ตาราง 30 ผลการตรวจสอบความตรงของเมตริกซ์ควิจากการทดลองวินิจฉัยด้วยโมเดล G-DINA MODEL.....	272
ตาราง 31 แสดงผลการวิเคราะห์เปรียบเทียบประสิทธิภาพความสอดคล้องเชิงสัมพัทธ์ของ โมเดลการวินิจฉัย G-DINA (Test Level Fit Statistics).....	273
ตาราง 32 แสดงผลการวิเคราะห์เปรียบเทียบประสิทธิภาพความสอดคล้องเชิงสัมพัทธ์ของ โมเดลการวินิจฉัย G-DINA (Test Level Fit Statistics).....	274
ตาราง 33 ค่าระดับความแม่นยำในการจำแนกของแบบสอบรายฉบับและรายคุณลักษณะ	275
ตาราง 34 แสดงผลการวิเคราะห์ความตรงและความเที่ยงเชิงจำแนกของแบบสอบราย คุณลักษณะ	276
ตาราง 35 แสดงผลการวิเคราะห์คุณภาพของข้อคำถามด้วยโมเดล G-DINA.....	277
ตาราง 36 แสดงผลการวินิจฉัยรายคุณลักษณะ (Attribute Prevalence).....	279
ตาราง 37 แสดงผลการวิเคราะห์ผลการจำแนกกลุ่มผู้สอบตามคุณลักษณะความสามารถ 64 แบบแผน	281
ตาราง 38 แสดงตัวอย่างผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจของนักเรียน ในระดับชั้นมัธยมศึกษาปีที่ 3.....	285

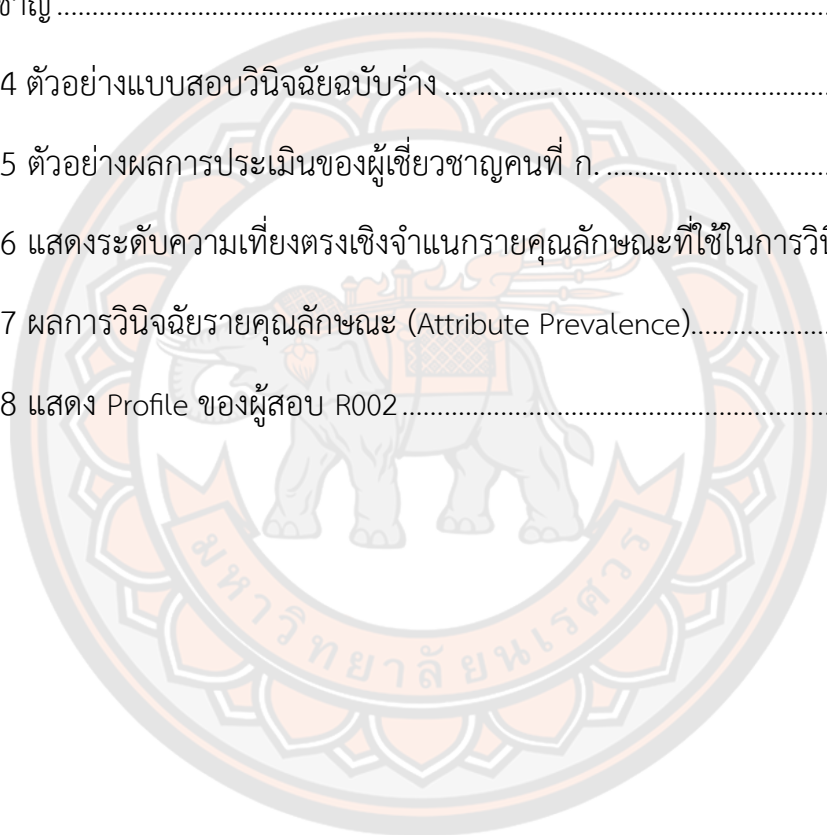
สารบัญภาพ

หน้า

ภาพ 1 ผลการทดสอบ O-NET ของนักเรียน ม.3 เปรียบเทียบกับผลสอบระดับประเทศ..	19
ภาพ 2 แสดงผลการทดสอบระดับชาตินักเรียนชั้น ม.3 จังหวัดเพชรบูรณ์.....	20
ภาพ 3 ชั้นผังสำหรับการวินิจฉัย 3 คุณลักษณะ.....	24
ภาพ 4 กระบวนการอ่านแบบ Bottom-up และ top-down.....	51
ภาพ 5 กระบวนการอ่านแบบปฏิสัมพันธ์.....	52
ภาพ 6 รูปแบบการอ่านภาษาอังกฤษเป็นภาษาที่สองของโคดี้.....	54
ภาพ 7 โมเดลมันนี่.....	66
ภาพ 8 โครงสร้างโมเดลวินิจฉัยเชิงจำแนก.....	100
ภาพ 9 โครงสร้างที่มีลักษณะพหุมิติของ CDMs.....	102
ภาพ 10 โมเดลวินิจฉัยทางพุทธิปัญญา.....	111
ภาพ 11 ชั้นผังสำหรับการวินิจฉัย 3 คุณลักษณะ.....	112
ภาพ 12 แสดงผลการพิจารณาเมตริกซ์คิวของผู้เชี่ยวชาญ.....	134
ภาพ 13 แสดงตัวอย่างเมตริกซ์คิวของข้อสอบ 20 ข้อ ที่ใช้วัดคุณลักษณะ 6 คุณลักษณะ.....	135
ภาพ 14 แสดงค่า PVAF ของข้อคำถามข้อที่ 3 และ 6 ก่อนปรับ.....	138
ภาพ 15 แสดง Final Q-matrix.....	139
ภาพ 16 แสดงตัวอย่างผลการวิเคราะห์สารสนเทศของแบบสอบ.....	142
ภาพ 17 แสดงตัวอย่างค่าความน่าจะเป็นในการตอบข้อสอบถูก.....	142
ภาพ 18 แสดงการใช้โปรแกรมคำนวณค่าความเชื่อมั่นด้วยวิธี KR-20.....	149

ภาพ 19 เมตริกซ์ควิที่แนะนำ (Suggest Q-matrix).....	150
ภาพ 20 แสดงความสัมพันธ์ระหว่างคะแนนสอบกับเกณฑ์	152
ภาพ 21 แสดงตัวอย่าง out put การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์โดยใช้โปรแกรม SPSS.....	153
ภาพ 22 โคว้นลักษณะข้อสอบแบบ 1 พารามิเตอร์ของตัวอย่างข้อสอบ 4 ข้อ	155
ภาพ 23 โคว้นลักษณะข้อสอบแบบ 2 พารามิเตอร์ของตัวอย่างข้อสอบ 4 ข้อ	156
ภาพ 24 โคว้นลักษณะข้อสอบแบบ 3 พารามิเตอร์ของตัวอย่างข้อสอบ 4 ข้อ	158
ภาพ 25 โมเดลการวัดแบบดั้งเดิมและโมเดลการตอบสนองข้อสอบ.....	168
ภาพ 26 พหุมิติระหว่างข้อคำถามและพหุมิติภายในข้อคำถาม	171
ภาพ 27 ลักษณะโครงสร้างของโมเดล UIRT, MIRT และ Bi-factor.....	174
ภาพ 28 โมเดลทดสอบทดสอบมาตรฐานสำหรับวัดความเข้าใจในการอ่าน EFL Test.....	175
ภาพ 29 ข้อสอบแบบทดสอบเลขวิชาภาษาไทยระดับชั้นมัธยมศึกษาปีที่ 3 (สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน), 2563).....	182
ภาพ 30 ข้อสอบแบบทดสอบเลขวิชาภาษาอังกฤษชั้นมัธยมศึกษาปีที่ 3	183
ภาพ 31 ข้อสอบเชิงสถานการณ์ข้อสอบ PISA วิชาคณิตศาสตร์	184
ภาพ 32 แสดงบัตรรายงาน DiagnOsis Scoring Reports.....	200
ภาพ 33 ตัวอย่าง Pattern Occurrence Probabilities การอ่านภาษาอังกฤษ.....	201
ภาพ 34 Component Skills Profile ความสามารถภาษาอังกฤษ 4 ด้าน.....	202
ภาพ 35 Skills Profile ความสามารถภาษาอังกฤษเพื่อการสื่อสารด้านการอ่าน.....	203
ภาพ 36 ตัวอย่าง Skill profile วิชาคณิตศาสตร์เรื่องความรู้สึกเชิงจำนวน	204
ภาพ 37 ตัวอย่าง A Diagnostic Score Report.....	205
ภาพ 38 แสดงตัวอย่างแบบรายงานคะแนนเชิงวินิจฉัย	206

ภาพ 39 กรอบแนวคิดการวิจัย.....	219
ภาพ 40 แสดงขั้นตอนของการศึกษาในระยะที่ 1.....	220
ภาพ 41 แสดงขั้นตอนของการศึกษาในระยะที่ 2.....	221
ภาพ 42 ตัวอย่างเครื่องมือที่ใช้เก็บรวบรวมข้อมูลในการวิจัยสำหรับผู้เชี่ยวชาญ.....	229
ภาพ 43 ตัวอย่างแบบประเมินความสอดคล้องของข้อความและคุณลักษณะสำหรับผู้เชี่ยวชาญ.....	230
ภาพ 44 ตัวอย่างแบบสอบถามวิจัยฉบับร่าง.....	231
ภาพ 45 ตัวอย่างผลการประเมินของผู้เชี่ยวชาญคนที่ ก.....	235
ภาพ 46 แสดงระดับความเที่ยงตรงเชิงจำแนกรายคุณลักษณะที่ใช้ในการวิจัย.....	275
ภาพ 47 ผลการวิจัยรายคุณลักษณะ (Attribute Prevalence).....	280
ภาพ 48 แสดง Profile ของผู้สอบ R002.....	287



บทที่ 1

บทนำ

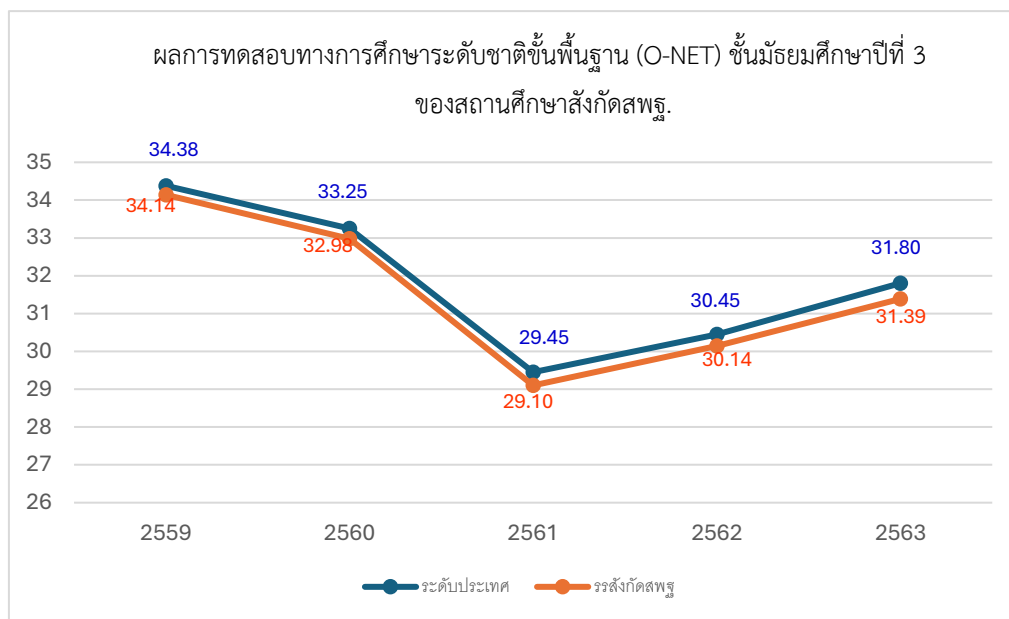
ความเป็นมาและความสำคัญของปัญหา

การอ่านเป็นสิ่งจำเป็นสำหรับคนทั่วไป โดยเฉพาะผู้ที่อยู่ในวัยศึกษาเล่าเรียน เพราะปัจจุบันเป็นยุคของข่าวสาร ซึ่งข้อมูลความรู้ในทุกด้านได้เผยแพร่ในรูปแบบของหนังสือและสิ่งตีพิมพ์ต่าง ๆ ผู้อ่านหนังสือมากจะได้รับคุณค่าทางปัญญา ผู้ที่อ่านด้วยความเข้าใจย่อมเกิดความรู้มากกว่า แต่ผู้มีความสามารถในการอ่านภาษาอังกฤษย่อมเกิดความรู้ มีความทันสมัย และมีโอกาสได้ข้อมูลที่แวดวงกว้าง จากผลความก้าวหน้าของเทคโนโลยีสมัยใหม่ส่งผลให้ผู้เรียนสามารถศึกษาค้นคว้าหาความรู้ได้ด้วยตนเองบนอินเทอร์เน็ต อาทิ บทความวิชาการ ในรูปแบบ Digital หรือ E-Book ที่การแลกเปลี่ยนความคิดผ่านการอ่านข้อมูลจาก Web blog รวมถึงแหล่งเรียนรู้บนอินเทอร์เน็ตซึ่งส่วนใหญ่ใช้ภาษาอังกฤษเป็นสื่อกลางในการเผยแพร่ข้อมูล เนื่องจากเป็นภาษาสากลที่มีการใช้อย่างแพร่หลายมากที่สุดภาษาหนึ่ง จึงนับได้ว่าการอ่านภาษาอังกฤษให้เข้าใจเป็นหนึ่งในความสามารถและทักษะที่สำคัญของผู้เรียนยุคสังคมแห่งการเรียนรู้ในศตวรรษที่ 21 และเป็นเครื่องมือที่ทุกคนใช้เพื่อเข้าถึงแหล่งของความรู้ที่ต้องการ (วิสาข์ จิตวิวัฒน์, 2543, หน้า 1; วรณีย์ โสมประยูร, 2544, หน้า 120; สุนันทา มั่นเศรษฐวิทย์, 2545, หน้า 1; พรทิพย์ ชาตะรัตน์, 2546, หน้า 59; ฉัตรปวีณ อัมภา, 2555, หน้า 1; จันทร์ฉาย เศษแสงศรี, 2562; Anderson, 1999; เบญจมาศ ศิริกมลเสถียร และคณะ, 2560, อ้างถึงใน ธนกร สุวรรณพฤติ, 2563, หน้า 196) สอดคล้องกับแนวทางการจัดการเรียนรู้ตามหลักสูตรแกนกลางของกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ พุทธศักราช 2551 ที่เน้นการจัดการเรียนการสอนเพื่อการสื่อสารต้องอาศัยทักษะย่อยของการอ่านภาษาอังกฤษในเรื่องความรู้คำศัพท์และโครงสร้างทางภาษาเป็นแนวทางและพื้นฐานในการพัฒนาทักษะอื่นๆ เช่น การฟังพูด หรือการเขียน และสอดคล้องกับ อติสรา กาศีป (1997 หน้า 51-52) กล่าวว่า แม้การอ่านจะเป็นทักษะสำคัญในการเรียนรู้ภาษาต่างประเทศและได้รับการบรรจุอยู่ในหลักสูตรการเรียนการสอนทั้งในโรงเรียนและมหาวิทยาลัย แต่การอ่านยังสร้างปัญหาให้นักเรียนไทยในการทำความเข้าใจข้อมูลจากการอ่าน กระบวนการในการทำความเข้าใจรวมทั้งกระบวนการแก้ปัญหาเมื่อเกิดความไม่เข้าใจยังคงเป็นปริศนาเนื่องจากความคิดในขณะที่อ่านเป็นกระบวนการที่เกิดขึ้นในสมอง บุคคลภายนอกไม่สามารถล่วงรู้ได้ จึงเป็นปมท้าทายสำหรับนักวิจัยด้านการอ่านให้เกิดแรงจูงใจที่จะแก้ปมปริศนาปมใหญ่นี้

ความสามารถการอ่านเพื่อความเข้าใจ (Reading Comprehension) ถือได้ว่าเป็นทักษะกระบวนการทางสมองที่ประกอบด้วยการถอดรหัสความหมายของภาษา (Decoding) เป็นการ

วิเคราะห์ตัวอักษร คำ กลุ่มคำ ประโยคของเนื้อหาที่อ่าน และความเข้าใจ (Comprehension) เป็น การสร้างความหมายจากการถอดรหัสตัวอักษร คำ กลุ่มคำ ประโยคของเนื้อหาที่อ่านโดยอาศัยพื้น ความรู้เดิม (Background Knowledge) 2 แบบได้แก่ 1) ความรู้ทางภาษา (Linguistic Knowledge) เช่น คำศัพท์ ไวยากรณ์ และ 2) ความรู้ทั่วไป (General Knowledge) เช่น วัฒนธรรม ความเป็นไป ของโลกปัจจุบัน (Villamin, 1994; Smith, 2004; Hammerberg ,2004; Harmer, 2006; Brown, 2007, อ้างถึงใน ธนกร สุวรรณพถุฒิ, 2563, หน้า 196-197; สถาบันภาษาอังกฤษ สำนักงาน คณะกรรมการการศึกษาขั้นพื้นฐาน, 2560: หน้า 97) การเปลี่ยนแปลงอย่างรวดเร็วของเทคโนโลยีทำ ให้ธรรมชาติของการอ่านได้เปลี่ยนแปลงตามไปด้วย ในอดีตหากนักเรียนต้องการหาข้อมูลหรือหา คำตอบใด ๆ ก็จะต้องค้นหาอ่านตามหนังสือหรือสารานุกรมและมักจะเชื่อว่าคำตอบที่พบนั้นถูกต้อง แล้ว แต่ตอนนี้หากนักเรียนสืบค้นข้อมูลบนสื่อออนไลน์ก็จะได้คำตอบนับล้านคำตอบทำให้ความ รวดเร็วในการเผยแพร่ข้อมูลจะมาก่อนคุณภาพของข้อมูล (PISA Thailand, 2564) ดังนั้นการสอน ของครูในปัจจุบันจึงไม่ใช่เพียงแค่การสอนนักเรียนเพื่อให้ความรู้อีกต่อไป แต่ควรเป็นการช่วย ให้ นักเรียนพัฒนาทักษะและความสามารถในการวิเคราะห์ข้อมูลจากการอ่านอีกด้วย สอดคล้องกับ การศึกษาของฉัตรปวีณ อัมภา (2555: หน้า 122) พบว่า งานวิจัยด้านการสอนอ่านภาษาอังกฤษของ นักเรียนชั้นมัธยมศึกษาตอนต้นเทคนิคและวิธีการสอนที่นิยมใช้มากที่สุดคือวิธีการสอนอ่าน ภาษาอังกฤษแบบมุ่งประสบการณ์และวิธีการสอนอ่านภาษาอังกฤษเพื่อความเข้าใจ นักวิจัยในยุค ศตวรรษที่ 21 จึงมุ่งหาปัจจัยเชิงสาเหตุและวิธีในการพัฒนาความสามารถทางการอ่าน ชัยวัฒน์ บวรวัฒน์เศรษฐ์ (2557) ได้ศึกษาปัจจัยเชิงสาเหตุที่มีอิทธิพลต่อความสามารถในการอ่านภาษาอังกฤษ เพื่อความเข้าใจของนักเรียนระดับมัธยมศึกษาตอนต้น มนัสวี ดวงลอย (2558) ปัจจัยที่มีผลต่อปัญหา ในการอ่านภาษาอังกฤษของนักศึกษาของมหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ จันทร์ฉาย เศษ แสงศรี (2562) ปัจจัยเชิงสาเหตุที่ส่งผลต่อความสามารถในการอ่านภาษาอังกฤษอย่างมีวิจารณญาณ วราภรณ์ พรหมอินทร์ (2561) ใช้กลวิธีการอ่านแบบร่วมมือพัฒนาความสามารถในการอ่าน ภาษาอังกฤษเพื่อความเข้าใจ ลียานา พันธพงศ์ธรรม (2561) การใช้กลวิธีการอ่านภาพเพื่อส่งเสริม ทักษะการอ่านและการคิดอย่างมีวิจารณญาณ

แม้ว่านักวิชาการ ครูผู้สอน หรือหน่วยงานทางการศึกษาของประเทศได้ให้ความสำคัญกับ ประสิทธิภาพด้านความสามารถทางภาษาอังกฤษของนักเรียนไทยมาอย่างต่อเนื่องยาวนาน กลับ พบว่ายังไม่ประสบความสำเร็จเท่าที่ควร เห็นได้จากผลการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ชั้นมัธยมศึกษาปีที่ 3 ของสถานศึกษาสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน เปรียบเทียบกับผลการทดสอบระดับประเทศย้อนหลัง 5 ปี ตั้งแต่ปีการศึกษา 2563 ถึง 2559 หลังจากนั้น (ปี 2564 ถึงปัจจุบัน) ผู้เรียนสอบตามความสนใจมีค่าเฉลี่ยแสดงในภาพ 1

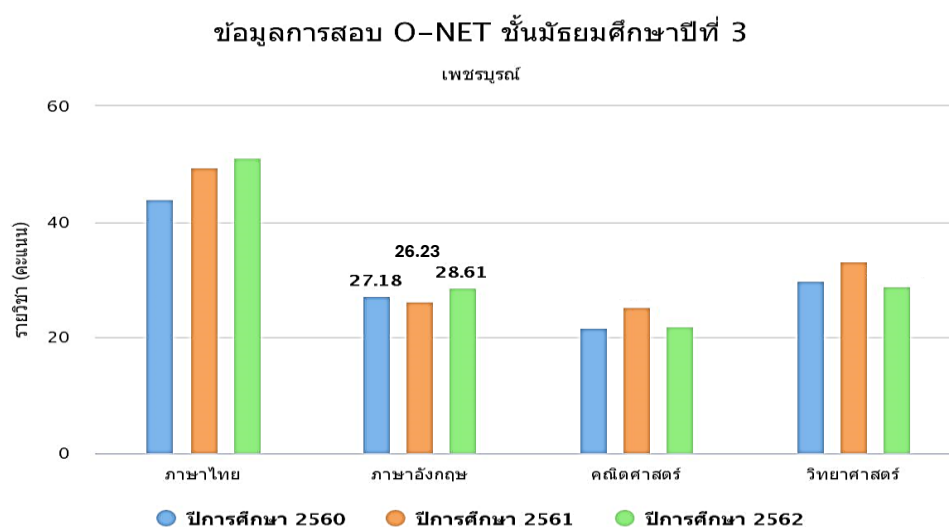


ภาพ 1 ผลการทดสอบ O-NET ของนักเรียน ม.3 เปรียบเทียบกับผลสอบระดับประเทศ

จากภาพ 1 แสดงให้เห็นอย่างชัดเจนว่าในช่วง 5 ปีการศึกษาที่ผ่านมา ผลการทดสอบ O-NET รายวิชาภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ศึกษาอยู่ในสถานศึกษาสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานอยู่ในระดับที่ต่ำกว่าผลทดสอบระดับประเทศมาโดยตลอด นอกจากนี้ปัญหาดังกล่าวพบว่ามีแง่มุมที่น่าสนใจเกี่ยวกับผลการสอบ PISA ความสามารถด้านการอ่านของนักเรียนอายุ 15 ปี (ในประเทศไทยเป็นนักเรียนในระดับชั้นมัธยมศึกษาปีที่ 3) ครั้งล่าสุดมีคะแนนเฉลี่ย 393 คะแนน ซึ่งต่ำกว่าการประเมินในทุกรอบที่ผ่านมา โดยสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) เป็นผู้ดำเนินงานในประเทศไทยในฐานะศูนย์แห่งชาติได้ดำเนินการจัดสอบเมื่อเดือนสิงหาคม พ.ศ. 2561 ที่ผ่านมา (สำนักข่าว BBC NEWS ไทย, 2564) สสวท. ระบุว่า ผลการประเมินการอ่านของนักเรียนไทยตั้งแต่ปี 2000 จนถึงครั้งล่าสุดในปี 2018 มีแนวโน้มลดลงอย่างต่อเนื่อง มีข้อสังเกตที่สำคัญคือนักเรียนไทยทั้งกลุ่มที่มีคะแนนสูงและกลุ่มที่มีคะแนนต่ำล้วนมีจุดอ่อนอยู่ที่ผลการประเมินในด้านความฉลาดรู้เรื่องการอ่าน (Reading Literacy) อย่างไรก็ตามถึงจะมีการแปลแบบทดสอบจากต้นฉบับที่เป็นภาษาอังกฤษให้เป็นภาษาไทย แต่ยังไม่ปรากฏความสัมพันธ์ระหว่างภาษากับวัฒนธรรมของเจ้าของภาษา (ภาษาอังกฤษ) ทุกเรื่องหรือทุกบทอ่านที่ถูกบรรจุอยู่ในแบบทดสอบ หมายความว่าหากนักเรียนมีพื้นความรู้เดิมเกี่ยวกับวัฒนธรรมของเจ้าของภาษา (ภาษาอังกฤษ) จะช่วยให้สามารถทำความเข้าใจเรื่องที่อ่านได้ดีกว่า สอดคล้องกับทฤษฎีวัฒนธรรมต้นทนิส (2550, หน้า 121) สรุปไว้ว่า การอ่านเป็นกระบวนการที่ผู้อ่านสร้างความหมายโดยการปฏิสัมพันธ์ระหว่างพื้นความรู้เดิมของผู้อ่าน บทอ่าน และบริบทที่เกี่ยวข้องซึ่งรวมความคิด ความเชื่อ และประสบการณ์ในอดีต โดยที่ผู้อ่านนำมาใช้ในการสร้างความหมายให้แก่เรื่องที่อ่าน ดังนั้นครูผู้สอน

และผู้มีส่วนเกี่ยวข้องทางการศึกษาทุกภาคส่วนควรให้ความสำคัญกับการพัฒนาความสามารถการอ่านเพื่อความเข้าใจของนักเรียนเป็นอันดับต้น

ในช่วงก่อนสถานการณ์โรคระบาดโควิด-19 จนถึงปัจจุบันหน่วยงานทางการศึกษาในจังหวัดเพชรบูรณ์ดำเนินการปฏิรูปการเรียนการสอนภาษาอังกฤษอย่างต่อเนื่องแต่กลับพบว่าผลการประเมินการอ่านของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนในจังหวัดเพชรบูรณ์ยังไม่เป็นที่น่าพอใจ ภาพที่ 2 แสดงคะแนนเฉลี่ยจากการทดสอบระดับชาติขั้นพื้นฐาน (O-NET)



ภาพ 2 แสดงผลการทดสอบระดับชาตินักเรียนชั้น ม.3 จังหวัดเพชรบูรณ์

จากภาพ 2 พบว่าคะแนนเฉลี่ยวิชาภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของสถานศึกษาสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานในจังหวัดเพชรบูรณ์ปีการศึกษา 2560 ถึง 2562 เท่ากับ 27.18, 26.23 และ 28.61 ตามลำดับ โดยต่ำกว่าร้อยละ 50 และต่ำกว่าคะแนนเฉลี่ยระดับประเทศอีกด้วย เนื่องจากข้อสอบดังกล่าววัดทักษะการคิดบนฐานของสิ่งที่อ่านเป็นหลัก (สัมพันธ์ พันธุ์พุกฤษ, 2556, อ้างถึงใน ลียานา พันธพงศ์ธรรม, 2561, หน้า 2) ในประเด็นดังกล่าว กันนิง (Gunning, 1992) ได้อธิบายไว้ว่า การที่ผู้เรียนไม่ประสบความสำเร็จในการอ่านภาษาอังกฤษ อาจเกิดจากปัจจัยหลายประการ เช่น ปัญหาด้านการเรียนรู้ ขาดทักษะความเข้าใจในระดับตัวอักษรหรือคำศัพท์ เป็นต้น ซึ่งเป็นสาเหตุต่อเนื่องที่ก่อให้เกิดปัญหาในการพัฒนาทักษะย่อยอื่นตามมา ด้วยเหตุนี้กระทรวงศึกษาธิการ (2557) หาแนวทางพัฒนาผู้เรียนและได้ประกาศนโยบายการปฏิรูปการเรียนการสอนภาษาอังกฤษโดยได้ใช้กรอบมาตรฐานการประเมินความสามารถทางด้านภาษาต่างประเทศหรือ กรอบอ้างอิง CEFR (Common European Framework of Reference for Languages) เป็นกรอบความคิดหลักในการจัดการเรียนการสอนภาษาอังกฤษของประเทศไทยเพื่อ

ยกระดับความสามารถทางด้านภาษาของนักเรียนไทยให้มีความพร้อมทั้งทางวิชาการและทักษะการสื่อสารทางภาษาในระดับที่ใช้งานได้ (Working Knowledge)

จากความสำคัญของปัญหาทำให้หลายฝ่ายร่วมกันหาแนวทางแก้ไขและพัฒนาผู้เรียนอย่างเร่งด่วน เนื่องจากเป็นช่วงชั้นสุดท้ายของการศึกษาภาคบังคับ อีกทั้งยังเป็นผู้เรียนที่มีช่วงอายุ 15 ปี ในโครงการประเมินผลนักเรียนนานาชาติ (PISA) ปัญหาดังกล่าวอาจมีสาเหตุมาจากปัจจัยต่าง ๆ ดังที่ Bloom (1976, อ้างถึงใน จันทร์ฉาย เศษแสงศรี, 2562, หน้า 3) ได้สรุปไว้ว่า ปัจจัยที่มีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนมี 3 ด้าน คือ ด้านพุทธิพิสัย ประกอบด้วย ความรู้พื้นฐานของผู้เรียน ความถนัดทางการเรียน เนื่องจากผู้เรียนที่มีพื้นฐานความรู้เดิมช่วยให้ได้รับความรู้ใหม่ได้มากกว่าผู้เรียนที่ไม่มีความรู้เดิมหรือใช้ประสบการณ์เดิมช่วยความเข้าใจสิ่งที่อ่าน ส่วนด้านที่สองคือ ปัจจัยด้านจิตพิสัย ประกอบด้วยเจตคติต่อวิชาที่เรียน ความตระหนักรู้ แรงจูงใจ และความภาคภูมิใจในตนเอง และปัจจัยด้านที่สามประกอบด้วย คุณภาพการสอน การเสริมแรงจากครู การแก้ไขข้อบกพร่อง และการให้ข้อมูลย้อนกลับ ดังนั้นการวัดผลด้วยแบบทดสอบทั่วไปที่มีจุดมุ่งหมายเพื่อจัดอันดับ หรือตัดสินผลการเรียนโดยสรุปรวมจึงไม่ตอบโจทยอันท้าทายข้อนี้

งานวิจัยในช่วงต่อมานักวิชาการด้านภาษาจึงให้ความสำคัญกับการศึกษาเกี่ยวกับการประเมินเชิงวินิจฉัยทางพุทธิปัญญาเนื่องจากมีผลการวิจัยที่สามารถบ่งชี้จุดแข็งจุดอ่อนของคุณลักษณะ (Attribute) หรือทักษะย่อย (Sub-skill) ที่มุ่งประเมิน รวมทั้งความสามารถการอ่านเพื่อความเข้าใจซึ่งเป็นกระบวนการที่เกิดขึ้นในสมองของผู้เรียนตามทฤษฎีโครงสร้างความรู้ (Schema Theory) ความรู้ของคนเราได้รับการรวบรวมเป็นหน่วยๆ เรียกว่าโครงสร้างความรู้ (Schemata) ซึ่งเป็นหน่วยความรู้ที่ได้รับการเรียบเรียงขึ้นจากความรู้ที่เรามีเกี่ยวกับผู้คนสิ่งของสถานที่เหตุการณ์ หากความรู้ที่ได้รับมาใหม่มีลักษณะผสมกลมกลืนจะถูกนำมาเพิ่มเข้ากับโครงสร้างเดิมที่มีอยู่ แต่ถ้าไม่อยู่ในลักษณะผสมกลมกลืนจะถูกสร้างเป็นโครงสร้างใหม่เรียงซ้อนกันคล้ายบล็อกซึ่งจะมีความซับซ้อนมากขึ้น (Piaget, 1970; Rumelhart, 1984, อ้างถึงใน ฉัตรปวีณ อ่ำภา, 2555, หน้า 11-12) ผู้เรียนที่มีความสามารถทางภาษาโดยรวมในระดับดี นำไปสู่การพัฒนาทักษะการอ่านที่ดีและสามารถใช้เทคนิคหรือกลวิธีต่าง ๆ ที่เหมาะสมเพื่อช่วยให้เข้าใจเนื้อหาที่อ่านง่ายขึ้น ดังนั้นความสามารถทางภาษามีความสัมพันธ์เชิงบวกกับความสามารถในการอ่านและการใช้กลวิธีการอ่าน (Kletzien, 1991, อ้างถึงใน ธนกร สุวรรณพถุณี, 2563, หน้า 207 ; Sheorey & Mokhatari, 2001) ความท้าทายจึงเป็นการตอบคำถามว่าผู้อ่านที่เก่งควรมีคุณลักษณะหรือทักษะย่อยอะไรบ้าง มีจุดแข็งและจุดอ่อนด้านใด หากครูหรือผู้เกี่ยวข้องได้รับสารสนเทศดังกล่าวรายบุคคลเสียแต่เนิ่น เช่น ระดับชั้นมัธยมศึกษาตอนต้น จะทำให้การหาแนวทางแก้ไขหรือพัฒนาจุดต่อเป็นไปอย่างทันทั่วทั้งที่ เนื่องจากการพัฒนาให้ผู้เรียนเกิดคุณลักษณะหรือทักษะทางภาษาเป็นเรื่องที่ต้องใช้ระยะเวลา

การประเมินเพื่อวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostic Assessment: CDA) เป็นรูปแบบการประเมินที่สามารถให้สารสนเทศเกี่ยวกับผลการสอบได้ทั้งในแง่ของผลการประเมินโดยสรุปในภาพรวม และผลการประเมินซึ่งเป็นข้อวินิจฉัยเกี่ยวกับจุดแข็งและจุดอ่อนของผู้เรียนในสิ่งที่ประเมินได้เป็นรายบุคคลแตกต่างจากการประเมินในรูปแบบเดิมที่ใช้กันโดยทั่วไป ซึ่งให้สารสนเทศเกี่ยวกับผู้สอบได้เฉพาะในภาพรวมหรือในรูปแบบคะแนน โดยมีวัตถุประสงค์สองประการคือเพื่อระบุประเด็นที่นักเรียนยังไม่รอบรู้และให้ข้อเสนอแนะเกี่ยวกับสาเหตุหรือเหตุผลที่เป็นไปได้ว่าเพราะเหตุใดนักเรียนถึงยังไม่รอบรู้ในประเด็นนั้นๆ เพื่อประโยชน์ในการพัฒนานักเรียนได้ตรงจุดที่บกพร่อง (Leighton & Gierl, 2007: 3, อ้างถึงใน สุมาลี มีสกุล, 2558: หน้า 5; Nitko and Brookhart, 2014: p.84,อ้างถึงใน วรางคณา ก้อมน้อย, 2563: หน้า 4) สอดคล้องกับงานเขียนของ สโนว์ และโลแมน (Snow & Lohman, 1989) ได้แสดงให้เห็นอย่างชัดเจนในการนำแนวคิดทฤษฎีพุทธิปัญญาเข้ามาใช้ในการวัดทางการศึกษาเรื่องการประยุกต์ใช้ศาสตร์ด้านจิตวิทยาทางปัญญา (Cognitively psychology) โดยชี้ให้เห็นประเด็นที่เป็นประโยชน์สำหรับการวัดทางการศึกษา ได้แก่ 1) ให้ข้อมูลการวิเคราะห์ข้อสอบที่อธิบายถึงโครงสร้างของความเข้าใจ 2) สร้างความกระจ่างในเป้าหมายของการสอบในเรื่องของความรู้และทักษะหรือกระบวนการเชิงพุทธิปัญญาที่ชี้ให้เห็นถึงความรอบรู้และความเข้าใจ และ 3) ยกระดับทฤษฎีของความถนัด ผลสัมฤทธิ์ และการเรียนรู้ (Leighton & Gierl, 2007, อ้างถึงใน แคทรียา แสงดาวเทียน, 2562: หน้า 1)

ในปัจจุบันมีการตื่นตัวในการนำการประเมินเพื่อวินิจฉัยทางพุทธิปัญญามาใช้ในการประเมินทางการศึกษามากขึ้น แต่ก็นับได้ว่ายังคงอยู่ในช่วงเริ่มต้น ทั้งที่ต้นกำเนิดของการประเมินในรูปแบบนี้มีมาเป็นเวลานาน เลตันและเกิร์ล (Leighton & Gierl, 2007: 343–345, อ้างถึงใน สุมาลี มีสกุล, 2558: หน้า 6) กล่าวว่า การประเมินเพื่อวินิจฉัยทางพุทธิปัญญาที่พบในปัจจุบันจะนำแบบทดสอบมาตรฐานทางการศึกษาที่มีอยู่แล้ว มากำหนดรหัสที่เป็นทักษะความสามารถทางพุทธิปัญญาที่ใช้ในการหาคำตอบกับข้อคำถามแต่ละข้อแล้วจึงทำการวิเคราะห์ข้อมูล (Post-hoc หรือ Retrofitting) ซึ่งวิธีการนี้นับว่าเป็นวิธีการทำให้ประสิทธิภาพในเรื่องของผลการวิเคราะห์ข้อมูลต่ำ เนื่องจากวิธีการที่ใช้ในการสร้างแบบทดสอบเพื่อประเมินทางการศึกษาโดยทั่วไปถูกกำหนดโดยขอบเขตของเนื้อหา มากกว่าสมรรถนะที่เป็นโมเดลการคิดแก้ปัญหาในแบบทดสอบทำให้การวิเคราะห์ข้อมูลในเชิงวินิจฉัยจากแบบทดสอบดังกล่าวได้ผลที่คลาดเคลื่อนและไม่สอดคล้องกับคุณลักษณะที่เป็นแบบการคิดแก้ปัญหาที่ถูกต้องของผู้เรียน ในประเด็นดังกล่าว สุมาลี มีสกุล (2558, หน้า 7) เสนอแนะว่า การนำองค์ความรู้ของการศึกษาวิจัยเกี่ยวกับการนำการประเมินเพื่อวินิจฉัยทางพุทธิปัญญามาใช้ในการประเมินทางการศึกษาทั้งของไทยและต่างประเทศนั้นยังคงทำอยู่ในวงจำกัดและคุณลักษณะที่ทำการวินิจฉัยยังไม่มีหลากหลายมากนัก ทั้งที่ในปัจจุบันมีความก้าวหน้าเกี่ยวกับวิธีการและรูปแบบที่

ใช้ในการประเมินดังกล่าวค่อนข้างมาก จากข้อเสนอแนะดังกล่าวผู้วิจัยจึงสนใจพัฒนาแบบสอบตามกระบวนการออกแบบการประเมินเชิงวินิจฉัยทางพุทธิปัญญาเพื่อใช้ในการศึกษาครั้งนี้

โมเดลวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostic Models: CDM) เป็นโมเดลทางสถิติในการวิเคราะห์ที่สามารถสกัดสารสนเทศเชิงวินิจฉัยจากข้อมูล นำสารสนเทศที่ได้มาพัฒนาหรือกำหนดข้อสอบที่วัดองค์ความรู้และทักษะที่เฉพาะเจาะจง และใช้โมเดลนี้โดยตรงกับการวิเคราะห์ทางจิตของรูปแบบการตอบสนองต่อข้อสอบของผู้ตอบ เพื่อส่งเสริมการวินิจฉัยที่เฉพาะเจาะจง (Leighton and Gierl, 2007 อ้างถึงใน Gierl and Zhou, 2008, อ้างถึงใน วราจคณา ก้อมน้อย, 2563: หน้า 4-5) ในปัจจุบันมีวิธีการที่ใช้สำหรับการประเมินเชิงวินิจฉัยทางพุทธิปัญญา (CDA) หลายวิธี แต่มีวิธีการสำคัญที่ใช้ซึ่งมีอยู่ทั้งหมด 5 วิธี คือ การใช้โมเดลการวินิจฉัยเชิงจำแนก (Diagnostic Classification Models : DCMs), การใช้แนวคิดการประเมินความรู้ (Knowledge-Space Theory) การใช้แนวคิดการประเมินกฎ (Rule-Space Method : RSM), การใช้โมเดลลำดับขั้นของคุณลักษณะ (Attribute Hierarchy Method : AHM) และวิธีการวิเคราะห์จำแนกกลุ่ม (Clustering Algorithms) ซึ่งวิธีการใช้โมเดลการวินิจฉัยเชิงจำแนก (Diagnostic Classification Models : DCMs) เป็นวิธีการประมาณค่าคุณลักษณะของผู้สอบวิธีการเดียวที่มีพื้นฐานการคำนวณอยู่บนความน่าจะเป็นในการตอบของผู้สอบ ส่วนวิธีการอีก 4 วิธีที่เหลือไม่ได้มีพื้นฐานการคำนวณอยู่บนความน่าจะเป็นในการตอบของผู้สอบและมีข้อจำกัดในการใช้บางประการทำให้วิธีการดังกล่าวไม่เป็นที่นิยมนำมาใช้ในปัจจุบัน (Rupp, Templin; & Henson, 2010: 96-108, อ้างถึงใน สุมาลี มีสกุล, 2558: หน้า 5) หากกล่าวในเชิงสถิติ โมเดลวินิจฉัยทางพุทธิปัญญา (CDMs) คือ ความน่าจะเป็นที่ได้จากยืนยันโมเดลตัวแปรแฝงพหุมิติด้วยโครงสร้างน้ำหนักองค์ประกอบเชิงเดี่ยวหรือน้ำหนักองค์ประกอบเชิงซ้อนที่เหมาะสมสำหรับการสร้างโมเดลตัวแปรตอบสนองแบบแบ่งกลุ่ม (Categorical Response Variables) และมีตัวแปรพยากรณ์เป็นตัวแปรแฝงแบบแบ่งกลุ่ม (Categorical Latent Predictor Variables) ที่นำไปสู่การสร้างชั้นแฝง (Rupp and Templin, 2007, p.8)

โมเดลวินิจฉัยทางพุทธิปัญญา (CDMs) เป็นวิธีการแบบสหวิทยาการ (Interdisciplinary Approach) โดยมีการเชื่อมโยงระหว่างจิตวิทยาพุทธิปัญญาและการวิเคราะห์ทางสถิติไว้ด้วยกัน ใช้สำรวจความสัมพันธ์ระหว่างกระบวนการทางจิตและกลวิธีที่ใช้ในการแก้ปัญหาข้อหนึ่งๆ ในแบบทดสอบ และการตอบสนองของผู้ตอบผ่านการวิเคราะห์ทางสถิติที่ซับซ้อนมีลักษณะเป็นโมเดลชั้นแฝง (Latent Class Model) ซึ่งจำแนกผู้ตอบแบบทดสอบเป็นชั้นแฝง (Latent Classes) ขึ้นอยู่กับการตอบสนองต่อแบบทดสอบของผู้ตอบโดยชั้นแฝงนั้นเป็นชั้นแฝงที่มีจำนวนจำกัด (Restricted Latent Classes) ขึ้นอยู่กับจำนวนของคุณลักษณะ (Attributes) ที่ใช้ในการตอบข้อสอบของแบบทดสอบ หากแบบทดสอบนั้นมียุคคุณลักษณะที่ต้องใช้จำนวน K คุณลักษณะ ผู้ตอบจะถูกจำแนกเป็น 2^K ชั้นแฝง โดยเลข 2 หมายถึงจำนวนของผลลัพธ์ที่เป็นไปได้สำหรับแต่ละคุณลักษณะ ซึ่ง

หมายถึง การรอบรู้ (Mastery) หรือไม่รอบรู้ (Non-mastery) ในคุณลักษณะนั้น ๆ (Ravand and Robitzsch, 2015, pp. 1-2, อ้างถึงใน วรจกณา ก้อมน้อย, 2563: หน้า 62) การคำนวณชั้นแฝงจะทำให้ทราบจำนวนโปรไฟล์แบบแผนการตอบที่เป็นไปได้ทั้งหมด ตัวอย่างเช่น แบบสอบชุดหนึ่งต้องใช้คุณลักษณะในการตอบแบบสอบจำนวน 3 คุณลักษณะ ($K=3$) จะได้ 2^3 นั่นคือ ผู้ตอบจะถูกจำแนกเป็น 8 ชั้นแฝง โดยที่ 1 หมายถึง รอบรู้ และ 0 หมายถึง ไม่รอบรู้แสดงในภาพที่ 3

$$\alpha = \{(0,0,0) (0,0,1) (0,1,0) (1,0,0) (0,1,1) (1,0,1) (1,1,0)(1,1,1)\}$$

ภาพ 3 ชั้นแฝงสำหรับการวินิจฉัย 3 คุณลักษณะ

ภาพ 3 แสดงให้เห็นจำนวนชั้นแฝง (Latent class) ของรูปแบบการเรียนรู้ที่อาจเกิดขึ้นจากการวินิจฉัยซึ่งผู้สอบจะถูกจัดเข้าชั้นแฝงชุดใดชั้นหนึ่ง ขึ้นอยู่กับความรอบรู้ของผู้สอบว่ามีความสามารถตามคุณลักษณะที่วัดจากแบบสอบชุดนั้นๆหรือไม่ (เอกราช ตีนาง, 2559: หน้า 4) ดังนั้นจึงสามารถนำมาใช้เพื่อประเมินจุดแข็ง (Strength) และจุดอ่อน (Weakness) ของนักเรียน ในแต่ละทักษะ (Skills) หรือคุณลักษณะ (Attributes) ที่จำเป็นต้องใช้ในการตอบแบบสอบได้ สารสนเทศจากการสอบจึงให้ทั้งข้อมูลการวัดการเรียนรู้ของผู้เรียนและพัฒนาการของผู้เรียน ใช้ปรับปรุงการเรียนการสอน และพัฒนาโครงการเพื่อแก้ไขปัญหาของนักเรียนบางคนและบางกลุ่ม (สังวรรณ ังดกระโทก, 2559)

เมื่อก้าวถึงวิธีการใช้โมเดลการวินิจฉัยเชิงจำแนก (Diagnostic Classification Models: DCMs) ในวรรณกรรมบางเล่มอาจเรียกว่าโมเดลวินิจฉัยทางพุทธิปัญญา (CDMs) ที่นำมาใช้ในการประเมินเพื่อวินิจฉัยทางพุทธิปัญญา พบว่าในปัจจุบันมีโมเดลที่ใช้วิเคราะห์ข้อมูลหลายโมเดลในแต่ละโมเดลมีสูตรในการคำนวณและวิธีประมาณค่าพารามิเตอร์ที่แตกต่างกัน The G-DINA model เป็นโมเดลที่ถูกพัฒนาขึ้นโดยเดอลาทอร์รี (De la Torre, 2011) เพื่อผ่อนปรนข้อจำกัดบางประการของโมเดลแบบชดเชยได้ และโมเดลแบบไม่ชดเชย G-DINA model หรือ Generalized deterministic input, noisy “and” gate model เป็นโมเดลเชิงวินิจฉัยทางพุทธิปัญญาแบบทั่วไป (General CDM) ที่มีการผ่อนปรนข้อตกลงเบื้องต้นเกี่ยวกับความน่าจะเป็นในการทำข้อสอบได้ โดยอนุญาตให้ในแบบทดสอบเดียวกัน มีการใช้ต่างโมเดลในการประมาณค่าแต่ละข้อ (Item-by-Item Model Estimation) และทดสอบสมมุติฐานได้ ขึ้นอยู่กับการออกแบบและเมตริกซ์ความสัมพันธ์ระหว่างข้อคำถามและคุณลักษณะหรือเมตริกซ์คิว (Q-matrix) และมีการประมาณค่าพารามิเตอร์โดยใช้อัลกอริทึม Marginalized Maximum Likelihood Estimation (MMLE) ซึ่งในการประมาณค่าวิธีนี้สามารถประมาณค่าพารามิเตอร์ได้รวดเร็วและมีประสิทธิภาพกว่าการประมาณค่าโดยอัลกอริทึม Markov

Chain Monte Carlo อีกทั้งการใช้โมเดลจีดีไอเอ็นเอ (G-DINA model) ยังสามารถช่วยลดรูปโมเดลจากโมเดลเต็มรูป และใช้การทดสอบวาลด์ (Wald test) ในการเปรียบเทียบโมเดลเต็มรูปและลดรูปสำหรับข้อที่มีพหุคุณลักษณะได้ ทำให้ในการประมาณค่าพารามิเตอร์สามารถประมาณค่าจากโมเดลลดรูปได้ (De la Torre, 2011, p.196, อ้างถึงใน วรางคณา ก้อมน้อย, 2563: หน้า 67) อีกทั้งยังเป็นโมเดลที่มีการประมาณค่าที่มีความยืดหยุ่นมากกว่าโมเดลอื่น โดยจะกำหนดให้ผู้สอบที่มีแบบแผนการคิดในการหาคำตอบของปัญหาแตกต่างกันมีค่าความน่าจะเป็นในการหาคำตอบได้อย่างถูกต้องไม่เท่ากัน ดังนั้นในการวิจัยนี้จึงเลือกใช้โมเดลจีดีไอเอ็นเอ (G-DINA model) ในการวินิจฉัยทางพุทธิปัญญาความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษ โดยเป็นกระบวนการที่ใช้ทักษะย่อย ๆ ร่วมกันในการทำความเข้าใจเพื่อให้ได้ความถูกต้องแม่นยำ ซึ่งข้อสอบเดี่ยว (Single Item) ไม่สามารถวัดทักษะที่ซับซ้อนได้อย่างครอบคลุม จึงมีแนวคิดเกี่ยวกับการจัดกลุ่มข้อสอบที่ใช้สิ่งเร้าร่วมกันและรวมกันอยู่ภายในแบบสอบฉบับเต็มโดยข้อสอบหนึ่งข้อสามารถวัดคุณลักษณะผู้สอบได้มากกว่าหนึ่งเรียกข้อสอบลักษณะนี้ว่าเทสต์เลท (Grabe & Stoller, 2002: p.17; Chen, 2014, p.11, อ้างถึงใน จงกล บัวแก้ว, 2560: หน้า 4)

ข้อสอบแบบเทสต์เลท (Testlet-test) เป็นลักษณะของข้อสอบเชิงสถานการณ์ที่พบได้บ่อยในทางภาษาและทางสังคมศาสตร์ ในการประมาณค่าความสามารถของบุคคลในความเข้าใจสิ่งสิ่งเร้าเดียวกัน เช่น บทความ กราฟ แผนภูมิ หรือตารางตัวเลข เป็นต้น โดยในการทำข้อสอบที่ใช้สถานการณ์การอ่านเหล่านี้ทำให้ผู้สอบใช้เวลาค่อนข้างมากถ้าจะตอบคำถามเพียงข้อเดียว ดังนั้นในแต่ละบทความจึงมีข้อความที่เกี่ยวข้องกับบทความเป็นเซตเดียวกัน อาจมีจำนวนข้อความ 3-5 ข้อหรือบางครั้ง 6-12 ข้อ โดยเรียงลำดับข้อความจะเรียงจากง่ายไปหายาก ซึ่งบทความจะมีความยาวมากกว่า 300 คำ (Wainer, Bradlow, & Wang, 2007: p. 60, อ้างถึงใน พิสิฐ พิณจสุกุล, 2563: หน้า 77; Wheadon, 2010: p. 106; อริสสา เตห์ลิ้ม, 2559: หน้า 2 ; จงกล บัวแก้ว, 2560: หน้า 4) ในการทำข้อสอบที่ใช้สถานการณ์ในการอ่านเหล่านี้ ใช้เวลาค่อนข้างมากถ้าจะตอบคำถามเพียงข้อเดียวและอาจเกิดสภาวะภาระทางปัญญา (Cognitive Load) ซึ่งจะเกิดขึ้นมากขึ้นอยู่กับระดับความยากง่ายของข้อมูลรวมไปถึงปริมาณข้อมูลและสื่อในการนำเสนอข้อมูล ซึ่งถ้าขณะเรียนรู้นั้นเกิดภาระทางปัญญามากเกินไป ก็จะทำให้ไม่เกิดการรับรู้ใด ๆ (กอบเกียรติ สระอุบล และพัลลภ พิริยะสุวรรณค์, 2557) ด้วยลักษณะเฉพาะของเทสต์เลท (Testlet-test) ที่สามารถวัดคุณลักษณะหลายด้านในข้อเดียวกัน จึงสามารถลดจำนวนข้อสอบที่ใช้ในการทดสอบได้ถึงร้อยละ 30-50 และใช้เวลาในการทำข้อสอบน้อยลงอีกด้วย (Senarat et al., 2013)

การที่ข้อสอบตอบสนองสิ่งเร้าหรือบทความเดียวกันทำให้ข้อสอบในรูปแบบของชุดเทสต์เลทนั้นอาจมีการพึ่งพาคำตอบกันในแต่ละข้อถือว่าเป็นการชี้นำคำตอบ ในประเด็นนี้หากลดจำนวนข้อสอบลงก็สามารถลดการชี้นำคำตอบซึ่งกันและกันของข้อสอบลงด้วย นอกจากนี้เนื่องจากข้อสอบ

หนึ่งข้อต้องใช้ความสามารถหลายอย่างในการตอบทำให้ข้อสอบมีความยากสูงซึ่งจะลดการชี้้นำคำตอบระหว่างข้อได้อีกระดับหนึ่ง (พิสิฐ พิณจสกุล, 2563: หน้า 81)

ปัญหาความไม่เป็นอิสระกันนั้นหากพิจารณาอย่างผิวเผิน แนวคิดเรื่องความเป็นอิสระของข้อคำถาม (IRT) อาจขัดแย้งกับแนวคิดเรื่องความสอดคล้องภายในของข้อคำถามตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ที่ระบุว่าข้อคำถามที่มีความสัมพันธ์กันมาก แสดงให้เห็นว่ามาตรวัดนั้นมีความเที่ยงในระดับสูง อย่างไรก็ตามเมื่อพิจารณาอย่างละเอียดจะพบว่า แนวคิดตามทฤษฎีทั้งสองนั้นมิได้ขัดแย้งกันเพราะความเป็นอิสระของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นความเป็นอิสระของข้อคำถามที่เกิดขึ้นภายหลังการควบคุมพารามิเตอร์บุคคลอย่างละเอียด คุณลักษณะแฝงภายในบุคคลต้องเป็นปัจจัยเดียวที่ทำให้การตอบข้อคำถามของผู้ทดสอบมีความแตกต่างกัน ความแตกต่างของการตอบข้อคำถามต้องไม่เกิดขึ้นจากความสัมพันธ์รูปแบบอื่น และพารามิเตอร์ข้อสอบที่มีอิทธิพลโดยตรงต่อการตอบข้อคำถามในโมเดลแล้ว (สุชาติ สกลกิจรุ่งโรจน์, 2558, หน้า 27-28) อริสฟา เตห์ลิ้ม (2559) กล่าวว่า อิทธิพลทดสอบ (Testlet Effect) เป็นผลกระทบของอิทธิพลแบบสุ่มที่เป็นปฏิสัมพันธ์ของการตอบข้อสอบกับทดสอบเลขซึ่งควรมีขนาดเล็ก (0.3-0.4) ข้อสอบทดสอบเลขที่มีทั้งที่เป็นปรนัยและอัตนัยจึงทำให้มีการสร้างแบบสอบมาตรฐานที่มีลักษณะทดสอบเลขมาใช้ในการประเมินทางการศึกษามากขึ้น โดยส่วนใหญ่ใช้รูปแบบข้อสอบแบบหลายตัวเลือก (Multiple Choice: MC) ซึ่งมีความเหมาะสมและเป็นประโยชน์เมื่อใช้ในการประเมินกับคนกลุ่มขนาดใหญ่ (Large Scale Assessment) การให้คะแนนมีความเป็นปรนัย ตรวจสอบให้คะแนนได้ง่าย ประหยัดเวลา และครอบคลุมเนื้อหา นอกจากนี้การใช้ชุดข้อสอบที่มีความสัมพันธ์กันภายในแบบสอบทำให้แบบสอบที่สร้างขึ้น มีความยืดหยุ่นมากขึ้นในการประเมินทางพุทธิพิสัยข้อสอบแบบทดสอบเลขวัดคุณลักษณะหลายด้านในข้อเดียวกัน ซึ่งข้อคำถามที่วัดหลายองค์ประกอบอาจเกิดโครงสร้างเป็นพหุมิติ (Multidimensional Structure)

ความสามารถทางภาษาศาสตร์นั้นถือว่าเป็นสิ่งที่ผสมผสาน มีความเกี่ยวข้องสัมพันธ์กันของความสามารถหลายด้านอย่างซับซ้อน (Wainer et al., 2000; De Mars, 2006, cited in Chen, 2014, p.11-12, อ้างถึงใน จงกล บัวแก้ว, 2560: หน้า 17; Wainer et al., 2007, p.110, อ้างถึงใน พิสิฐ พิณจสกุล, 2563: หน้า 77) การศึกษาของ Chen (2015) ทำการวิเคราะห์โครงสร้างการอ่านของ PISA เพื่อสำรวจความสัมพันธ์ของทักษะย่อยของการอ่านเพื่อความเข้าใจด้วยชุดของข้อสอบทดสอบเลขที่มีสิ่งเร้าร่วมกันเป็นบทอ่าน จำนวน 6 ทดสอบเลขด้วยข้อสอบ 20 ข้อ โดยใช้โมเดล G-DINA อย่างไรก็ตามการศึกษาของ Chen (2015) เป็นการนำข้อมูลที่มีอยู่แล้วมาวิเคราะห์ทางสถิติและแบบทดสอบที่ใช้กันไม่ได้สร้างขึ้นเพื่อการประเมินเชิงวินิจฉัยทางพุทธิปัญญา ด้วยเหตุนี้ผู้วิจัยจึงสนใจการนำข้อสอบชุดทดสอบเลขมาใช้ในการการศึกษาครั้งนี้ โดยทำการพัฒนาแบบสอบแบบทดสอบเลขด้วยกระบวนการออกแบบเพื่อประเมินเชิงวินิจฉัยทางพุทธิปัญญาด้วยโมเดลวินิจฉัยเชิงจำแนก CDMs

เลือกโมเดลจีดีไอเอ็นเอ (G-DINA) ซึ่งหากพบว่าสามารถแก้ปัญหาความไม่เป็นอิสระกันของข้อสอบในชุดทดสอบเดียวกันได้จะเป็นการเพิ่มแนวทางในการศึกษาวิจัยให้กว้างขวางมากขึ้น ทั้งนี้การให้ข้อมูลเกี่ยวกับผลการตอบหรือผลการกระทำแก่ผู้เรียนทำให้ผู้เรียนทราบว่าเขาทำได้ดีเพียงใดเพื่อที่จะได้พัฒนาและปรับปรุงจุดบกพร่องของตนเอง

การให้ข้อมูลย้อนกลับ (Feedback) เกี่ยวกับผลการตอบข้อสอบหรือผลการกระทำแก่ผู้เรียน เพื่อให้ผู้เรียนทราบความก้าวหน้าของตนเอง หากการตอบสนองของเขาไม่ถูกต้องหรือยังมีจุดบกพร่องเมื่อได้ทราบคำตอบที่ถูกต้อง ก็สามารถแก้ไขความเข้าใจได้ทันทีที่ไม่ทำให้เรียนรู้ไปอย่างผิดๆ หากผู้เรียนไม่ได้รับทราบผลการตอบอาจเกิดความไม่แน่ใจว่าตนเองทำถูกต้องมากน้อยเพียงใดทำให้เกิดความท้อแท้เบื่อหน่ายได้ การให้ข้อมูล ย้อนกลับจึงมีความสำคัญต่อการเรียนการสอนมากในด้านเป็นตัวเสริมแรงและเป็นกลไกสำคัญใน การเรียนรู้ของผู้เรียน จะเห็นได้ว่าการให้ข้อมูลย้อนกลับของครูเป็นสิ่งสำคัญในการเรียนภาษาอังกฤษของนักเรียน (ไชยยศ เรืองสุวรรณ, 2533, อ้างถึงใน ฮัชมีนี ยูโซะ, 2555, หน้า 4) สอดคล้องกับ Lee & Sawaki (2009) สรุปว่า การวินิจฉัยทางภาษาควรให้ข้อมูลย้อนกลับเกี่ยวกับระดับความรู้และทักษะในปัจจุบันของผู้เรียน เพื่อแก้ไขจุดอ่อนในการเรียนภาษาอังกฤษขั้นตอนในการวินิจฉัยทางพุทธิปัญญาในการอ่านภาษาอังกฤษเริ่มจากการระบุคุณลักษณะ จากนั้นสร้าง Q-matrix และวิเคราะห์ข้อมูลด้วย CDMs ขั้นตอนสุดท้ายเป็นการรายงานผลการสอบและให้ข้อมูลวินิจฉัยย้อนกลับด้วย Profile ที่ได้มาจากการวิเคราะห์แบบแผนการตอบ

ความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษเป็นทักษะพื้นฐานของสมรรถนะความฉลาดรู้ด้านการอ่านและการอ่านในระดับสูง แต่ผลการทดสอบทางการอ่านของนักเรียนยังไม่ประสบผลสำเร็จเท่าที่ควร ทั้งที่มีการศึกษาวิจัยเกี่ยวกับแนวคิดวิธีการพัฒนาความสามารถการอ่านและแบบทดสอบมาตรฐานเป็นจำนวนมาก อาจเนื่องมาจากแบบสอบมาตรฐานที่มีอยู่ให้ผลการประเมินสรุปรวมในรูปแบบคะแนนเฉลี่ย เพื่อตัดสินว่าสอบผ่านหรือไม่ผ่านหรือให้เกรดเท่านั้นทำให้ผู้เรียนที่มีผลการเรียนปานกลางจนถึงต่ำอาจพลาดโอกาสในการพัฒนาการเรียนรู้ กอปรกับการทำแบบสอบภาษาอังกฤษที่มีบทอ่านยาว ๆ และตอบข้อคำถามเพียงข้อเดียวทำให้ใช้เวลานานและอาจสร้างภาวะสมองล้าแก่ผู้เรียนได้

จากความสำคัญของปัญหาที่กล่าวมา ผู้วิจัยจึงสนใจศึกษาการพัฒนาแบบสอบแบบทดสอบเพื่อใช้ในการประเมินเชิงวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่กำลังศึกษาอยู่ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเพชรบูรณ์ ด้วยคุณสมบัติพิเศษของข้อสอบแบบทดสอบซึ่งเป็นกลุ่มของข้อสอบที่ตอบสนองสิ่งเร้าเดียวกัน โดยใช้จำนวนข้อสอบน้อย ๆ แต่สามารถวัดได้อย่างครอบคลุมมากกว่าหนึ่งคุณลักษณะจึงช่วยให้ผู้สอบไม่เกิดภาวะเหนื่อยล้าจากการทำข้อสอบจำนวนมากๆ สร้างขึ้นตามกรอบการประเมินเชิงวินิจฉัยทางพุทธิปัญญาแบบ CDS's Embretson (1998) โดยใช้โมเดลวินิจฉัยเชิงพุทธิ

ปัญญา G-DINA model ได้รูปแบบความรอบรู้ของผู้สอบรายบุคคล และให้ข้อมูลวินิจฉัยแบบย้อนกลับด้วย Profile ที่แสดงสารสนเทศเกี่ยวกับจุดแข็งและจุดอ่อนของผู้สอบรายบุคคล เพื่อให้ผู้เกี่ยวข้องนำไปใช้ประโยชน์ในการพัฒนาผู้เรียนมิให้เกิดปมด้อยทางภาษาอังกฤษที่ยากต่อการแก้ไข เมื่อผู้เรียนศึกษาต่อในระดับชั้นที่สูงขึ้น

คำถามการวิจัย

1. แบบสอบแบบทดสอบ (Testlet-test) เพื่อวินิจฉัยทางพุทธิปัญญาความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษของผู้เรียนระดับชั้นมัธยมศึกษาปีที่ 3 มีการสร้างและตรวจสอบคุณภาพอย่างไร

2. ความรอบรู้ของนักเรียนในด้านความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ที่ได้รับการวินิจฉัยด้วยแบบสอบแบบทดสอบ (Testlet-test) มีลักษณะอย่างไร และ สารสนเทศจากการวินิจฉัยแบบย้อนกลับ (Feedback) มีลักษณะอย่างไร

วัตถุประสงค์การวิจัย

1. เพื่อสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3

2. เพื่อวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ

ความสำคัญของการวิจัย

1. การวิจัยครั้งนี้ทำให้ได้แบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษที่มีจำนวนข้อคำถามไม่มาก โดยมีลักษณะเป็นกลุ่มข้อคำถามที่ใช้สิ่งเร้าร่วมกันหรือเนื้อหาเดียวกันในการตอบข้อสอบ และแต่ละข้อคำถามวัดได้หลายคุณลักษณะ เมื่อนำผลการวัดมาวิเคราะห์โดยโมเดลวินิจฉัยเชิงพุทธิปัญญาจะทำให้ได้สารสนเทศความสามารถในการอ่านเพื่อความเข้าใจของผู้สอบรายบุคคลและภาพรวม

2. การวิจัยครั้งนี้ทำให้ได้รูปแบบความรอบรู้ของนักเรียนในด้านความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ที่ได้รับการวินิจฉัยด้วยแบบสอบแบบทดสอบ (Testlet-test) และ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับการอ่านเพื่อความเข้าใจของนักเรียนหรือผู้สอบรายบุคคลและภาพรวมทั้งในเชิงปริมาณและเชิงคุณภาพ เกี่ยวกับ

จุดแข็ง จุดอ่อน คุณลักษณะหรือทักษะย่อยที่จำเป็นต้องได้รับการปรับปรุงแก้ไข เพื่อครูผู้สอนหรือผู้เกี่ยวข้องนำไปใช้ออกแบบกิจกรรมการเรียนรู้ที่ตอบสนองทุกรูปแบบความรู้

ขอบเขตของการวิจัย

1. ขอบเขตด้านแหล่งข้อมูล

แหล่งข้อมูลในงานวิจัยประกอบด้วย ผู้เชี่ยวชาญ และกลุ่มตัวอย่างมีรายละเอียด ดังนี้ ผู้เชี่ยวชาญ ทำการตรวจสอบคุณภาพเครื่องมือวิจัยด้านความตรงเชิงเนื้อหาของข้อคำถามกับคุณลักษณะที่มุ่งวัด (Q-matrix) จำนวน 5 ท่าน ประกอบด้วย ผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษและขอบเขตของเนื้อหาภาษาอังกฤษ ผู้ซึ่งมีคุณสมบัติเป็นครูชำนาญการขึ้นไป และมีวุฒิทางการศึกษาปริญญาตรีหรือปริญญาโทในสาขาวิชาเอกภาษาอังกฤษ จำนวน 3 ท่าน ผู้เชี่ยวชาญด้านภาษาอังกฤษหรืออาจารย์มหาวิทยาลัย จำนวน 1 ท่าน และผู้เชี่ยวชาญด้านการวัดผลทางการศึกษาที่เป็นอาจารย์มหาวิทยาลัย จำนวน 1 ท่าน

ประชากร ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนที่กำลังศึกษาอยู่ในระดับชั้นมัธยมศึกษาปีที่ 3 ในปีการศึกษา 2565 จากโรงเรียนมัธยมศึกษาในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเพชรบูรณ์ทั้งหมด 39 แห่ง จำนวน 4,882 คน

กลุ่มตัวอย่าง กลุ่มตัวอย่างในการวิจัยครั้งนี้เป็นนักเรียนที่กำลังศึกษาอยู่ในระดับชั้นมัธยมศึกษาปีที่ 3 ในปีการศึกษา 2565 ของโรงเรียนมัธยมศึกษาในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเพชรบูรณ์เขต 40 ซึ่งแบ่งออกเป็น 3 กลุ่มตามวัตถุประสงค์และขั้นตอนของการวิจัยดังนี้

1. กลุ่มตัวอย่างที่ใช้ในการตรวจสอบคุณภาพของแบบสอบขั้นต้นตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) จำนวน 511 คน

2. กลุ่มตัวอย่างที่ใช้ในการตรวจสอบคุณภาพของแบบสอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ด้วยโมเดลโลจิสติกแบบ 3 พารามิเตอร์ และใช้ตรวจสอบคุณภาพในเชิงการวินิจฉัยของข้อคำถามด้วยโมเดลวินิจฉัยเชิงจำแนก G-DINA ด้านความตรงของ Q-matrix อำนาจจำแนก และความสอดคล้องของโมเดลเชิงสัมพันธ์จำนวน 1,015 คน ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน (Multi-stage sampling)

3. กลุ่มตัวอย่างใช้ทำการวินิจฉัยและจัดอันดับความสามารถในแต่ละคุณลักษณะย่อยของความสามารถการอ่านเพื่อความเข้าใจ จำนวน 396 คน ได้มาโดยการสุ่มแบบหลายขั้นตอน (Multi-stage sampling)

2. ขอบเขตด้านเนื้อหา

ขอบเขตด้านเนื้อหาแบ่งออกเป็น 3 ส่วนตามเครื่องมือวิจัยดังนี้

2.1 ขอบเขตด้านเนื้อหาที่ใช้ในการสร้างแบบสอบ ประกอบด้วย ขอบเขตด้านเนื้อหาอิงเป้าหมายของกรอบอ้างอิงความสามารถทางภาษา CEFR ด้านการอ่านระดับ A2 คือ ผู้เรียนสามารถเข้าใจบทอ่านที่สั้น ๆ ง่าย ๆ เป็นเรื่องใกล้ตัวและใช้ภาษาที่พบบ่อยในชีวิตประจำวันประกอบด้วย

1) การอ่านคำแนะนำ (Reading Instruction) สามารถเข้าใจคำแนะนำที่ใช้ภาษาง่ายๆและขั้นตอนการใช้ เช่น วิธีกดเงินจากตู้ ATM หรือวิธีซื้อเครื่องดื่มจากตู้อัตโนมัติ เครื่องจำหน่ายตัวสารธารณะ ข้อมูลเกี่ยวกับความปลอดภัย และการบอกทิศทาง เช่น Passive Voice ที่ใช้ในโครงสร้างประโยคง่าย ๆ is/are + past participle คำสันธาน (conjunction) เช่น and/ but/ or/ before/ after/ because etc. ตัวเชื่อม (connective words) เช่น First, Second, Third, Fourth, Next, Then, Finally,... etc.

2) การอ่านข้อมูลเบื้องต้น (Reading for Orientation) สามารถค้นหาข้อมูลที่ต้องการและเข้าใจประเด็นหลัก, ใจความสำคัญ (main idea) ของข้อมูลนั้นๆ เช่น ราคา ขนาด ของโฆษณา หน้าเว็บ (Web page) แผ่นพับ หน้าเว็บ (web page) แคตตาล็อก ตารางเวลา เช่น หัวข้อเรื่อง ใจความสำคัญ รายละเอียดสนับสนุน คำถามเกี่ยวกับใจความสำคัญของเรื่อง เช่น ใคร ทำอะไร ที่ไหน เมื่อไร อย่างไร ทำไม ใช่หรือไม่ - Yes/No Question - Wh-Question etc.

3) การอ่านข้อมูลและข้อโต้แย้ง (Reading for Information & Argument) สามารถเข้าใจประเด็นหลักในบทความสั้น ๆ จากหนังสือพิมพ์ นิตยสาร เกี่ยวกับบุคคลที่มีชื่อเสียงในสังคม หรือเหตุการณ์ปัจจุบัน

4) การอ่านวรรณกรรม (Read Literature) สามารถเข้าใจประเด็นสำคัญ (Main idea) ของงานเขียน เช่น เรื่องสั้น (Story) บทกลอน (Poem) เกี่ยวกับชีวิตประจำวันซึ่งใช้ภาษาง่ายๆ และมีสื่อหรือรูปภาพประกอบ

5) การอ่านเอกสารโต้ตอบหรือจดหมาย (Reading Correspondence) สามารถเข้าใจข้อความง่าย ๆ ในจดหมายอิเล็กทรอนิกส์ (E-mail) การสนทนาทางเว็บหรือแอปพลิเคชัน หรือข้อความจากเพื่อนหรือเพื่อนร่วมงานเกี่ยวกับเหตุการณ์ต่างๆในชีวิตประจำวัน เช่น การนัดหมายเวลาไปรับประทานอาหารกลางวัน /อาหารเย็น We shall meet at six p.m. at BN café. /When shall we meet? เป็นต้น

2.2 ขอบเขตด้านเนื้อหาของสิ่งที่มีุงวินิจฉัย ได้แก่ ความสามารถในการอ่านเพื่อความเข้าใจ ประกอบด้วยคุณลักษณะย่อย (Attributes) 6 คุณลักษณะ ดังนี้ มีความรู้คำศัพท์ (Vocabulary Knowledge) มีความสามารถในการระบุข้อมูลที่ต้องการ (Identifying Explicit Information) มีความรู้ทางไวยากรณ์ (Syntactic Knowledge) มีความสามารถในการระบุใจความสำคัญของเรื่อง (Generalization Main Idea) มีความสามารถในการอนุมาน (Getting Inference) และมีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน (Synthesizing Information from Text)

3. ขอบเขตด้านตัวแปร

ขอบเขตด้านตัวแปรแยกตามวัตถุประสงค์การวิจัย มีรายละเอียดดังนี้

3.1 การสร้างและตรวจสอบคุณภาพของแบบสอบถามแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญา

1) ตัวแปรที่แบบสอบถามวินิจฉัย คือ ความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษ ในงานวิจัยนี้มีกรอบองค์ประกอบหรือคุณลักษณะย่อย (Attributes) ของความสามารถในการอ่านเพื่อความเข้าใจ ซึ่งผู้วิจัยสังเคราะห์จากเอกสารที่เกี่ยวข้องกับทักษะหรือคุณลักษณะความสามารถในการอ่านเพื่อความเข้าใจซึ่งนักวิชาการได้ศึกษาและเสนอไว้เป็น 6 คุณลักษณะ ดังนี้

๔ มีความรู้คำศัพท์ (Vocabulary Knowledge) พิจารณาจากความเข้าใจเกี่ยวกับความหมายคำศัพท์ และกลุ่มคำโดยอาศัยพื้นฐานความรู้เดิม หรือเข้าใจความหมายของคำศัพท์ที่ยากหรือไม่คุ้นเคยโดยอาศัยการคาดเดาความหมายคำศัพท์จากบริบทหรือคำบ่งชี้ที่พบในข้อความ (Context Clue) ประกอบด้วยการนิยามความหมาย การกล่าวซ้ำ การเปรียบเทียบความเหมือนหรือความต่าง และการยกตัวอย่าง

๕ มีความสามารถในการระบุข้อมูลที่ต้องการ (Identifying Explicit Information) พิจารณาได้จากทักษะในการค้นหารายละเอียดหรือข้อมูลเฉพาะที่ต้องการซึ่งปรากฏชัดเจนในบทอ่าน หลังจากที่ได้อ่านแบบผ่าน (Scan) และอ่านแบบพินิจพิเคราะห์ (Skim) ทำให้ทราบว่าข้อมูลอยู่ที่ใดของเรื่อง เช่น จำนวน ราคา วันเวลา สถานที่ ซึ่งข้อมูลดังกล่าวจะปรากฏอย่างชัดเจนผู้อ่านสามารถเข้าใจได้โดยมีสติความ

๖ มีความรู้ทางไวยากรณ์ (Syntactic Knowledge) พิจารณาได้จากความเข้าใจเกี่ยวกับวากยสัมพันธ์ซึ่งหมายถึงการศึกษาความสัมพันธ์ของคำในประโยคหรือแบบแผนองค์ประกอบของประโยคในภาษาอังกฤษ และความสัมพันธ์ระหว่างประโยคและสามารถวางตำแหน่งของคำหรือกลุ่มคำประเภทต่าง ๆ ในประโยคได้อย่างถูกต้องตามโครงสร้างทางไวยากรณ์

๗ มีความสามารถในการระบุใจความสำคัญของเรื่อง (Generalization Main Idea) พิจารณาได้จากการทำความเข้าใจเพื่อค้นหาหัวเรื่อง (Topic) นำมารวมกับข้อความที่ถูกควบคุมให้กล่าวถึงสิ่งใดสิ่งหนึ่ง (Controlling Idea) และลงข้อสรุปเป็นใจความสำคัญ หรือแนวคิดหลักที่อยู่ในบทอ่านซึ่งเป็นส่วนที่ครอบคลุมย่อหน้านั้น ๆ หรือเนื้อหาทั้งหมด

๘ มีความสามารถในการอนุมาน (Getting Inference) พิจารณาได้จากความสามารถในการตีความจากเรื่องที่อ่าน โดยใช้ความรู้และประสบการณ์ตรงของผู้อ่าน เพื่อทำความเข้าใจข้อมูลที่ไม่ได้ถูกระบุไว้อย่างชัดเจนซึ่งเป็นจุดมุ่งหมาย ทศนคติของผู้เขียน หรือข้อโต้แย้งโดยอาศัยเบาะแสร่องรอยในเนื้อความซึ่งเป็นคำหรือประโยคชี้แนะ หรือทำการตรวจสอบข้อเท็จจริง

ทั้งหมดของเรื่องที่อ่าน เพื่อพิจารณาหรือตัดสินว่าสิ่งที่อ่านทั้งหมดนั้นหมายความว่าถึงสิ่งใดและสามารถเชื่อมโยงไปถึงข้อมูลภายนอกซึ่งไม่ได้กล่าวไว้ในบทความ

๘ มีความสามารถในการสังเคราะห์ข้อมูลจากบทความ (Synthesizing Information from Text) พิจารณาได้จากความสามารถในการแยกประโยคแนวคิดหลัก หรือใจความสำคัญออกจากประโยครองหรือข้อความซึ่งเป็นรายละเอียดเสริมจากบทความ นำประเด็นหลักที่รวบรวมได้ มาเชื่อมโยงเชื่อมต่อกัน โดยพิจารณาความเกี่ยวข้อง ความเป็นเหตุเป็นผล ทำการจัดลำดับข้อความ และทำการลงข้อสรุปหรือเรียบเรียงประโยคขึ้นใหม่โดยคงความหมายเดิม

2) ตัวแปรคุณภาพของแบบสอบ ประกอบด้วย

2.1 ความตรงเชิงเนื้อหา (Content validity) ค่าดัชนีความสอดคล้อง (IOC) ระหว่างข้อสอบ (ฉบับร่าง) กับคุณลักษณะการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในแต่ละด้าน

2.2 ความตรงของ Q-matrix ในเชิงประจักษ์เป็นการตรวจความถูกต้องและเหมาะสมของ Q-matrix ซึ่งเป็นเมทริกซ์แสดงความสัมพันธ์ระหว่างลำดับที่ของข้อคำถามกับคุณลักษณะซึ่งเป็นความสามารถที่ใช้ในการหาคำตอบของข้อคำถามนั้นให้สอดคล้องกับความเป็นจริงมากที่สุด

2.3 ค่าพารามิเตอร์ความยาก (b -parameter) ค่าพารามิเตอร์อำนาจจำแนก (a -parameter) พารามิเตอร์โอกาสการเดา (c -parameter) ที่ได้จากการคำนวณตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ด้วยโมเดลตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (IRT 3PL)

2.4 ดัชนีความสอดคล้องของโมเดลเชิงสัมพัทธ์ ได้แก่ ค่า $-2LL$ (-2Loglikelihood) ค่า AIC (Akaike's information criterion) ค่า BIC (Bayesian information criterion) ค่า RMSEA (Root Mean Square Error of Approximation)

2.5 ดัชนีค่าความถูกต้อง (Test Level Accuracy)

2.6 ค่าระดับคุณลักษณะ (Attribute Level Accuracy)

2.7 Latent class analysis (LCA) ได้แก่ ค่าพารามิเตอร์การเดา (Guessing Parameter) ค่าพารามิเตอร์ความสะเพร่า (Slipping Parameter) ค่าดัชนีอำนาจจำแนก (Generalization Discrimination Index: GDI) ค่าความน่าจะเป็นของการรอบรู้แต่ละคุณลักษณะ (Skill Probabilities)

3.2 การวินิจฉัยความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษ

- ตัวแปรที่ศึกษา

1) รูปแบบความรู้ของนักเรียน

2) Profile สารสนเทศจากการวินิจฉัยย้อนกลับรายบุคคล

นียมศัพท์เฉพาะ

1) การประเมินวินิจฉัยทางพุทธิปัญญา หมายถึง กระบวนการที่เป็นระบบซึ่งใช้สำหรับการค้นพบสารสนเทศซึ่งเป็นทักษะความสามารถทางปัญญาเฉพาะด้าน เพื่อค้นหาว่าผู้เรียนได้เรียนรู้อะไรมาบ้างเกี่ยวกับสิ่งที่จะเรียน สิ่งที่ยู่มาก่อนนี้ถูกต้องหรือไม่ จึงเป็นการใช้ในลักษณะประเมินก่อนเรียน เพื่อตรวจสอบความก้าวหน้า นอกจากนี้ยังใช้เพื่อให้เหตุผล ควบคุม และทำการ ตัดสินใจอย่างเหมาะสมเกี่ยวกับการกระทำที่เป็นผลอันเกิดมาจากลักษณะทางจิต หรือหาสาเหตุของปัญหาหรืออุปสรรคต่อการเรียนรู้ของผู้เรียนเป็นรายบุคคล

2) โมเดลวินิจฉัยเชิงพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจ หมายถึง โครงสร้างคุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ ประกอบด้วย 6 คุณลักษณะย่อย คือ มีความรู้คำศัพท์ (Vocabulary Knowledge) มีความสามารถในการระบุข้อมูลที่ต้องการ (Identifying Explicit Information) มีความรู้ทางไวยากรณ์ (Syntactic Knowledge) มีความสามารถในการระบุใจความสำคัญของเรื่อง (Generalization Main Idea) มีความสามารถในการอนุมาน (Getting Inference) และมีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน (Synthesizing Information from Text)

3) กรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรป (CEFR) ด้านการอ่านในระดับ A2 คือ เป้าหมายความสามารถด้านการอ่านภาษาอังกฤษซึ่งเป็นมาตรฐานการประเมินที่สหภาพยุโรปกำหนดขึ้นเพื่อใช้เป็นแนวทางในการจัดการเรียนการสอน โดยเป็นความสามารถในการทำความเข้าใจเนื้อหาในเอกสาร ตำราเรียน บทความง่าย ๆ เว็บไซต์ ฯลฯ เข้าใจประเด็นสำคัญหรือใจความสำคัญ สามารถค้นหาข้อโต้แย้งง่าย ๆ ที่ปรากฏในเรื่องที่อ่าน สามารถเข้าใจเรื่องด้วยการคาดเดาจากรูปภาพ ประกอบได้ และสามารถรับรู้ความหมายของคำจากพื้นความรู้เดิม หรือการเดาความหมายจากบริบท

4) แบบสอบสำหรับใช้ในการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ หมายถึง แบบสอบแบบทดสอบ (Testlet-test) ที่ใช้ในการวินิจฉัยความสามารถของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ด้านการอ่านเพื่อความเข้าใจภาษาอังกฤษ ประกอบด้วยข้อสอบแบบทดสอบเป็นกลุ่มหรือชุดของข้อสอบหลายข้อที่มีการใช้สถานการณ์ หรือข่าวเหตุการณ์ในปัจจุบันเป็นสิ่งเร้าร่วมกัน ในลักษณะของบทอ่าน (Passage) โดยใช้ขอบเขตเนื้อหาอิงเป้าหมายกรอบอ้างอิงความสามารถทางภาษา CEFR ระดับ A2 มีการตรวจให้คะแนนแบบสองค่า (Dichotomous) เป็นข้อสอบแบบปรนัย 4 ตัวเลือก ซึ่งผู้วิจัยสร้างขึ้นตามขั้นตอนการสร้างแบบสอบอิงกรอบการออกแบบการประเมินเพื่อวินิจฉัยทางพุทธิปัญญาของเอ็มเบรทสัน (CDS framework's Embretson, 1984,1998, p.383-384)

5) คุณภาพของแบบสอบวินิจฉัยทางพุทธิปัญญา หมายถึง คุณภาพของแบบสอบที่ใช้ในการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ โดยพิจารณาจาก

5.1 ความตรงเชิงเนื้อหา (Content Validity) โดยพิจารณาจากดัชนีความสอดคล้องระหว่างข้อสอบ ในแบบสอบ (ฉบับร่าง) กับคุณลักษณะการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในแต่ละด้าน พิจารณาโดยผู้เชี่ยวชาญ (IOC) และความตรงของเมตริกซ์คิว (Q-matrix validation) พิจารณาโดยการทดสอบวาร์ด วิธี Stepwise Wald Test ด้วยโมเดลวินิจฉัยเชิงพุทธิปัญญา (CDM) ใช้ G-DINA Model

5.2 ค่าพารามิเตอร์ความยาก (b -parameter) ค่าพารามิเตอร์อำนาจจำแนก (a -parameter) พารามิเตอร์โอกาสการเดา (c -parameter) ที่ได้จากการคำนวณตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ด้วยโมเดลตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (IRT 3PL)

5.3 ดัชนีค่าความถูกต้อง (Test Level Accuracy)

5.4 ค่าระดับคุณลักษณะ (Attribute Level Accuracy)

5.5 Latent class analysis (LCA) ได้แก่ ค่าพารามิเตอร์การเดา ค่าพารามิเตอร์ความสะเพร่า ค่าดัชนีอำนาจจำแนก (IDI)

5.6 ค่าความน่าจะเป็นของการรอบรู้แต่ละคุณลักษณะ (Attribute Probabilities)

6) ประสิทธิภาพของโมเดลที่ใช้วินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ หมายถึง ความสามารถในการวิเคราะห์ข้อมูลเชิงการวินิจฉัยได้อย่างถูกต้อง และสอดคล้องกับข้อมูลเชิงประจักษ์โดยพิจารณาจากดัชนีความสอดคล้องของโมเดล ต่อไปนี้

6.1 ดัชนีความสอดคล้องของโมเดลเชิงสัมพัทธ์ พิจารณาจากค่าดัชนีความสอดคล้อง -2 Log Likelihood (-2LL), Akaike's Information Criterion (AIC) และ Bayesian Information Criterion (BIC) โดยค่าที่ต่ำกว่าแสดงให้เห็นถึงประสิทธิภาพในการประมาณค่าของโมเดลที่ดีกว่า (ในกรณีที่ทำกรทดสอบเปรียบเทียบ 2 โมเดล ขึ้นไป)

6.2 ค่าดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ $RMSEA_2$ ตรวจสอบความเหมาะสมของข้อสอบเป็นรายข้อด้วยดัชนีความคลาดเคลื่อนในการประมาณ ค่าพารามิเตอร์ ($RMSEA_2$) ตรวจสอบโดยใช้เกณฑ์ของ Maydeu – Olivares, & Joe (2014) ค่า $RMSEA_2 < 0.089$ หมายถึง ข้อสอบข้อนั้นมีความเหมาะสมเพียงพอที่จะสามารถนำไปใช้ได้ ค่า $RMSEA_2 < 0.05$ หมายถึง ข้อสอบข้อนั้น มีความคลาดเคลื่อนต่ำมากและมีความเหมาะสมเป็นอย่างมากที่จะนำไปใช้วินิจฉัยผู้สอบ

7) แบบแผนการคิดที่ใช้ในการตอบข้อสอบ หมายถึง คุณลักษณะเฉพาะที่เป็นการนำทักษะย่อยของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษมาใช้ในการตอบข้อสอบ ซึ่งถูกกำหนด

ด้วยเมทริกซ์ขนาด $1 \times n$ เมื่อ n แทนจำนวนความสามารถทั้งหมดที่ใช้ในการตอบข้อสอบ โดยสมาชิกในเมทริกซ์มีค่าเท่ากับ 1 เมื่อนักเรียนใช้คุณลักษณะ j ในการหาคำตอบของคำถามที่กำหนดและสมาชิกในเมทริกซ์มีค่าเท่ากับ 0 เมื่อนักเรียนไม่ใช้คุณลักษณะ j ในการหาคำตอบ เมื่อ $j \in \{1, 2, \dots, 3\}$

8) รูปแบบความรู้ของนักเรียน หมายถึง ผลการวินิจฉัยความรอบรู้เรื่องการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียน ตัวอย่างเช่น นักเรียนคนหนึ่ง มีรูปแบบการรอบรู้เป็น 110101 หมายความว่า นักเรียนผู้นี้ มีความรอบรู้ในสี่คุณลักษณะของการอ่านเพื่อความเข้าใจ คือ คุณลักษณะที่ 1 2 4 และ 6 แต่ไม่รอบรู้ในคุณลักษณะที่ 3 และ 5 ซึ่งในงานวิจัยนี้ การอ่านเพื่อความเข้าใจประกอบด้วย 6 คุณลักษณะจำนวนรูปแบบความรู้ที่คาดหวังได้จาก 2^n (n =จำนวนคุณลักษณะ) ดังนั้น 2^6 จึงประกอบด้วย 64 รูปแบบความรู้

9) สารสนเทศจากการวินิจฉัยย้อนกลับรายบุคคล (Profile) คือ ข้อมูลเกี่ยวกับความรอบรู้และความไม่รอบรู้ของนักเรียนเกี่ยวกับคุณลักษณะที่เป็นทักษะความสามารถที่นักเรียนนำมาใช้ในการตอบข้อคำถามของแบบสอบแบบทดสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อสร้างและตรวจสอบคุณภาพแบบสอบถามแบบทดสอบเพื่อวินิจฉัยเชิงพุทธิปัญญาความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษและวินิจฉัยความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ผู้วิจัยได้ทำการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องโดยมีรายละเอียดแบ่งออกเป็น 5 ตอน ดังต่อไปนี้

ตอนที่ 1 มโนทัศน์เกี่ยวกับความสามารถการอ่านเพื่อความเข้าใจ

- 1.1 ความหมายของความสามารถการอ่านเพื่อความเข้าใจ
- 1.2 รูปแบบการประเมินความเข้าใจในการอ่าน
- 1.3 แนวคิดทฤษฎีที่เกี่ยวข้องกับความสามารถการอ่านเพื่อความเข้าใจ
- 1.4 คุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจ
- 1.5 กรอบอ้างอิงความสามารถทางภาษา CEFR

ตอนที่ 2 มโนทัศน์เกี่ยวกับการประเมินเชิงวินิจฉัยทางพุทธิปัญญา

- 2.1 ความหมายของการประเมินเชิงวินิจฉัยทางพุทธิปัญญา
- 2.2 โมเดลวินิจฉัยทางพุทธิปัญญา
- 2.3 แนวคิดสำหรับการออกแบบการประเมินเพื่อวินิจฉัย
- 2.4 การกำหนดคุณลักษณะของโมเดลการวินิจฉัยเชิงพุทธิปัญญา
- 2.5 แบบสอบวินิจฉัยทางพุทธิปัญญา

ตอนที่ 3 มโนทัศน์เกี่ยวกับข้อสอบแบบทดสอบ

- 3.1 ลักษณะสำคัญของข้อสอบแบบทดสอบ
- 3.2 แนวคิดทฤษฎีที่เกี่ยวข้องกับข้อสอบแบบทดสอบ
- 3.3 ตัวอย่างข้อสอบแบบทดสอบและงานวิจัยที่เกี่ยวข้อง

ตอนที่ 4 มโนทัศน์เกี่ยวกับการให้ข้อมูลย้อนกลับ (Feedback)

- 4.1 ความหมายของข้อมูลย้อนกลับ
- 4.2 ประเภทของการให้ข้อมูลย้อนกลับ
- 4.3 แนวทางการให้ข้อมูลย้อนกลับ

ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง

- 5.1 งานวิจัยในประเทศ
- 5.2 งานวิจัยในต่างประเทศ

ตอนที่ 6 กรอบแนวคิดการวิจัย

ตอนที่ 1 มโนทัศน์เกี่ยวกับความสามารถการอ่านเพื่อความเข้าใจ (Reading Comprehension)

การอ่านถือเป็นทักษะที่สำคัญสำหรับนักเรียนที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศหรือภาษาที่สอง มากกว่าทักษะอื่น การอ่านเป็นบ่อเกิดแห่งองค์ความรู้ต่างๆ บุคคลที่สามารถอ่านภาษาอังกฤษได้ดีย่อมสามารถพัฒนาตนเองและสืบค้นความรู้ได้อย่างกว้างขวางและหลากหลายด้าน โดยเฉพาะอย่างยิ่งความสามารถในการอ่านภาษาอังกฤษเพื่อความเข้าใจ เนื่องจากบุคคลที่สามารถเข้าใจเรื่องที่อ่านจะสามารถประมวลผลข้อมูลจากการอ่านในลักษณะอื่นๆ ในขั้นสูงขึ้นไปได้อีกด้วย เช่น การอ่านอย่างมีวิจารณญาณ เป็นต้น ดังนั้นการอ่านจึงมิใช่เพียงเป็นการรับสารเท่านั้น หากแต่เป็นกระบวนการทางความคิดที่ซับซ้อน ในการศึกษาครั้งนี้นำเสนอความหมายของการอ่านเพื่อความเข้าใจ องค์ประกอบของความสามารถการอ่านเพื่อความเข้าใจ แนวคิดทฤษฎีที่เกี่ยวข้อง และคุณลักษณะการอ่านเพื่อความเข้าใจจากการสังเคราะห์คุณลักษณะและทักษะย่อยของการอ่านเพื่อความเข้าใจของนักการศึกษาด้านภาษาศาสตร์ โดยมีรายละเอียดตามหัวข้อต่าง ๆ ดังต่อไปนี้

1. ความหมายและองค์ประกอบของความสามารถการอ่านเพื่อความเข้าใจ

ความหมายของความสามารถการอ่านเพื่อความเข้าใจที่ได้จากข้อสรุปของนักวิชาการทางภาษาศาสตร์ทั้งไทยและต่างประเทศ มีรายละเอียดดังนี้

Carrell (1983, p. 569) ได้กล่าวไว้ว่า การอ่านความเข้าใจเป็นกระบวนการโต้ตอบระหว่างข้อความและความรู้เบื้องต้นของผู้อ่าน ความเข้าใจในการอ่านเกี่ยวข้องกับความรู้ของโลกซึ่งอาจจะขึ้นอยู่กับวัฒนธรรมและอคติทางวัฒนธรรม

Goodman (1988) กล่าวไว้ว่า การอ่านเพื่อความเข้าใจ หมายถึง กระบวนการรับรู้ทางภาษาซึ่งเป็นปฏิสัมพันธ์ระหว่างกระบวนการคิดของผู้อ่านกับข้อความในบทอ่านโดยการสร้างความหมายในสิ่งที่อ่าน ผู้อ่านที่มีความสามารถในการตีความจะเข้าใจจุดมุ่งหมายหรือสิ่งที่ผู้เขียนต้องการสื่อสาร

Anderson (1985, p. 372, อ้างถึงใน อรอนงค์ อุ้นแก้ว, 2562, หน้า 10-11) กล่าวว่า การอ่านเป็นกระบวนการค้นหาความหมายในหลายระดับตั้งแต่ระดับตัวอักษร โครงสร้างภาษา ไปจนถึงระดับความหมายการอ่านเพื่อจับใจความเกิดจากปฏิริยาสัมพันธ์ระหว่างเนื้อหาและโครงสร้างของเรื่องกับความรู้เดิมของผู้อ่านซึ่งจะช่วยเสริมสร้างสมมุติฐานเกี่ยวกับเรื่องที่อ่าน ผู้อ่านที่อ่านได้คล่องแคล่วจึงสามารถทำนายและอ้างอิงข้อมูลจากการอ่านได้

ฮีตัน (Heaton, 1991) อธิบายว่าการอ่านเพื่อความเข้าใจถูกแบ่งออกเป็นทักษะย่อยๆ ได้หลายทักษะ จากผลการศึกษาเชิงคุณภาพเกี่ยวกับอนุกรมวิธานทักษะการอ่านจับใจความได้ดำเนินการโดยผู้เชี่ยวชาญด้านการอ่านเป็นหลัก ผู้เชี่ยวชาญด้านการอ่านมักพยายามระบุทักษะหรือ

ความสามารถในการเข้าใจในการอ่านโดยการให้ชุดบทความ (Passage) และถามคำถามที่มีจุดประสงค์เพื่อทดสอบความเข้าใจในระดับต่างๆของข้อความในลักษณะนี้

ดัฟฟี และโรเลอร์ (Duffy & Roehler, 1993, p. 29, อ้างถึงใน วราภรณ์ พรหมอินทร์, 2561, หน้า 27) กล่าวว่า การอ่านเพื่อความเข้าใจหมายถึง ความสามารถในการตีความหมายจากบทอ่านได้อย่างถูกต้อง รวมทั้งสามารถทำความเข้าใจกับความหมายนั้น การตีความจะอาศัยพื้นฐานความรู้เดิมของผู้อ่านเป็นสิ่งสำคัญ หรือความเข้าใจในการอ่านจะเกิดขึ้นได้เมื่อผู้อ่านนำประสบการณ์เดิมที่มีอยู่แล้วมาสัมพันธ์กับข้อเขียนที่อ่าน

กรอบอ้างอิงความสามารถทางภาษา CEFR (2544) อธิบายว่า การอ่านคือการสื่อความรู้ ความรู้สึกนึกคิดระหว่างผู้เขียนและผู้อ่าน โดยผู้อ่านเข้าใจในสัญลักษณ์ เครื่องหมายรูปภาพ ประโยค ข้อความ ตัวอักษร คำ และข้อความที่พิมพ์หรือเขียนขึ้นมาด้วยการสังเกตและพิจารณาซึ่งมีความหมายตรงกับผู้สื่อสารเขียน โดยเป็นกระบวนการปฏิสัมพันธ์และการตีความระหว่างผู้เขียนและผู้อ่านซึ่งจะเข้าใจมากหรือน้อยแค่ไหนขึ้นอยู่กับความรู้เดิมและการทำนาย ความรู้เกี่ยวกับเรื่อง ความหมายของศัพท์ โครงสร้างทางไวยากรณ์ ตลอดจนความเข้าใจในเรื่องความสอดคล้องต่อเนื่องของประโยคแต่ละประโยคและโครงสร้างของข้อเขียนด้วย

ถิรวัดน์ ต้นทนิส (2550, หน้า 121) การอ่านเพื่อความเข้าใจ เป็นการสร้างความหมายจากภาษาเขียนโดยมีการ เชื่อมโยงเนื้อเรื่องที่อ่าน กับพื้นความรู้เดิมของผู้อ่าน ซึ่งมีลักษณะเป็นกระบวนการเชิง ปฏิสัมพันธ์ระหว่างพื้นความรู้เดิมกับความรู้จากภาษาเขียน ความหมายที่ผู้อ่านสร้างขึ้นอาจคล้ายคลึงกับความหมายในความคิดของผู้เขียน ผู้อ่านแต่ละคนย่อมสร้างความหมายแตกต่างกัน และในการอ่านนอกจากจะต้องมีปฏิสัมพันธ์กับบทอ่านแล้วยังต้องมี ปฏิสัมพันธ์กับองค์ประกอบอื่นนอกเหนือบทอ่าน เพื่อให้ผู้อ่านเกิดความเข้าใจในบท อ่านนั้นๆ อีกด้วย

สถาบันภาษาอังกฤษ สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (2560, หน้า 97) ได้อธิบายว่า การอ่านเพื่อความเข้าใจคือการสื่อความรู้ความรู้สึกนึกคิดระหว่างผู้เขียนและผู้อ่านโดยผู้อ่านเข้าใจในสัญลักษณ์ เครื่องหมาย รูปภาพ ประโยค ข้อความ ตัวอักษร คำ และข้อความที่พิมพ์หรือเขียนขึ้นมาด้วยการสังเกตและพิจารณาซึ่งมีความหมายตรงกับผู้สื่อสารเขียนโดยเป็นกระบวนการปฏิสัมพันธ์และการตีความระหว่างผู้เขียนและผู้อ่านซึ่งจะเข้าใจมากหรือน้อยขึ้นอยู่กับความรู้เดิมและการทำนายความรู้เกี่ยวกับเรื่องความหมายของคำศัพท์ โครงสร้างไวยากรณ์ ตลอดจนความเข้าใจในเรื่องความสอดคล้องต่อเนื่องของประโยคแต่ละประโยค

ณัฐมน วงศ์ทาทอง (2560, หน้า 53) การอ่านเพื่อความเข้าใจ หมายถึง กระบวนการทางความคิดที่ใช้ความรู้ ประสบการณ์เดิมของผู้อ่านมาสร้างความหมายให้เข้าใจเรื่องที่อ่าน ให้ตรงตามวัตถุประสงค์ของผู้เขียน โดยผู้อ่านจะต้องสามารถตีความ และสรุปความในสิ่งที่ผู้เขียนต้องการสื่อให้ผู้อ่านรับรู้

วรารณณ์ พรหมอินทร์ (2561, หน้า 27) กล่าวว่า ความเข้าใจในการอ่านเป็นความสามารถในการทำความเข้าใจทั้งในระดับตัวอักษร และระดับความหมายโดยการตีความจะอาศัยปฏิสัมพันธ์ระหว่างเนื้อเรื่องที่อ่านกับความรู้หรือประสบการณ์เดิมของผู้อ่านเป็นหลักจึงทำให้เกิดความรู้ใหม่ ความเข้าใจเป็นพื้นฐานที่จะนำไปสู่ความสามารถทางการอ่าน ทำให้การอ่านประสบความสำเร็จ อันเป็นจุดมุ่งหมายสูงสุดของการอ่าน และทำให้ผู้อ่านสามารถนำความรู้อันเกิดจากความเข้าใจในการอ่านนั้นไปพัฒนาคุณภาพชีวิตของตนเองและสังคมได้

อรอนงค์ อุ่นแก้ว (2562, หน้า 16) การอ่านเพื่อความเข้าใจ คือ กระบวนการเรียนรู้และพฤติกรรมที่มีความเชื่อมโยงต่อกระบวนการเรียนรู้เพื่อให้เกิดความเข้าใจและเป็นการกระตุ้นให้แสดงออกเพื่อสะท้อนความคิดที่ได้จากเรื่องที่อ่าน เช่น ความเข้าใจในเนื้อหาต้นคำ ประโยค วลี การจับประเด็นสำคัญจากข้อความที่ระบุชัดเจน การแปลความจากสัญลักษณ์ หรือเนื้อหาที่มีความซับซ้อน รวมทั้งการสะท้อนความคิดจากเรื่องที่อ่านแสดงความคิดเห็นจากเรื่องที่อ่านได้และจะต้องอาศัยความรู้เดิม รวมทั้งทักษะต่างๆในการอ่านเพื่อให้เกิดความเข้าใจในเรื่องที่อ่านได้

Primary National Strategy (2005, p. 2) การอ่านเพื่อความเข้าใจ เป็นการที่ผู้อ่านอ่านแล้วสามารถแปลความจากสิ่งที่อ่านได้ตามความเข้าใจของตนเองการอ่านที่ดีต้องเข้าใจในสิ่งที่อ่านได้อย่างครอบคลุม เพื่อช่วยให้บรรลุเป้าหมายในการอ่านและเกิดความเข้าใจในการอ่านในมุมกว้างและเป็นการกระตุ้นให้สะท้อนกลับให้เห็นในสิ่งที่เข้าใจจากเรื่องที่อ่าน

เกา (Gao, 2006, p. 13) ได้กล่าวไว้ว่า ความสามารถในการอ่านเป็นคุณลักษณะหรือทักษะย่อย ๆ ที่หลากหลายครอบคลุมถึงการจดจำแนวคิดหลัก การทำความเข้าใจความสัมพันธ์ระหว่างประโยคและส่วนต่างๆ ของข้อความ การตีความตามข้อความการสังเคราะห์ ทำความเข้าใจจุดประสงค์หรือทัศนคติของผู้เขียน และจดจำคำศัพท์ในบริบท.

จาง (Jang, 2009, p. 3) ชี้แจงว่าการอ่านเป็นความสามารถที่มีความซับซ้อนไม่สามารถมองเห็นกระบวนการทางสมองที่เกิดขึ้นขณะอ่านเป็นเรื่องเกี่ยวกับจิต และกิจกรรมทางสังคมที่มีปฏิสัมพันธ์กับมนุษย์และสิ่งแวดล้อม

เดลต้า (DELTA, 2014, p.64) ความสามารถของนักเรียนในการอ่านและทำความเข้าใจชนิดหรือประเภทของบทอ่านที่ถูกเขียนภาษาอังกฤษที่พวกเขาจะอ่านเพื่อการเรียนรู้ภาษาอังกฤษโดยทั่วไปมากขึ้น ข้อความที่ผู้อ่านทำการอ่านเป็นข้อความต่างๆประเภท (เช่น บทความวิชาการ บทความข่าว และบทความเด่น) เหล่านี้คือประเภทของข้อความที่คาดว่าจะระบุระดับความสามารถในการอ่านของผู้อ่านแต่ละข้อความเขียนเป็นภาษาอังกฤษมาตรฐาน

จาวิดเนมาห์และซาหรับ (Javidanmehr & Sarab, 2019, p.6-8) กล่าวว่าไว้ว่า การอ่านเพื่อความเข้าใจเป็นความสามารถทางภาษาที่ผู้อ่านใช้ในการทำความเข้าใจประกอบด้วยความรู้

ด้านคำศัพท์ ความรู้ด้านไวยากรณ์ ข้อมูลที่ระบุอย่างชัดเจนในเรื่องที่อ่าน การอนุมานการเชื่อมโยง และสังเคราะห์ข้อมูลจากเรื่องที่อ่าน

โตนีกาโบนี (Tonekaboni F. R., 2021, p.60) การอ่านเป็นกระบวนการทางปัญญา แบบไดนามิกที่เกี่ยวข้องกับความเข้าใจข้อความที่ถูกเขียนขึ้น การพัฒนา การอนุมานและการใช้ ความหมายที่เหมาะสมกับประเภทของข้อความ วัตถุประสงค์ และสถานการณ์

จากทัศนะของนักการศึกษาทั้งไทยและต่างประเทศ สรุปได้ว่าการอ่านเพื่อความเข้าใจ เป็นความสามารถในการสื่อสารความคิดหรือความรู้ระหว่างผู้เขียนและผู้อ่านโดยเป็นกระบวนการ ปฏิสัมพันธ์ระหว่างผู้อ่านกับข้อความโดยอาศัยพื้นความรู้เดิม กระบวนการคิด และกลวิธีในการ ตีความหรือสร้างความหมายจากตัวอักษร สัญลักษณ์ คำ ประโยค บทอ่าน

องค์ประกอบของความสามารถการอ่านเพื่อความเข้าใจ การทำความเข้าใจในการอ่านนั้นต้องอาศัย องค์ประกอบหลายอย่างเพื่อช่วยให้ ผู้อ่านเกิดความเข้าใจในการอ่าน ซึ่งองค์ประกอบของการอ่าน เพื่อความเข้าใจจากทัศนะของนักศึกษามีรายละเอียดดังต่อไปนี้

Pearson (1978: 8 – 10, อ้างถึงใน ธรรมนูญ วงษ์ทาทอง, 2560, หน้า 58) ได้กล่าวว่า การอ่านเพื่อความเข้าใจ เกี่ยวข้องกับภาษา แรงจูงใจ การเรียนรู้ การพัฒนาความคิดรวบยอด และ ประสบการณ์องค์ประกอบที่มีอิทธิพลต่อการอ่านเพื่อความเข้าใจ 2 ประเภท คือ

1. องค์ประกอบภายในสมอง ได้แก่ ความสามารถทางภาษาศาสตร์ รวมทั้ง องค์ประกอบด้านอารมณ์ ความสนใจ แรงจูงใจ และความสามารถในการอ่าน

2. องค์ประกอบภายนอกสมอง ซึ่งแบ่งเป็น

2.1 องค์ประกอบในหน้ากระดาษ นั่นคือ คุณลักษณะของเนื้อความรวมทั้งความยาก ง่ายของเนื้อหา และการจัดระบบในเนื้อหา

2.2 คุณภาพของสิ่งแวดล้อมของผู้อ่าน รวมทั้งองค์ประกอบ เกี่ยวกับกิจกรรมที่ครู กระทำก่อนการอ่าน ระหว่างการอ่าน หรือ หลังการอ่าน เพื่อช่วยให้ผู้เรียนเกิดความเข้าใจเนื้อเรื่อง

Bond and Tinker (1979: 325 – 350, อ้างถึงใน ธรรมนูญ วงษ์ทาทอง, 2560, หน้า 59) กล่าวว่า การ ที่ผู้อ่านจะเกิดความเข้าใจได้นั้นต้องอาศัยองค์ประกอบดังนี้

1. เข้าใจความหมายของคำ (Word meaning) การเข้าใจความหมายคำศัพท์ ช่วยให้เข้าใจประโยคได้ดี

2. เข้าใจหน่วยความคิด (Thought unit) การอ่านทีละคำ จะเป็นอุปสรรคต่อ ความเข้าใจ เพราะขาดความต่อเนื่อง ดังนั้นผู้เรียนจึงต้องได้รับการฝึกการอ่านเป็นหน่วยความคิด

3. เข้าใจประโยค (Sentence comprehension) เมื่อผู้อ่านสามารถเอา หน่วยความคิดย่อยสัมพันธ์กันจนได้ใจความประโยคแล้ว ผู้อ่านจะต้องมองเห็นความสัมพันธ์ระหว่าง คำ และระหว่างกลุ่มคำใน ประโยคด้วย

4. เข้าใจอนุเฉท (Paragraph) ผู้อ่านต้องสามารถใจความสำคัญของแต่ละประโยคแล้วนำใจความสำคัญนั้นมาสัมพันธ์กัน ดังนั้นผู้อ่านจึงควรเรียนรู้ประเภท หรือโครงสร้างแบบต่างๆ ของอนุเฉทเพื่อให้สามารถระบุประโยคที่เป็นใจความหลักในอนุเฉทนั้น จึงมองหาความสัมพันธ์ของประโยคอื่นๆ ซึ่งเป็นรายละเอียดในอนุเฉท

5. เข้าใจเนื้อหาทั้งหมด (Comprehension for larger unit) การเข้าใจเรื่องราวต่างๆทั้งหมดผู้อ่านต้องความสำคัญระหว่างข้อความสำคัญๆ ในแต่ละตอน เพื่อลำดับความคิดของเรื่องและต้องทราบว่าคุณเขียนแต่ละประเภทมีโครงสร้างข้อความแตกต่างกัน

Phillips (1984, p. 285, อ้างถึงใน อรอนงค์ อุ๋นแก้ว, 2562, หน้า 18) ได้เสนอองค์ประกอบสำคัญที่จะช่วยให้ผู้เรียนสามารถอ่านภาษาต่างประเทศด้วยความเข้าใจ 3 ประการ คือ

1. ความรู้ความสามารถด้านภาษา
2. สติปัญญาของผู้อ่านหรือความสามารถในการใช้เหตุผลและความคิดต่อสิ่งที่อ่าน
3. ประสบการณ์เดิมรวมทั้งความรู้รอบตัวทั่วไปของผู้อ่าน

จิตภา ฉันทานนท์ (2541, หน้า 52-59, อ้างถึงใน วราภรณ์ พรหมอินทร์, 2561, หน้า 26) ได้กล่าวว่า การอ่านเพื่อความเข้าใจมีองค์ประกอบอยู่ 3 ด้าน คือ

1. พื้นความรู้ด้านภาษา (Linguistic schema) หมายถึง ความรู้เกี่ยวกับตัวอักษร การสะกดคำคำศัพท์ โครงสร้างไวยากรณ์ และสัมพันธ์ภาพในข้อความ
2. พื้นความรู้ด้านเนื้อหา (Content schema) หมายถึง ความรู้เกี่ยวกับวัฒนธรรม ความรู้ รอบตัวทั่วไป และความรู้ในเนื้อหาเฉพาะด้านตามที่ปรากฏหรือเกี่ยวข้องกับเนื้อหาในบทอ่าน
3. พื้นความรู้ด้านโครงสร้างบทอ่าน (Formal schema) หมายถึง ความรู้เกี่ยวกับโครงสร้างของงานเขียนประเภทต่าง ๆ ที่ใช้ในการนำเสนอเนื้อหาในบทอ่าน ซึ่งมีอยู่หลายประเภท เช่น ประเภทบรรยาย (Descriptive) ประเภทเล่าเรื่อง (Narrative) ประเภทเปรียบเทียบ (Comparison)

ถิรวัฒน์ ตันทนิส (2550, หน้า 121) กล่าวว่าไว้ว่า การทำความเข้าใจในการอ่านนั้นต้องอาศัยองค์ประกอบหลายอย่างเพื่อช่วยให้ผู้อ่านเกิดความเข้าใจในการอ่าน ซึ่งองค์ประกอบของการอ่านเพื่อความเข้าใจสรุปได้ เป็น 3 องค์ประกอบใหญ่ๆ ได้แก่

องค์ประกอบที่ 1 ผู้อ่าน ซึ่งหมายถึงพื้นความรู้เดิมของผู้อ่าน รวมถึง ความสามารถของผู้อ่านในการใช้กลวิธีในการอ่าน

องค์ประกอบที่ 2 เนื้อเรื่องที่อ่าน ซึ่งหมายถึงความเข้าใจเนื้อเรื่องที่อ่าน เช่น คำศัพท์ ข้อเท็จจริงต่างๆ รวมถึงความเข้าใจ สิ่งที่อยู่นอกเหนือเนื้อเรื่องที่อ่าน เช่น ความหมายแฝง จุดมุ่งหมาย หรืออารมณ์ของ ผู้เขียนที่แฝงอยู่ในเนื้อเรื่อง

องค์ประกอบที่ 3 บริบทของการอ่าน เช่น ลักษณะ ของบทอ่าน รวมถึงสถานการณ์การอ่าน ซึ่งองค์ประกอบเหล่านี้จะมีการเปลี่ยนแปลง ไปตามวัยของผู้อ่าน นอกจากองค์ประกอบดังกล่าว จะมีผลต่อความเข้าใจในการอ่าน แล้วยังพบว่ามืองค์ประกอบบางประการในตัวผู้อ่านและจากบทอ่าน ที่อาจขัดขวาง ความเข้าใจในการอ่านได้

จากที่กล่าวมา สรุปได้ว่าการศึกษาที่ผู้อ่านจะเกิดความเข้าใจมืองค์ประกอบหลายอย่างได้แก่ พื้นความรู้เดิมหรือประสบการณ์เดิมและความสามารถของผู้อ่านในการใช้กลวิธีในการอ่าน เนื้อเรื่องที่อ่านซึ่งเป็นความเข้าใจเนื้อเรื่องที่อ่าน เช่น คำ42ศัพท์ ข้อเท็จจริงต่างๆ รวมถึงความเข้าใจ สิ่งที่อยู่ นอกเหนือเนื้อเรื่องที่อ่าน เช่น ความหมายแฝง จุดมุ่งหมายหรืออารมณ์ของ ผู้เขียนที่แฝงอยู่ในเนื้อ เรื่อง รวมทั้งบริบทของการอ่าน เช่น ลักษณะ ของบทอ่าน สถานการณ์การอ่าน

2. รูปแบบการประเมินความเข้าใจในการอ่าน

เนื่องจากการสอนอ่านในปัจจุบัน เน้นการทำทำความเข้าใจต่อเรื่องที่อ่านเป็นหลัก การวัด และประเมินผลการอ่าน จึงออกมาในรูปแบบของการทดสอบความเข้าใจในการอ่าน ซึ่งมีนักวิชาการ ให้ข้อเสนอแนะไว้ดังรายละเอียดต่อไปนี้

Hubley (2012) เสนอว่าการอ่านประเมินเป็น 5 รูปแบบ ในแต่ละแบบให้ประเมิน ทักษะย่อย (Subskills) ดังต่อไปนี้

1) ประเมินเรื่องทั้งหมด (Whole passage) ผู้สอนสามารถประเมินทักษะย่อย เช่น การ จับใจความหลัก การประเมินชนิดของบทอ่าน การระบุประเภทของผู้อ่านในบทอ่านชนิดต่าง ๆ การ เข้าใจจุดมุ่งหมายของผู้เขียน การเข้าใจทัศนคติของผู้เขียน การเรียงลำดับเหตุการณ์ การสรุปประเด็น สำคัญ และการแยกแยะข้อเท็จจริงออกจากความคิดเห็น เป็นต้น

2) ประเมินในส่วนวรรค (Section) เป็นการประเมินในประเด็นความเข้าใจหัวเรื่อง ความสัมพันธ์ของโครงสร้างเรื่องที่อ่าน และการจัดหมวดหมู่ข้อมูล

3) ประเมินในส่วนย่อหน้า (Passage) เป็นการประเมินความสามารถของผู้อ่านในการ ระบุใจความสำคัญและข้อความสนับสนุนหลัก เข้าใจประโยคใจความสำคัญ สามารถระบุการอ้างอิงใน ย่อหน้า และแยกความแตกต่างระหว่างข้อมูลทั่วไปกับตัวอย่างได้

4) ประเมินความสัมพันธ์ระหว่างประโยค (Across sentence boundaries) เป็นการ ประเมินความเข้าใจในการใช้คำเชื่อมความ และความสามารถในการระบุการอ้างอิงคำนามคำสรรพนาม เป็นต้น

5) ประเมินในส่วนของประโยค (Sentence) เป็นการประเมินทักษะย่อย เช่น การเดา ความหมายของคำศัพท์จากบริบท ความเข้าใจในการใช้คำบุพบทและสันธาน ความหมายจากการ เรียงลำดับคำ คำพ้องต่าง ๆ

สุภัทรา อักษรานุเคราะห์ (2531, หน้า 59-61, อ้างถึงใน พิมพ์นทร เทพปัญญา, 2556, หน้า 27) อธิบายว่า ทักษะการอ่านเป็นทักษะที่เสริมและพัฒนาการสอนทักษะอื่นๆ คือ ฟัง-พูด และ เขียน ฉะนั้น การประเมินผลอาจทำได้โดยการให้ผู้เรียนอ่านแล้วแสดงความเข้าใจในเรื่องที่อ่าน แสดง พฤติกรรมออกมาโดยใช้ทักษะการฟัง-พูด หรือ เขียนได้ดังตัวอย่างต่อไปนี้

- 1) แสดงความเข้าใจในเรื่องที่อ่านโดยการหาความหมายของคำศัพท์
- 2) แสดงความเข้าใจในเรื่องที่อ่านโดยการบอกเกี่ยวกับความสัมพันธ์ระหว่างประโยค
- 3) แสดงความเข้าใจในเรื่องที่อ่านโดยการหาความหมายจากประโยคที่กำหนดให้
- 4) แสดงความเข้าใจในเรื่องที่อ่านโดยการตอบคำถามเกี่ยวกับเรื่องที่อ่านโดยใช้คำถาม

แบบเลือกตอบ (Multiple Choices)

- 5) แสดงความเข้าใจในเรื่องที่อ่านโดยการเรียงลำดับสิ่งที่อ่านให้ถูกต้อง
- 6) แสดงความเข้าใจในเรื่องที่อ่านโดยการเติมคำลงในช่องว่างโดยใช้วิธีโคลซ (Cloze Test)
- 7) แสดงความเข้าใจในเรื่องที่อ่านโดยการแสดงความคิดเห็นเกี่ยวกับเรื่องที่อ่านอย่าง ง่าย ๆ
- 8) แสดงความเข้าใจในเรื่องที่อ่านโดยการตั้งชื่อเรื่อง

สุมิตรา อังวัฒนสกุล (2537, หน้า 215-220, อ้างถึงใน พิมพ์นทร เทพปัญญา, 2556, หน้า 27-28) ได้เสนอรูปแบบการวัด และการประเมินผลการอ่าน ดังนี้ การทดสอบขององค์ประกอบย่อย ๆ ของการอ่านเป็นการทดสอบย่อย ๆ ทางภาษา ได้แก่ คำศัพท์อันเป็นพื้นฐานในการอ่าน เหมาะ สำหรับผู้เรียนการอ่านในระดับเริ่มต้น หรือ ผู้ที่มีความสามารถในการอ่านจากการทดสอบในลักษณะ นี้ ได้แก่ การที่ครูอ่านออกเสียงคำศัพท์ และให้นักเรียนเลือกคำศัพท์ที่ตรงกับคำที่ออกเสียงหรือการใช้ เทคนิค “Same-Different” ให้นักเรียนเลือกหรือการให้คำหลัก (Key Words) มาให้แล้วนักเรียน เลือกคำที่ต่างออกไปจากคำหลักซึ่งมีจำนวนข้อมากและใช้เวลารวดเร็ว สรุปได้ว่า รูปแบบการวัดและ การประเมินผลการอ่าน 3 แบบ ได้แก่

1) การทดสอบความเข้าใจประโยค (Sentence Comprehension) เป็นการอ่านที่ เหมาะสำหรับนักเรียนที่เริ่มต้นฝึกฝนการอ่าน ความเข้าใจในประโยคหลายๆประโยคเรียง เรียงตามลำดับเรื่องราวและความสัมพันธ์เป็นบทความ

2) การทดสอบความเข้าใจในบทอ่าน (Passage Comprehension) การเลือกบทอ่านที่นักเรียนจะได้พบเห็นมีโอกาสใช้จริง ได้แก่ การประกาศโฆษณา เช่น ประกาศรับสมัครงาน จดหมายธุรกิจ อนุญาตขับซิ่งรถ ป้ายประกาศ

3) แบบทดสอบชนิดโคลซแบบมีตัวเลือก (Multiple Choice Cloze) แบบทดสอบโคลซชนิด Standard Cloze ซึ่งไม่มีตัวเลือกนั้นไม่เหมาะที่จะใช้วัดความเข้าใจในการอ่านเพราะแบบทดสอบชนิดนี้จะเน้นไปทาง Productive Skill เสียมากกว่าจึงมีการดัดแปลงเป็นแบบมีตัวเลือกเหมาะสมกับทักษะอ่านซึ่งเป็น Receptive Skill นอกจากนี้ ยังมีแบบทดสอบโคลซอีกแบบหนึ่ง แต่มีลักษณะตรงกันข้าม เรียกว่า Editing Test กล่าวคือ จะกำหนดข้อความมาให้ให้นักเรียนอ่านแล้วเลือกคำเกินมาที่ไม่เกี่ยวข้องกับความหมายในประโยคออกจากบทอ่าน เหมาะสำหรับวัดความสามารถในการอ่านเร็ว

สมุทฺร เซ็นชานนิต (2549, หน้า 73, อ้างถึงใน วราภรณ์ พรหมอินทร์, 2561, หน้า 26) ได้แบ่งความเข้าใจออกเป็น 2 ประเภทใหญ่ ๆ ดังนี้

1. ความเข้าใจแบบทันที (Receptive Comprehension) เป็นความเข้าใจที่ต้องอาศัยการรู้ความหมายของคำศัพท์ต่าง ๆ ที่ผู้เขียนใช้เป็นส่วนใหญ่ ต้องรู้คำศัพท์ต่าง ๆ มากพอ จะต้องเข้าใจความหมายที่สำคัญ ๆ ของสำนวนและประโยคต่าง ๆ ได้ถูกต้องตามจุดมุ่งหมายของผู้เขียน ขณะที่อ่านต้องมีสมาธิแน่วแน่จึงจะเข้าใจสิ่งที่อ่านได้โดยตลอด

2. ความเข้าใจแบบไตร่ตรอง (Reflective Comprehension) คือความเข้าใจที่ต้องอาศัยความรอบรู้และทักษะความสนใจหลาย ๆ ด้าน เป็นหลักใหญ่ เพื่อนำมาใช้เป็นเครื่องมือช่วยทำความเข้าใจเรื่องได้อย่างถูกต้องตามจุดมุ่งหมายของผู้เขียนว่าต้องการอธิบายชี้แจง สั่งสอน ชักจูง หรือให้ความเพลิดเพลิน นอกจากนี้ยังสามารถพินิจพิจารณาและไตร่ตรองหาข้อมูลสรุปตามความคิดที่ผู้เขียนต้องการความเข้าใจแบบนี้ต้องพึงเหตุผล และอาศัยการเปรียบเทียบระหว่างประสบการณ์ต่าง ๆ หรือสิ่งที่เคยพบเห็นมาแล้วในชีวิตเป็นสำคัญ

อัจฉรา วงศ์โสธร (2539, หน้า 154-155) กล่าวถึง การประเมินทักษะการอ่านสามารถใช้เกณฑ์ที่กำหนดตามส่วนประกอบของภาษาแบบแยกย่อย และเกณฑ์ที่กำหนดตามความสามารถในการรับสาร หรือ อาจใช้เกณฑ์ทั้งสองแบบประยุกต์เข้าด้วยกันโดยใช้น้ำหนักเกณฑ์แบบรวมมากกว่าเกณฑ์แบบย่อย ดังรายละเอียดต่อไปนี้

1) เกณฑ์กำหนดความสามารถทางการอ่านที่เป็นเกณฑ์แบบย่อย ได้แก่

1.1 ความรู้ในด้านศัพท์ หมายถึง ความสามารถในการเข้าใจคำศัพท์และสำนวนต่าง ๆ

1.2 ความรู้ในด้านไวยากรณ์ หมายถึง ความสามารถในการใช้ความรู้ด้านไวยากรณ์ ในการทำความเข้าใจเกี่ยวกับคำสรรพนาม ความเชื่อมโยงของเนื้อความ เช่น การใช้คำสันธานคำบุพบทที่กำหนดหน้าที่ของภาษาว่าเป็นการขอร้อง เชื้อเชิญ หรือขออนุญาต เป็นต้น

2) เกณฑ์กำหนดความสามารถทางการอ่านแบบรวม ได้แก่

2.1 ความสามารถในการเรียบเรียงความ หมายถึง ความสามารถในการทำความเข้าใจบทอ่าน และสามารถตอบคำถามที่ให้เรียบเรียงถ้อยคำใหม่ โดยให้ได้ใจความเดิมหรือสามารถตอบคำถามแบบเลือกตอบ และแบบเรียงลำดับข้อความได้

2.2 ความสามารถในการอ่านข้อมูลที่เป็นรายละเอียด หมายถึง ความสามารถในการโยงรายละเอียดที่เกี่ยวข้องเข้ากับใจความสำคัญของเรื่องได้ว่า เป็นรายละเอียดสนับสนุนหรือเป็นรายละเอียดที่ขัดแย้งกันเพื่อให้ข้อมูลตรงกันข้าม ตลอดจนเข้าใจความสัมพันธ์ระหว่างรายละเอียดต่าง ๆ

2.3 ความสามารถในการอ่านจับใจความสำคัญ หมายถึง ความสามารถในการระบุแก่นเรื่อง หัวเรื่อง และใจความสำคัญของเรื่องที่อ่านได้

2.4 ความสามารถในการวิเคราะห์และประเมินความสัมพันธ์ของเนื้อความ และสุนทรียศาสตร์ของการใช้ภาษา หมายถึง ความสามารถในการใช้ความรู้ด้านศัพท์ ไวยากรณ์ ความเข้าใจสิ่งที่อ่านและความรู้เกี่ยวกับรูปแบบลีลาภาษาที่ใช้ในบทอ่านที่เป็นตัวกระตุ้น วิเคราะห์ ประเมิน และสรุปบทอ่านได้ว่าเป็นสารประเภทใด ใช้ลีลาภาษาแบบเป็นทางการหรือไม่ เข้าใจเจตนาที่ศนคคติของผู้เขียนที่แฝงอยู่ สามารถวิพากษ์วิจารณ์ถึงเหตุและผลที่เกิดขึ้นได้ ตลอดจนสามารถประเมินบทอ่านได้ว่า มีความชัดเจนเข้าสู่ประเด็นอย่างไรไม่อ้อมค้อมและใช้ภาษาได้กระชับ ความสามารถในระดับนี้เป็นระดับสูง ซึ่งต้องอาศัยความรู้ในระดับต้น ๆ เป็นพื้นฐาน

ผัสนพรรณ ธนอมพงษ์ชาติ และคณะ (2556, หน้า 60, อ้างถึงใน อรอนงค์ อุ่นแก้ว, 2562) กำหนดพฤติกรรมการอ่านภาษาอังกฤษเพื่อความเข้าใจ ดังนี้

Literal Comprehension หมายถึง ความสามารถในการจับประเด็นสำคัญของเรื่องที่อ่านจากข้อความที่ระบุชัดเจนอยู่ในบทที่อ่าน

Interpretative Comprehension หมายถึง ความสามารถในการแปลความพิจารณาสาเหตุ และผลสรุปลงความเห็นจากสิ่งที่ไม่ได้ระบุไว้ในเรื่องที่อ่าน พฤติกรรมที่ผู้เรียนสามารถนำเนื้อหาของเรื่องที่อ่านมาประยุกต์ใช้ให้เกิดประโยชน์กับชีวิตของตนเอง และพฤติกรรมที่ผู้เรียนสามารถบูรณาการเชื่อมโยงหลักการสำคัญของเรื่องที่อ่าน เข้ากับวิชาอื่นๆและเข้ากับชีวิตของตนเอง

Applied Comprehension หมายถึง ความสามารถในการนำความรู้ไปสัมพันธ์กับความรู้เดิมของตนหรือสัมพันธ์กับความรู้จากวิชาต่างๆและนำไปใช้ในสถานการณ์ใหม่ได้ สำหรับ

พฤติกรรมที่ผู้เรียนเข้าใจเพื่อนมนุษย์ชุมชน และสังคมดีขึ้นเป็นความเข้าใจและเห็นอกเห็นใจที่เกิดขึ้น ภายหลังจากการอ่านเนื้อเรื่องและพฤติกรรมที่ผู้เรียนเกิดความใส่ใจ และมองเห็นคุณค่าของสิ่งที่เรียน ภายหลังจากการอ่านเนื้อเรื่อง

Evaluation Comprehension หมายถึง ความสามารถในการตัดสินใจประเมินผล สิ่งที่อ่าน โดยอาศัยความรู้ ประสบการณ์ของผู้อ่านในเรื่องดังกล่าว

2.1 รูปแบบของแบบทดสอบในการประเมินความเข้าใจในการอ่านภาษาอังกฤษ

ฟิน็อคเชียโร และซาโก (Finocchiaro & Sako, 1983, p. 16, อ้างถึงใน วรา ภรณ์ พรหมอินทร์, 2561, หน้า 29) ได้กล่าวไว้ว่า รูปแบบของแบบทดสอบที่ใช้ในการประเมินผลความ เข้าใจในการอ่านภาษาอังกฤษที่นิยมใช้มี 2 แบบ คือ

1) แบบทดสอบอัตนัย (Subjective test) ได้แก่ แบบทดสอบความเรียงที่ให้ ผู้เรียนตอบคำถามจากเรื่องที่อ่าน โดยเขียนคำตอบเป็นประโยคหรือข้อความยาว ๆ

2) แบบทดสอบแบบปรนัย (Objective test) ได้แก่ แบบทดสอบแบบเลือกตอบ แบบถูกผิด แบบจับคู่ และแบบเติมคำ เป็นต้น

แมดสัน (Madsen, 1983, p. 97) กล่าวว่า ในทางปฏิบัติมักนิยมใช้แบบทดสอบ ปรนัยชนิดเลือกตอบ (Multiple choice) เพราะการตรวจให้คะแนนสะดวก รวดเร็ว มีความเที่ยงตรง ต่อการคิดค่าคะแนน ใครเป็นผู้ตรวจก็ได้ คะแนนที่ออกมามีความเที่ยงตรง รวมทั้งเป็นแบบทดสอบ ที่มีค่าความเชื่อมั่นสูง สะดวกในการให้คะแนน และใช้เวลาในการทำแบบทดสอบน้อย นอกจากนั้น ยังสามารถถามได้ครอบคลุมเนื้อหาหลาย ๆ ด้าน และสามารถใช้วัดกับผู้สอบทุกระดับ

ฟิน็อคเชียโรและซาโก (Finocchiaro & Sako, 1983, pp. 131-135, อ้างถึงใน วราภรณ์ พรหมอินทร์, 2561, หน้า 30) ได้เสนอแนวทางในการวัดและประเมินผลความเข้าใจในการ อ่านควรครอบคลุมถึงสิ่งต่อไปนี้

1) คำศัพท์ การเดาคำศัพท์โดยใช้ข้อความบริบทที่อยู่รอบ ๆ เพื่อให้ได้ ความหมายที่ถูกต้องและเหมาะสมกับคำศัพท์นั้นๆ

2) ใจความสำคัญหรือใจความหลักของเรื่องที่อ่าน

3) รายละเอียดในเนื้อเรื่องที่อ่าน

4) การสรุปเรื่องและตีใจความที่ปรากฏโดยนัย หรือเรื่องที่ไม่ปรากฏอยู่ในข้อความ

5) อารมณ์ ความรู้สึก และเจตนาอารมณ์ของผู้เขียน

6) ความสัมพันธ์ของข้อความหรือประโยคที่ปรากฏในเนื้อเรื่อง

7) โครงสร้างของประโยค โดยการแปลความประโยคหรือโดยการกำหนด ประโยคมาให้ 1 ประโยค แล้วเลือกประโยคเทียบเคียงที่มีความหมายตรงกับประโยคที่กำหนดให้

วารภรณ์ พรหมอินทร์ (2561, หน้า 30) การวัดและประเมินผลการอ่านสามารถครอบคลุมถึงด้านวลี คำ ประโยค ข้อความ ใจความสำคัญ และการตีความในการประเมิน สามารถใช้เกณฑ์ความสามารถแบบแยกย่อยหรือแบบรวมก็ได้ หรืออาจใช้เกณฑ์ทั้งสองแบบประยุกต์รวมกัน สิ่งสำคัญที่ควรคำนึงถึงคือ รูปแบบในการประเมินและเนื้อหา ต้องเหมาะสมกับความสามารถในการใช้ภาษา ประสบการณ์เดิมและพื้นฐานความรู้ของผู้อ่าน รวมทั้งควรพิจารณาให้สอดคล้องกับจุดประสงค์ของรายวิชา

จากที่กล่าวมา การประเมินความสามารถการอ่านเพื่อความเข้าใจ มีหลากหลายรูปแบบไม่ว่าจะเป็นเกณฑ์ความสามารถแบบแยกย่อยหรือแบบรวม ประเมินความเข้าใจที่เกิดขึ้นทันทีหรือแบบไต่ตรง หรือประเภทประเมินแบบแบ่งเป็นส่วน ๆ เช่น ประเมินความเข้าใจจากเนื้อเรื่องทั้งหมด ประเมินในวรรคเดียวกัน ประเมินในส่วนย่อหน้า ประเมินความสัมพันธ์ระหว่างประโยค และประเมินในส่วนประโยค นิยมใช้แบบทดสอบปรนัยชนิดเลือกตอบ (Multiple choice) เพราะการตรวจให้คะแนนสะดวก รวดเร็ว มีความเที่ยงตรงต่อการคิดค่าคะแนน เป็นแบบทดสอบที่มีค่าความเชื่อมั่นสูง ที่สำคัญคือใช้เวลาในการทำแบบทดสอบน้อยสามารถถามได้ครอบคลุมเนื้อหาหลาย ๆ ด้าน และสามารถใช้วัดกับผู้สอบทุกระดับ

2.2 ลักษณะของข้อคำถามการอ่านเพื่อความเข้าใจ

สมิตรา อังวัฒนสกุล (2537, หน้า 215-220, อ้างถึงใน พิมพ์นทร เทพปัญญา, 2556, หน้า 27-28) ได้เสนอลักษณะของข้อคำถามเกี่ยวกับการอ่าน (Type of questions)

1) สำหรับการทดสอบความเข้าใจในระดับประโยค (Sentence Comprehension) ลักษณะของข้อคำถามความเข้าใจในระดับประโยคมีดังต่อไปนี้ เลือกรูปภาพให้ตรงกับประโยคที่กำหนดให้ ตอบคำถามโดยใช้รูปภาพประกอบ เลือกใช้วลีและประโยคเป็นตัวแทน เช่น การสรุปถ้อยแถลง/ การถอดความ (Paraphrase) ประโยคหรือวลีที่กำหนด อาทิ การสรุปถ้อยแถลงจากภาพเป็นประโยคจากประโยคเป็นภาพ การถอดความจากประโยคเดิมถูกเรียบเรียงเป็นประโยคใหม่ที่มีความหมายเดียวกัน เป็นต้น

2) สำหรับการทดสอบความเข้าใจในบทอ่าน (Passage Comprehension) ประเภทของคำถามที่สำคัญ ได้แก่ ข้อคำถามแบบให้สรุปเนื้อเรื่อง (Paraphrase) ข้อคำถามแบบให้สังเคราะห์สิ่งที่ได้จากการอ่าน (Synthesis) โดยผู้อ่านจะต้องประมวลข้อมูลต่างๆ ที่ได้จากรื่องเข้าด้วยกัน เพื่อตอบคำถามข้อคำถามให้สรุปอ้างอิงจากสิ่งที่อ่าน (Inference)

3) สำหรับแบบทดสอบชนิดโคลซแบบมีตัวเลือก (Multiple Choice Cloze) แบบทดสอบโคลซชนิด Standard Cloze ซึ่งไม่มีตัวเลือกนั้นไม่เหมาะที่จะใช้วัดความเข้าใจในการอ่านเพราะแบบทดสอบชนิดนี้จะเน้นไปทาง Productive Skill เสียมากกว่าจึงมีการดัดแปลงเป็นแบบมีตัวเลือกเหมาะสมกับทักษะอ่านซึ่งเป็น Receptive Skill นอกจากนี้ ยังมีแบบทดสอบโคลซ

อีกแบบหนึ่ง แต่มีลักษณะตรงกันข้าม เรียกว่า Editing Test กล่าวคือ จะกำหนดข้อความมาให้ให้นักเรียนอ่านแล้วเลือกคำเกินมาที่ไม่เกี่ยวข้องกับความหมายในประโยคออกจากบทอ่าน เหมาะสำหรับวัดความสามารถในการอ่านเร็ว

สถาบันจุฬาลงกรณ์มหาวิทยาลัย (2558) คำถามเกี่ยวกับการอ่าน (Type of questions) โดยทั่วไปแบ่งออกเป็น 3 ประเภท ได้แก่

1) คำศัพท์ (Vocabulary) หรือที่รู้จักกันในทางเทคนิคว่าคำศัพท์ในคำถามตามบริบท จะถามคุณเกี่ยวกับคำที่อยู่ในเนื้อเรื่อง เช่น จากเรื่องคำว่า 'heading' (บรรทัดที่ 5) หมายถึง ข้อใดเป็นต้นคำศัพท์ในคำถามตามบริบทเหมาะสำหรับผู้สอบ เพราะเป็นข้อคำถามที่ตรงไปตรงมาและตอบได้รวดเร็ว

2) คำถามเฉพาะเจาะจง (Specific question) คือคำถามที่มีคำตอบอยู่ในส่วนหนึ่งส่วนใดอยู่ในข้อความหนึ่งเท่านั้น โดยที่คำถามจะถามถึงบางสิ่งบางอย่างไม่น้อยกว่าหนึ่งคำ แต่ยังคงอยู่ในขอบเขตที่แคบและเป็นสิ่งที่มุ่งเน้น ได้แก่ คำถามระบุเฉพาะในบรรทัดหรือในแต่ละย่อหน้า คำถามระบุเฉพาะส่วนหนึ่งของข้อมูลที่ไม่ได้ระบุแน่นอนด้วยหมายเลขบรรทัด ผู้สอบต้องค้นหาข้อมูลเอง โดยการอนุมานเชื่อมโยง หรือสรุปเกี่ยวกับข้อมูลเฉพาะในบทความ

3) สารสำคัญของเรื่อง (Main Idea) เป็นคำถามเกี่ยวกับสารสำคัญของเรื่องหรือใจความสำคัญของเรื่องเป็นคำถามการอ่านที่ยากที่สุดในการทดสอบ และมักใช้เวลาในการตอบนานที่สุด เพราะวัดความเข้าใจเนื้อเรื่องโดยรวม

3. แนวคิดทฤษฎีที่เกี่ยวข้องกับการอ่านเพื่อความเข้าใจ

การอ่านเพื่อความเข้าใจเป็นทักษะทางปัญญาซึ่งมีความเกี่ยวข้องกับการเรียนรู้ตามทัศนะของนักจิตวิทยาในกลุ่มคอนสตรัคติวิสต์เชิงปัญญา (Cognitive Constructivism) ซึ่งเชื่อว่าการอ่าน คือ กระบวนการทางปัญญาที่บุคคลใช้ในการค้นหาความหมายของสิ่งที่อ่านโดยอาศัยความรู้พื้นฐานหรือโครงสร้างความรู้ที่สะสมไว้หรือมวลประสบการณ์ เพื่อความเข้าใจและเกิดการเรียนรู้ (Graves and Graves, 1998, อ้างถึงใน ฉัตรปวีณ อัมภา, 2555, หน้า 11) ทฤษฎีการเรียนรู้ตามแนวคอนสตรัคติวิสต์ (Constructivism) ที่สำคัญมีสองกลุ่มใหญ่ๆ คือ กลุ่มแนวคิดคอนสตรัคติวิสต์เชิงสังคม (Social Constructivism) และกลุ่มแนวคิดคอนสตรัคติวิสต์เชิงปัญญา (Cognitive Constructivism) ในวรรณกรรมบางเล่มเรียกปัญญานิยมหรือพุทธิปัญญานิยม เพื่อเข้าใจตรงกันในการศึกษาครั้งนี้ใช้คำว่าทฤษฎีการเรียนรู้พุทธิปัญญานิยม

3.1 ทฤษฎีการเรียนรู้พุทธิปัญญานิยม (Cognitive Constructivism)

มีรากฐานทางปรัชญาของทฤษฎีมาจากความพยายามที่จะเชื่อมโยงประสบการณ์เดิมกับประสบการณ์ใหม่ด้วยกระบวนการที่พิสูจน์อย่างมีแนวคิดว่ามีแนวคิดที่ว่า มนุษย์เราต้อง “สร้าง” ความรู้ด้วยตนเองโดยผ่านทางประสบการณ์ ซึ่งประสบการณ์เหล่านี้จะกระตุ้นให้ผู้เรียนสร้างโครงสร้างทางปัญญาหรืออาจกล่าวได้ว่า การสร้างความรู้เป็นกระบวนการที่เกิดขึ้นภายในของผู้เรียน ผู้เรียนสร้างความรู้โดยการนำประสบการณ์หรือสิ่งที่พบเห็นในสิ่งแวดล้อมหรือสารสนเทศใหม่ที่ได้รับมาเชื่อมโยงกับความรู้ความเข้าใจที่มีอยู่เดิมเรียงตัวซ้อนเป็นชั้นของความเข้าใจถูกเรียกว่า โครงสร้างทางปัญญา (Cognitive structure) หรือ สกีม่า (Schema) แต่ละบุคคลอาจสร้าง ความหมายที่แตกต่างกัน เนื่องจากมีประสบการณ์หรือพื้นความรู้เดิม (Background Knowledge) ที่แตกต่างกัน สกีม่าเหล่านี้สามารถเปลี่ยนแปลงได้ (Change) ขยายออกได้ (Enlarge) และซับซ้อนขึ้นได้โดยผ่านกระบวนการการดูดซึม (Assimilation) และการปรับเปลี่ยน (Accommodation) (อนุชา โสมาบุตร, 2557)

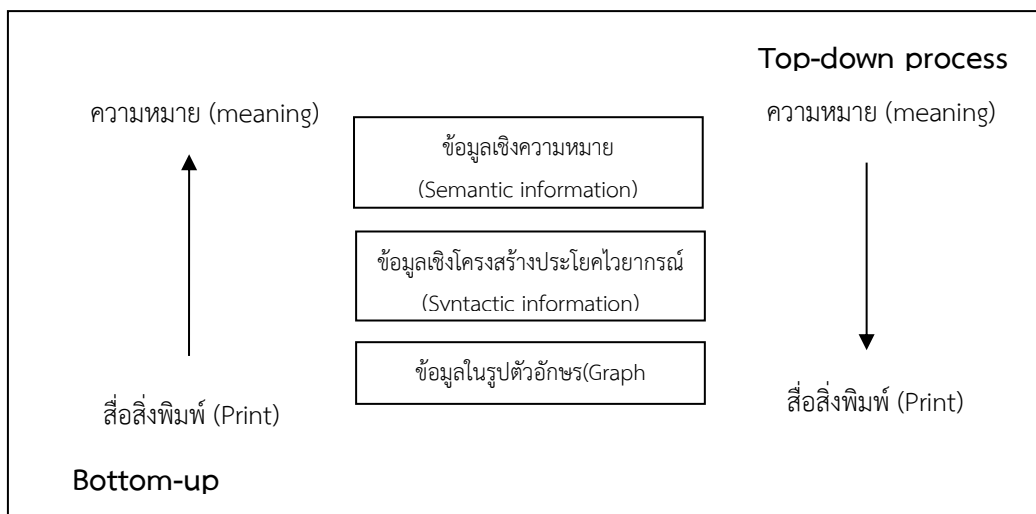
นักจิตวิทยาภาษาศาสตร์มองว่าการอ่านเพื่อความเข้าใจเป็นกระบวนการปฏิสัมพันธ์ระหว่างเรื่องที่อ่านกับความรู้เดิมและความรู้ทั่วไปของผู้อ่าน การอ่านยึดผู้อ่านเป็นศูนย์กลางความรู้ใหม่ มโนทัศน์ใหม่จะมีความหมายก็ต่อเมื่อ สิ่งเหล่านั้นสัมพันธ์กับสิ่งที่ผู้อ่านได้รู้มาแล้ว (Carrell and Eisterhold, 1983, pp. 553 -573) ผู้เขียนจะถ่ายทอดความคิดของตนผ่านตัวอักษร และผู้อ่านพยายามทำความเข้าใจความคิดของผู้เขียนจากตัวอักษรนั้น ทำให้เกิดปฏิสัมพันธ์ระหว่างภาษาและความคิดขึ้น ผู้ที่มีความสามารถในการอ่านดีจะสร้างความหมายให้สอดคล้องกับความคิดที่ผู้เขียนต้องการสื่อ รู้จักใช้กลวิธีในการอ่านเลือกใช้ตัวชี้แนะ (Clue) ที่เหมาะสม รวมถึงการใช้พื้นความรู้เดิมและความรู้ทางภาษาที่สัมพันธ์ทำในการทำความเข้าใจเรื่อง ที่อ่าน นอกจากนี้ ความเข้าใจในการอ่านยังขึ้นอยู่กับความสนใจ แรงจูงใจ เจตคติ ความเชื่อ ความรู้สึก และจุดประสงค์ของผู้อ่านด้วย แต่ผู้ที่ประสบปัญหาในการอ่านจะไม่สามารถตีความหมายเรื่อง ที่อ่านให้ตรงกับความคิดของผู้เขียน เนื่องมาจากพื้นความรู้เดิม กลวิธี วัฒนธรรมของผู้เขียนและผู้อ่านแตกต่างกัน ในขณะที่ อ่านกระบวนการเริ่มต้นด้วยการรับรู้ข้อความตามตัวอักษร ทำความเข้าใจและสร้างความหมาย ความคิดรวบยอด (Concept) ของข้อความที่อ่าน ประเมิน และปรับความรู้ใหม่ให้สอดคล้องกับพื้นความรู้เดิมและประสบการณ์ของผู้อ่าน ซึ่งในการปฏิสัมพันธ์นี้ประกอบด้วย การดึงข้อมูล ความหมาย ไวยากรณ์ การเรียงลำดับคำในวลีและประโยค อีกส่วนเป็นการตอบสนองความคิดโดยพยายามทำความเข้าใจและตัดสินสารที่อ่าน (ถิรวัฒน์ ต้นทนิส, 2550, หน้า 122-123) ทฤษฎีการอ่านที่ได้รับอิทธิพลของแนวคิดคอนสตรัคติวิสต์เชิงปัญญาคือทฤษฎีโครงสร้างความรู้เป็นทฤษฎีการอ่านแนวจิตวิทยาภาษาศาสตร์ การอ่านเป็นกระบวนการที่ผู้อ่านต้องใช้ความคิดเป็นอย่างมาก เนื่องจากความหมายของข้อความที่สรรค์สร้างขึ้นจากบทอ่านนั้น ได้รับอิทธิพลอย่างมากจาก

กระบวนการทางสมองที่ผู้อ่านทุ่มเทให้กับกระบวนการอ่าน การเรียนรู้เกิดขึ้นเมื่อผู้เรียนสามารถเชื่อมโยงความคิดกับโครงสร้างทางปัญญาโดยอาศัยความเข้าใจและความ ระหว่างสิ่งที่รู้อยู่เดิมกับสิ่งที่เรียนรู้ใหม่เพื่อใช้ในการทำความเข้าใจข้อมูลที่ได้รับจากการอ่าน (Graves and Graves, 1998, อ้างถึงใน ฉัตรปวีณ อัมภา, 2555, หน้า 10)

คาร์เรลและไอสเตอร์โฮลด์ (Carrell & Eisterhold, 1988, อ้างถึงใน ถิรวัฒน์ ตันทนิส, 2550, หน้า 124-125) อธิบายว่า รูปแบบการอ่านภายใต้แนวคิดโครงสร้างความรู้ 3 รูปแบบ คือ กระบวนการอ่านจากองค์ประกอบย่อยสู่ใจความหลัก (Bottom-up process) กระบวนการอ่านจากใจความหลักสู่องค์ประกอบย่อย (Top-down process) และกระบวนการอ่านแบบปฏิสัมพันธ์ (Interactive model) มีรายละเอียดดังต่อไปนี้

1) กระบวนการอ่านจากองค์ประกอบย่อยสู่ใจความหลัก (Bottom-up process) หรือกระบวนการอ่านจากฐานขึ้นไปยอด เกิดจากการตีความโดยอาศัยสารในบทอ่าน สารในบทอ่าน จึงเป็นตัวกระตุ้นให้ เกิดกระบวนการอ่าน จึงเรียกอีกอย่างหนึ่งว่า กระบวนการยึดเนื้อหาในบทอ่าน เป็นหลัก (Text-based processing) กระบวนการอ่านแบบนี้เป็นการวิเคราะห์จากบทอ่านสู่สมอง หรือเรียกว่าเป็นรูปแบบการผลักดันข้อมูล (Data-driven model) ผู้อ่านจะเริ่มจากการเข้าใจ ตัวอักษร คำ ประโยค ไปจนถึงหน่วยของภาษาระดับข้อความ แล้วถอดรหัส (Decode) ข้อความที่อ่านโดยอาศัย ความรู้ด้านคำศัพท์ โครงสร้าง ไวยากรณ์ และความสัมพันธ์ระหว่างประโยคในการทำให้เกิดความเข้าใจ ในบทอ่าน ดังนั้น ถ้าผู้อ่านไม่สามารถทำความเข้าใจในสิ่งที่อ่าน อาจเนื่องมาจากผู้อ่านขาดความรู้ ความสามารถในการอ่านด้านภาษา ผู้อ่านจะเข้าใจได้โดยการรับข้อมูลเข้ามา ข้อมูลเหล่านี้ต้องมีคุณสมบัติ เหมาะสมกับโครงสร้างความรู้เพิ่มระดับพื้นฐานโครงสร้างความรู้เดิมเหล่านี้ จะจัดเรียงเป็นชั้น ๆ ขึ้นไป ยอด ในงานวิจัยบางเรื่องจึงถูกเรียกว่า รูปแบบการประมวลผลความจากฐานขึ้นไปยอด

2) กระบวนการอ่านจากใจความหลักสู่องค์ประกอบย่อย (Top-down process) หรือกระบวนการอ่านจากยอดลงมาฐาน เป็นกระบวนการอ่านที่ยึดพื้นความรู้ของผู้อ่านเป็นหลัก (Knowledge-based processing) กระบวนการอ่านแบบนี้เป็นการวิเคราะห์จากสมองสู่บทอ่าน หรือเรียกว่าเป็นรูปแบบการผลักดันความคิด ผู้อ่านจะใช้พื้นความรู้และประสบการณ์เดิม ตลอดจน ความรู้รอบตัวในเรื่องที่เกี่ยวข้องกับบทอ่านเป็นส่วนสำคัญในการสร้างความหมาย และทำความเข้าใจ ในสิ่งที่อ่านโดยอาศัยข้อมูลทางภาษา (คำ ประโยค และข้อความ) จากบทอ่านเป็นเพียงองค์ประกอบ พื้นฐานในการคาดเดาเรื่องที่อ่านเท่านั้น ซึ่งความหมายที่แท้จริงจะอยู่ที่ตัวผู้อ่าน



ภาพ 4 กระบวนการอ่านแบบ Bottom-up และ top-down

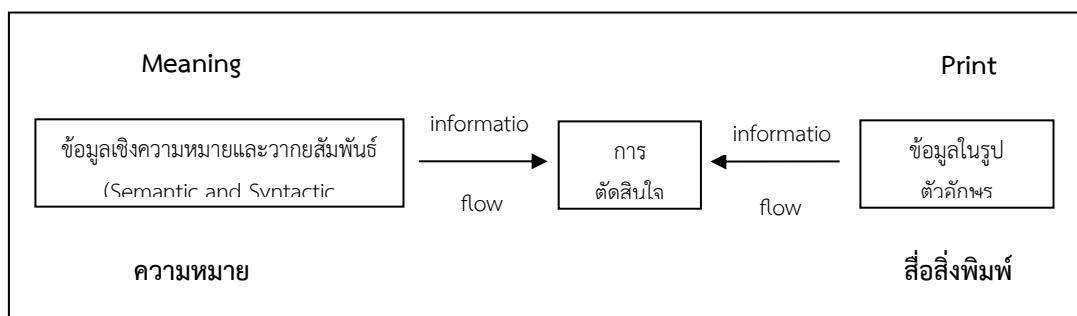
(ฉลอง รัตนพงษ์, 2557, หน้า 55)

จากภาพ 4 สรุปได้ว่า Bottom-up process เป็นการอ่านที่เน้นเนื้อหาที่อ่านเริ่มจากการเข้าใจ ความหมาย ตัวอักษร คำ ประโยค ไปจนถึงหน่วยของภาษาระดับข้อความ แล้วถอดรหัสโดยอาศัยความรู้ด้านคำศัพท์ โครงสร้าง ไวยากรณ์ และความสัมพันธ์ระหว่างประโยคในการทำให้เกิดความเข้าใจในบทอ่าน และ Top-down process เป็นการอ่านที่เน้นการใช้พื้นความรู้และประสบการณ์เดิมสร้างความหมายและทำความเข้าใจในสิ่งที่อ่าน โดยอาศัยข้อมูลทางภาษา อาทิ คำ ประโยค และข้อความ

อย่างไรก็ตามการอ่านที่จะประสบความสำเร็จได้นั้น ไม่ได้ใช้เฉพาะกระบวนการอ่านจากองค์ประกอบย่อยสู่ใจความหลัก หรือจากใจความหลักสู่องค์ประกอบย่อยเพียงอย่างเดียวหนึ่งเท่านั้น แต่จะต้องใช้ทั้ง 2 กระบวนการอย่างมีปฏิสัมพันธ์กัน

3) กระบวนการอ่านแบบปฏิสัมพันธ์ (Interactive Model) รูปแบบการอ่านแบบปฏิสัมพันธ์ (Interactive Approach) ในการอ่านภาษาที่ 1 เป็นกระบวนการอ่านที่ผู้อ่านต้องใช้ความรู้ในเรื่องโครงสร้างประโยค(Syntactical) ความหมาย (Semantic) คำศัพท์ (Lexical) และการสะกดคำ (Orthographic) มาเป็นตัวทำให้เกิด ปฏิสัมพันธ์ ระหว่างผู้อ่าน (Reader) กับสาร (Text) ที่อ่าน (Rumelhart, 1977, อ้างถึงใน ธีรวัฒน์ ต้นทนิส, 2550, หน้า 124-125) การอ่านแบบปฏิสัมพันธ์ จึงเป็นพฤติกรรมทางความรู้ความคิดที่เริ่มต้นกระบวนการอ่านด้วย การที่ผู้อ่านใช้โครงสร้างทางความรู้ความคิดในสมอง (Cognitive structure) โดยที่ผู้อ่าน จะใช้ความรู้ทางภาษา ตั้งแต่ตัวอักษร คำ ประโยค ความสัมพันธ์ทางไวยากรณ์ ในการ วิเคราะห์ ในขณะที่เดียวกันผู้อ่านก็จะใช้พื้นความรู้และประสบการณ์เดิมที่มีอยู่ในการ ทำความเข้าใจเรื่องที่อ่านด้วย ปฏิสัมพันธ์ที่เกิดขึ้นใน

ลักษณะของการอ่านแบบนี้เป็น ปฏิสัมพันธ์ของความรู้ในตัวผู้อ่าน กับสารกระบวนการอ่านในลักษณะนี้เป็นกระบวนการ อ่านที่มีความสอดคล้องกันระหว่างสารกับโครงสร้างความรู้เดิมที่มีอยู่



ภาพ 5 กระบวนการอ่านแบบปฏิสัมพันธ์

(ฉลอง รัตนพงษ์, 2557, หน้า 56)

จากภาพ 5 กระบวนการอ่านแบบปฏิสัมพันธ์ (Interactive Model of Reading) เกิดจากการนำรูปแบบทั้งสองได้แก่ โมเดล Bottom-up process และ Top-down process มารวมกันเป็นการลดจุดอ่อนหรือลดข้อจำกัดเพื่อให้ผู้อ่านประสบความสำเร็จในการอ่านเพื่อความเข้าใจมากขึ้น

การอ่านแนวจิตวิทยาภาษาศาสตร์ มีแนวคิดที่ว่า ความเข้าใจในการอ่านเป็นกระบวนการปฏิสัมพันธ์ระหว่างความรู้เดิมและความรู้ทั่วไปของผู้อ่าน การอ่านแนวนี้ยึดผู้อ่านเป็นศูนย์กลางความรู้ใหม่จะมีความหมายก็ต่อเมื่อสิ่งเหล่านี้สอดคล้องกับสิ่งที่ผู้อ่านได้รู้มาแล้ว ทฤษฎีการอ่านตามแนวจิตวิทยาภาษาศาสตร์ การเรียนภาษาอังกฤษเป็นภาษาต่างประเทศหรือการเรียนภาษาอังกฤษเป็นภาษาที่สอง (English as a Foreign Language; EFL/ /English as a Second Language; ESL) ทฤษฎีนี้ได้รับอิทธิพลมาจากแนวความคิดของนักจิตวิทยาภาษาศาสตร์ โดยเฉพาะอย่างยิ่ง Goodman's psycholinguistic model of reading หรือรูปแบบการอ่านทางภาษาศาสตร์เชิงจิตวิทยาในการอ่านภาษาที่หนึ่งของ กูดแมน (Goodman, 1987, pp. 8-9, อ้างถึงใน ธีรวัฒน์ ตันพนิส, 2550, หน้า 125-127) ลักษณะของการอ่านในรูปแบบนี้ได้อธิบายแนวคิดที่ว่า การอ่านเป็นเกมการเดาทางภาษาศาสตร์เชิงจิตวิทยา กล่าวคือ การอ่านเป็นการสร้างความหมายซึ่งผู้อ่านใช้ความรู้และประสบการณ์เดิมทำความเข้าใจในเรื่องที่อ่าน ในกระบวนการอ่าน การสร้างความหมายของผู้อ่านจะดำเนินไปอย่างต่อเนื่องเป็นวงจรเป็นขั้นตอนเริ่มจากการที่ผู้อ่านใช้การเดาเรื่อง การจับใจความ การยืนยัน และการแก้ไขตนเองขณะอ่าน หรืออาจกล่าวได้ว่า ผู้อ่านใช้ วิธีการเดาหรือตั้งสมมติฐานในการอ่านและตรวจสอบการเดาของตัวเองว่าถูกต้องหรือไม่ หรือจำเป็นต้องมีการแก้ไขเพื่อให้ได้ความหมายที่ถูกต้อง ผู้ที่มีความสามารถในการอ่าน ที่ดีไม่จำเป็นต้องใช้ตัวชี้แนะทั้งหมด ผู้อ่านที่ดีสามารถเดาความหมายในเรื่องที่อ่านได้ อย่างถูกต้อง แต่ผู้ที่มีความสามารถในการอ่านน้อย

จำเป็นจะต้องกลับไปอ่านเรื่องซ้ำ และตรวจสอบการเดาเรื่องของตนเองอีก กระบวนการอ่านในรูปแบบการอ่านทางภาษาศาสตร์เชิงจิตวิทยาในการอ่านภาษาที่หนึ่ง ของ Goodman ประกอบด้วย 5 ขั้นตอน ดังนี้

1. ขั้นเริ่มจำตัวอักษร (Recognition-initiation) ในขั้นนี้สมองจะต้องจำตัวอักษรใน ภาษาเขียนให้ได้และเริ่มต้นการอ่าน โดยปกติในการอ่านจะต้องมีขั้นนี้หนึ่งครั้ง ถึงแม้ จะมีกิจกรรมอื่นๆ สอดแทรกก็ตาม เช่น การดูภาพประกอบการอ่าน

2. ขั้นการเดา (Prediction) เป็นขั้นที่สมองทำการคาดคะเนและเดา เพื่อหาลำดับความสำคัญของข้อมูลที่ได้รับรู้ในขั้นนี้ผู้อ่านจะเดาเหตุการณ์ในเรื่องก่อนการอ่าน

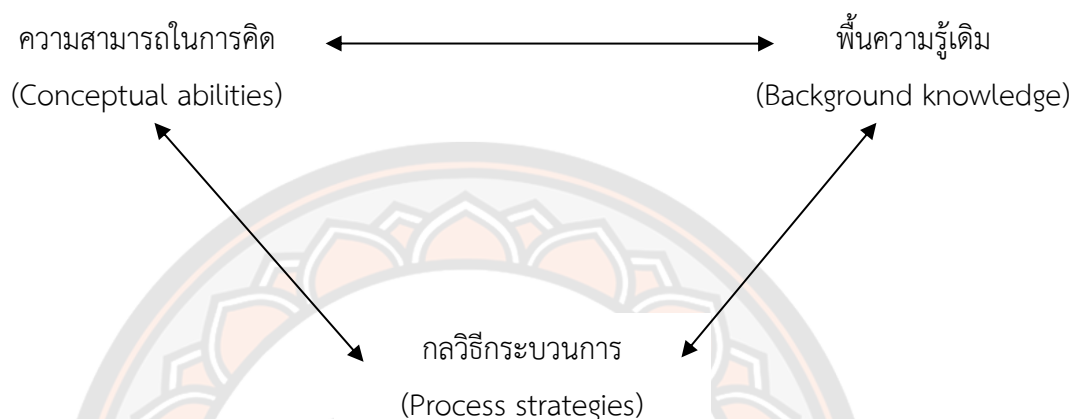
3. ขั้นการยืนยัน (Confirmation) เมื่อสมองเดาเรื่องแล้วจะต้องมีการพิสูจน์การเดาว่าถูกต้องหรือไม่ ในขั้นนี้จะต้องหาข้อมูลมายืนยันหรือปฏิเสธการเดา เพราะฉะนั้น ผู้อ่านจะต้องจับความหมายของเรื่องให้อ่านให้ได้

4. ขั้นการแก้ไข (Correction) ในขั้นนี้เมื่อปรากฏว่าสิ่งที่ตนเดาไว้นั้นไม่ถูกต้องหรือไม่ได้รับการยืนยัน สมองก็ต้องเริ่มต้นกระบวนการคิดและจำใหม่ นั่นคือผู้อ่านจะ ต้องเริ่มต้นกลับไปอ่านทบทวนอีกครั้งเพื่อให้ได้ข้อมูลที่แท้จริง

5. ขั้นสิ้นสุดการอ่าน (Termination) สมองจะยุติการอ่านเมื่อกิจกรรมการอ่านเสร็จสิ้นสมบูรณ์ แต่ผู้อ่านบางคนอาจล้มเลิกการอ่านเพราะไม่สามารถจับใจความของเรื่องให้อ่านได้หรือไม่ก็เป็นเพราะเรื่องนั้นไม่น่าสนใจหรือไม่ตรงตามจุดมุ่งหมายของการอ่าน จากที่กล่าวมาสรุปได้ว่าการอ่านคือการเลือกสรรหรือการทำนายโดยผู้อ่าน เลือกคำชี้แนะจากตัวอักษรในข้อความที่เก็บอยู่ในความจำระยะสั้น แล้วผู้อ่านตัดสินใจ เลือกความหมายของคำที่อ่านแล้วถ่ายโอนข้อมูลไปยังความจำระยะกลาง จนกระทั่งสะสมความหมายได้เพียงพอ แล้วตรวจสอบความสอดคล้องของความหมาย ถ้าเหมาะสมกับสิ่งที่คาดการณ์ไว้ล่วงหน้าในด้านความหมายของคำและโครงสร้างประโยค ก็จะยอมรับความหมายและถ่ายโอนไปยังความจำระยะยาว อย่างไรก็ตามถ้าความหมายไม่ตรงกับที่คาดการณ์ไว้ ผู้อ่านจะกลับไปอ่านข้อความเดิมเพื่อเลือกคำชี้แนะต่อไปและดำเนินไปตามขั้นตอนข้างต้น

กระบวนการอ่านอีกรูปแบบหนึ่งที่แนวความคิดเพิ่มเติมในรูปแบบการอ่านของกลุ่มจิตวิทยาภาษาศาสตร์นั้นคือ กระบวนการอ่านของ Coady (1979: 56 -58, เตือนใจ ตันงามตรง, 2529, หน้า 96-97, อ้างถึงใน อรอนงค์ อุ่นแก้ว, 2019, หน้า 19-20) ข้อคิดเห็นเกี่ยวกับรูปแบบการอ่านตามแนวความคิดของนักจิตวิทยาภาษาศาสตร์เดิมเป็นเพียงปัจจัยหนึ่งของการอ่านเท่านั้นและในขณะที่ผู้อ่านกำลังดำเนินการอ่านสมองของผู้อ่านมิได้ว่างเปล่าหรือไม่รู้เรื่องเกี่ยวกับเนื้อความนั้นเสียทีเดียว ผู้อ่านจะนำความรู้เดิมที่สะสมไว้เอามารวมเข้ากับข้อความใหม่ที่ได้รับจากการอ่าน และนำมาผสมผสานกันกลายเป็นข้อ ความที่เกิดขึ้นในขณะนั้น ดังจะเห็นว่าผู้อ่านสามารถเข้าใจสิ่งที่อ่านได้เป็น

เพราะตัวผู้อ่านเองมากกว่าตัวหนังสือหรือเรื่องี่อ่าน โดยที่ผู้อ่านจะสร้างความหมายของสิ่งที่อ่านโดยอาศัยความรู้ เจตคติและความเชื่อที่มีอยู่ในตัวผู้อ่าน ทักษะในการอ่านขึ้นอยู่กับปฏิสัมพันธ์ระหว่างความรู้ด้านภาษาและความรู้ทั่วไปของผู้อ่าน โคดี้ทำการประยุกต์แนวคิดทฤษฎีของ Goodman ที่ใช้กับภาษาที่ 1 มาเป็นโมเดลที่ใช้กับภาษาที่ 2 โดยอธิบายว่า การอ่านในเชิงจิตวิทยาเกิดขึ้นได้ด้วยปฏิสัมพันธ์ของปัจจัย 3 ประการ ดังแผนภูมิต่อไปนี้



ภาพ 6 รูปแบบการอ่านภาษาอังกฤษเป็นภาษาที่สองของโคดี้ (Coady's 1979)

รูปแบบการอ่านภาษาอังกฤษเป็นภาษาที่สองของโคดี้ (Coady's 1979 model of the ESL reader) เป็นกระบวนการอ่านอีกรูปแบบหนึ่งที่ยึดตามแนวคิดของกลุ่มภาษาศาสตร์เชิงจิตวิทยา รูปแบบการอ่านในลักษณะนี้ได้้อธิบายแนวคิดที่ว่าความเข้าใจในการอ่านจะเกิดขึ้นเนื่องจากการปฏิสัมพันธ์ของปัจจัย 3 ประการคือ

1) ความสามารถในการคิด (Conceptual abilities) หมายถึง ความสามารถทางสติปัญญาทั่ว ๆ ไป

2) พื้นความรู้เดิม (Background knowledge) มีผลต่อความเร็วในการอ่าน ผู้ที่มีความบกพร่องทางภาษาสามารถเข้าใจเรื่องี่อ่านได้ หากผู้นั้นมีความสนใจและมีพื้นความรู้เดิมประกอบกันก็จะทำให้สามารถเข้าใจเรื่องี่อ่าน และส่งผลให้การอ่านมีประสิทธิภาพยิ่งขึ้น

3) กลวิธีกระบวนการ (Process strategies) ประกอบด้วยทักษะย่อยต่างๆ เช่น ความสัมพันธ์ระหว่างหน่วยเสียงและหน่วยคำ ความหมายของคำศัพท์ และบริบท กลวิธีการเรียนรู้รวมทั้งตัวแปรทางด้านจิตวิทยา กลวิธีกระบวนการในการอ่านนี้จะเริ่มดำเนินการอย่างเป็นขั้นตอนคือผู้อ่านต้องมีทักษะในการอ่านเริ่มจากกระบวนการอ่านแบบง่ายระดับรูปธรรมไปตามลำดับจนกระทั่งถึงระดับยากซับซ้อนในขั้นนามธรรมในระยะเริ่มต้นการอ่านผู้อ่านจะเน้นความสำคัญที่ตัวอักษร พยางค์ คำ และประโยค แต่เมื่อผู้อ่านเริ่มพัฒนาความสามารถในการอ่านมากขึ้น กระบวนการนี้ก็จะ

หมดความสำคัญไป และผู้อ่านจะเปลี่ยนไปใช้กลวิธีขั้นนามธรรม โดยอาศัยบริบท พื้นความรู้เดิม ความรู้เกี่ยวกับหน้าที่และโครงสร้างของหน่วยภาษา ในระดับนี้ผู้อ่านจะรู้จักเลือกใช้กลวิธีการอ่านได้อย่างเหมาะสมและถูกต้อง

ถิรวัดน์ ตันทนิส (2550, หน้า 126-132) ได้สรุปแนวคิดเกี่ยวกับการอ่านของ โคตันัน แสดงให้เห็นว่าผู้อ่านจะต้องเลือกใช้และปรับเปลี่ยนกลวิธีในการอ่านให้เหมาะสมกับรูปแบบของเนื้อหา ระดับความยาก ของเนื้อหา และจุดประสงค์ในการอ่านแต่ละครั้ง นอกจากนี้ยังมุ่งเน้นว่า ผู้อ่านจำเป็นต้องใช้ทั้งความสามารถทางความคิดและพื้นความรู้เดิมเป็นตัวช่วยในการอ่าน ซึ่งเกี่ยวข้องกับแนวคิดตามทฤษฎีโครงสร้างความรู้ (Schema Theory)

จากที่กล่าวมา การอ่านเพื่อความเข้าใจตามทฤษฎีการอ่านแนวจิตวิทยา ภาษาศาสตร์เป็นกระบวนการปฏิสัมพันธ์ระหว่างเรื่องทีอ่านกับพื้นความรู้เดิมและความรู้ทั่วไปของผู้อ่าน โดยใช้ความสามารถทางปัญญารวมทั้งการใช้กลวิธีเพื่อช่วยในการทำความเข้าใจ

3.2 ทฤษฎีโครงสร้างความรู้ (Schema Theory)

ทฤษฎีโครงสร้างความรู้เป็นมโนคติหนึ่งที่สำคัญของจิตวิทยาพุทธิปัญญานิยม โดยที่ทฤษฎีนี้ให้ความสนใจทางด้านความรู้ที่มีอยู่ในสมองของคนเราและความสำคัญของความรู้เหล่านี้ที่มีต่อการเรียนรู้สิ่งใหม่ๆ โครงสร้างความรู้จะช่วยอันดับแรกก็คือช่วยให้สิ่งที่อ่านมีความหมายสามารถเชื่อมความสัมพันธ์ระหว่างความรู้ใหม่กับความรู้เดิมพิจารณาความสำคัญของความสัมพันธ์ของข้อมูลในบทอ่านการ สรุปโดยนัยและการจดจำ โครงสร้างความรู้ที่มีความเกี่ยวข้องกับการอ่านมี 3 ชนิด ได้แก่ ความรู้เกี่ยวกับสิ่งต่างๆบนโลก ความรู้เกี่ยวกับลักษณะการเรียบเรียงบทอ่าน และความรู้เกี่ยวกับเนื้อหาวิชา (Adams and Bruce, 1982; Anderson and Pearson, 1984, อ้างถึงใน ฉัตรปวีณ อัมภา, 2555, หน้า 12) โดยเชื่อว่าความรู้จะถูกรวบรวมในโครงสร้างที่จัดไว้อย่างเป็นระเบียบเรียกว่า โครงสร้างความรู้ (Schemata)

ทฤษฎีโครงสร้างความรู้มีความเชื่อว่าความรู้ของคนเราจะได้รับการรวบรวมเป็นหน่วยๆ เรียกว่าโครงสร้างความรู้ (Schemata) ซึ่งโครงสร้างภายในของความรู้ที่มนุษย์มีอยู่นั้นจะมีลักษณะเป็นโนหนดหรือกลุ่มที่มีการเชื่อมโยงกันถูกเรียบเรียงขึ้นเป็นโครงสร้างข้อมูลภายในสมองของมนุษย์ซึ่งรวบรวมความรู้เกี่ยวกับวัตถุ ลำดับเหตุการณ์ รายการกิจกรรมต่างๆ เอาไว้ หน้าที่ของโครงสร้างรู้นี้ก็คือ การนำไปสู่การรับรู้ข้อมูล (Perception) การรับรู้ข้อมูลนั้นไม่สามารถเกิดขึ้นได้หากขาดโครงสร้างความรู้ (Schema) ทั้งนี้ก็เพราะการรับรู้ข้อมูลนั้นเป็นการสร้างความหมายโดยการถ่ายโอนความรู้ใหม่เข้ากับความรู้เดิม ภายในกรอบความรู้เดิมที่มีอยู่และจากการกระตุ้นโดยเหตุการณ์หนึ่ง ๆ ที่ช่วยให้เกิดการเชื่อมโยงความรู้นั้นๆเข้าด้วยกัน การรับรู้เป็นสิ่งสำคัญที่ทำให้เกิดการเรียนรู้ เนื่องจากไม่มีการเรียนรู้ใดที่เกิดขึ้นได้โดยปราศจากการรับรู้ เช่น การไปดูการแข่งขันฟุตบอลเป็นโครงสร้างความรู้เกี่ยวกับเหตุการณ์ เราตีความประสบการณ์ไม่ว่าจะโดยตรงจากการ

เผชิญกับสิ่งต่างๆ หรือโดยอ้อมจากการอ่าน โดยการเปรียบเทียบหรือการจับคู่ประสบการณ์นั้นๆ เข้ากับโครงสร้างความรู้ถ้าหากผู้อ่านมีเครือข่ายโครงสร้างความรู้ที่รวบรวมไว้อย่างหลากหลายก็จะมีความพร้อมสูงในการเลือกดึงมาใช้ถ้าหากเรามีความรู้เกี่ยวกับสิ่งใดสิ่งหนึ่ง ยิ่งมากก็จะง่ายขึ้นในการทำความเข้าใจเกี่ยวกับหัวข้อเรื่องนั้นๆ และก็จะยิ่งเรียนรู้ได้มากขึ้น การใช้ทฤษฎีโครงสร้างความรู้เข้ามาบิบทบาทในส่วนของการตีความสิ่งที่อ่าน การทำความเข้าใจสิ่งที่อ่าน จะช่วยทำให้ผู้อ่านเข้าใจโครงสร้างของข้อความที่มีความซับซ้อน เข้าใจเนื้อเรื่องอย่างมีเหตุการณ์ที่มนุษย์จะเกิดความเข้าใจในการอ่านได้นั้น ผู้อ่านจะต้องมีความรู้ทางภาษาและความรู้ทั่วไปมาช่วยในการทำความเข้าใจเนื้อเรื่อง เพราะข้อเขียนต่าง ๆ เป็นเพียงข้อมูลทางภาษาเท่านั้น ผู้อ่านจะต้องพยายามสร้างความหมายขึ้นมาเองโดยนำประสบการณ์เดิมที่มีอยู่เข้ามาสัมพันธ์กับข้อความในเรื่องที่อ่าน (Rumelhart and Ortony, 1977; อ้างถึงใน ถิรวัฒน์ ต้นพนิส, 2550, หน้า 120; Rumelhart, 1980, 1984, อ้างถึงใน ฉัตรปวีณ อากาศ, 2555, หน้า 12)

สคิม่า (Schema) เป็นกลุ่มของโครงสร้างความรู้ที่ผู้อ่านมีอยู่เดิมแล้วที่เรียงรายกันอยู่เป็นลำดับชั้น จัดเป็นกลุ่มตามคุณสมบัติที่คล้ายคลึงกันและมีความสัมพันธ์ซึ่งกันและกัน กลุ่มโครงสร้างความรู้มีหน้าที่เกี่ยวกับการอ่านคือเป็นสิ่งชี้แนะเรื่องให้กับผู้อ่าน และจัดโครงสร้างเรื่องไว้สำหรับเรื่องใหม่ ในขณะที่ผู้อ่านทำการอ่านเรื่อง ข้อความเหล่านี้จะบรรจุในโครงสร้างความรู้เดิม (Schemata) ที่จัดไว้ โดยสร้างขึ้นมาจากประสบการณ์เดิมของผู้อ่านและเป็นเครื่องช่วยชี้แนะผู้อ่านในขณะที่ต้องการเข้าใจสิ่งต่าง ๆ ทั้งยังช่วยผู้อ่านคาดการณ์ล่วงหน้าเกี่ยวกับเหตุการณ์ในเรื่องด้วย ผู้อ่านจะนำโครงสร้างความรู้เดิมนี้ไปใช้ประกอบการอ่าน ซึ่งมีความสำคัญต่อความเข้าใจในการอ่านมากกว่าโครงสร้างภาษาและกระบวนการทางภาษาที่ใช้ในข้อเขียน (ฉัตรสุดา คงพลอย, 2526 : 1-15) โครงสร้างความรู้ซึ่งประกอบไปด้วยมโนคติเกี่ยวกับสิ่งของและเหตุการณ์ที่มีคุณลักษณะทั่วไปหรือคุณลักษณะเฉพาะบางประการร่วมกันเมื่อคนเรามีประสบการณ์เกี่ยวกับสิ่งใหม่หรือสิ่งที่เราไม่คุ้นเคยจะเกิดการเปรียบเทียบกับสิ่งที่ได้จัดจำพวกไว้แล้วถ้ามีลักษณะผสมกลมกลืนกันก็จะเพิ่มความรู้นี้เข้ากับโครงสร้างที่มีอยู่เดิม (Assimilation) แต่ถ้าข้อมูลใหม่นั้นไม่อยู่ในลักษณะความรู้ที่มีอยู่แล้วก็จะเกิดการปรับความรู้ด้วยการสร้างโครงสร้างความรู้ขึ้นใหม่หรือเปลี่ยนโครงสร้างความรู้เก่าด้วยความรู้ใหม่ (Accommodation) โดยวิธีการนี้โครงสร้างความรู้ที่จัดตั้งขึ้นจะมีลักษณะซับซ้อนมากขึ้น

Mayer (1975: 11, อ้างถึงใน ณฐมน วงศ์ทาทอง, 2560, หน้า 71-75) ได้ให้คำจำกัดความโครงสร้างความรู้ว่าเป็นโครงสร้างของประสบการณ์เดิมในแต่ละด้านที่มีอิทธิพลต่อการรับรู้ในปัจจุบันของคน ๆ นั้น ข้อมูลใหม่ ๆ ที่ได้มาจะผสมกลมกลืนกับความรู้เดิมที่มีอยู่แล้ว ทั้งนี้ข้อมูลบางอย่างอาจสูญเสียลักษณะเฉพาะ แบ่งโครงสร้างเรื่องด้านรูปแบบออกเป็น 4 ชนิด คือ

1. โครงสร้างเรื่องแบบบรรยาย (Collection of Description) โครงสร้างแบบนี้เป็นการบรรยายเรื่องตามเวลาหรือเหตุการณ์ที่เกิดขึ้น
2. โครงสร้างเรื่องแบบเป็นเหตุเป็นผล (Causation) ประกอบด้วยเหตุการณ์ที่เป็นเหตุเป็นผล เรียงลำดับเหตุการณ์ที่เกิดขึ้น
3. โครงสร้างเรื่องแบบปัญหาและการแก้ไข (Problem/Solution) โครงสร้างเรื่องแบบนี้เป็นลักษณะของเหตุและผล ผสมกับปัญหาและการแก้ไขปัญหา
4. โครงสร้างเรื่องแบบเปรียบเทียบ (Comparison) เป็นการจัดเนื้อหาแบบเปรียบเทียบสิ่งที่เหมือนหรือแตกต่างกันโครงสร้างเรื่องมีอิทธิพลต่อการอ่านของผู้ที่อ่านภาษาอังกฤษเป็นภาษาต่างประเทศ แตกต่างกันไปตามเชื้อชาติ ทั้งนี้เนื่องจากโครงสร้างเรื่องในภาษาแม่แตกต่างจากโครงสร้างเรื่องในภาษาอังกฤษทำให้ความรู้เกี่ยวกับโครงสร้างเรื่องในภาษาแม่มาแทรกแซงเวลาที่เขาอ่านภาษาอังกฤษ

Carrell and Eisterhold (1983, p. 557, อ้างถึงใน อรอนงค์ อุ๋นแก้ว, 2562, หน้า 20-21) อธิบายเพิ่มเติมว่า โครงสร้างความรู้เดิมนี้อาจแบ่งออกได้ 2 ชนิด คือ

1. โครงสร้างความรู้เดิมแบบรูปนัย (Formal schemata) หมายถึง การที่ผู้อ่านมีความรู้เกี่ยวกับลักษณะลีลาการเขียนและโครงสร้างของเรื่องมาก่อน เช่น การเขียนเชิงบรรยาย นิทาน วิทยาศาสตร์ หนังสือพิมพ์ ถ้าผู้อ่านมีความรู้สึกไวต่อลักษณะโครงสร้างการเขียนและรู้จักใช้ความรู้เดิมเหล่านี้ให้เป็นประโยชน์ ในขณะที่อ่านจะช่วยอย่างมากทางด้านความเข้าใจและความจำลักษณะของการเขียนเรื่องราวต่าง ๆ จะแตกต่างกันไป และมักจะมีโครงสร้างการเขียนของแต่ละรูปแบบ เช่น นิทาน การบรรยาย การพรรณนา เป็นต้น และในบรรดา โครงสร้างการเขียนทั้งหมด โครงสร้างเกี่ยวกับการเขียนเชิงบรรยายค่อนข้างจะง่ายและเป็นที่ยอมรับกันมากจนอาจจะกล่าวได้ว่า เป็นโครงสร้างแบบสากล เพราะลักษณะการเขียนประเภทนี้มักจะ ไม่ค่อยมีวัฒนธรรมของแต่ละชาติเข้าไปแทรกแซงได้มากเหมือนกับลักษณะการเขียนประเภทอื่น ๆ ผู้อ่านส่วนมากมักจะมีโครงสร้างความรู้เพิ่มเกี่ยวกับการเขียนบรรยาย (Narrative Schemata) มาก่อน การอ่านจึงไม่มีปัญหา

2. โครงสร้างความรู้เชิงเนื้อหา (Content schemata) หมายถึง การที่ผู้อ่านมีความรู้เดิมเกี่ยวกับเนื้อหาของเรื่องในสาขาวิชาใดวิชาหนึ่งมาก่อน เช่น เศรษฐกิจ การแพทย์ ธุรกิจ เป็นต้น จะช่วยทำให้เกิดความเข้าใจดีขึ้นผู้อ่านที่มีโครงสร้างความคิดแขนงนี้จะรับเรื่องได้เร็วกว่าผู้ที่ไม่เคยมีประสบการณ์หรือความรู้ทางเนื้อหาเหล่านี้มาก่อน ผู้อ่านควรมีโครงสร้างความรู้ทั้งสองประเภท เพราะการมีความรู้เกี่ยวกับเรื่องมาก่อน จะช่วยเพิ่มความสามารถในการเดาเรื่องได้ดีขึ้นและถ้ามีความรู้เดิมเกี่ยวกับรูปลักษณะการเขียนดี จะได้เปรียบ เพราะประสิทธิภาพการอ่านจะต้องแปรไปตามลักษณะประโยค อนุเฉท และการจัดระเบียบการเขียนเรื่อง

ดังนั้นครูผู้สอนควรจัดเตรียมปัจจัยดังกล่าวไว้ให้กับนักเรียนที่ยังขาดโครงสร้างความรู้เดิมเหล่านี้และซ่อมเสริมสำหรับผู้เรียนที่ยังมีไม่ครบถ้วน

จากที่กล่าวมา ทฤษฎีโครงสร้างความรู้เป็นมโนทัศน์หนึ่งที่สำคัญของกลุ่มแนวคิดคอนสตรัคติวิสต์เชิงปัญญา ให้ความสนใจทางด้านความรู้ที่มีอยู่ในสมองและความสำคัญของความรู้เหล่านี้ที่มีต่อการเรียนรู้สิ่งใหม่ๆจากการอ่าน โครงสร้างความรู้ช่วยให้สิ่งที่อ่านมีความหมายสามารถเชื่อมความสัมพันธ์ระหว่างความรู้ใหม่กับความรู้เดิมพิจารณาความสำคัญของความสัมพันธ์ของข้อมูลในบทอ่าน กลุ่มของโครงสร้างความรู้ที่ผู้อ่านมีอยู่เดิมแล้วที่เรียงรายกันอยู่เป็นลำดับชั้น (Schema) จัดเป็นกลุ่มตามคุณสมบัติที่คล้ายคลึงกันและมีความสัมพันธ์ซึ่งกันและกัน

3.3 ทฤษฎีอภิปัญญา (Metacognitive Theory)

ทฤษฎีอภิปัญญา (Metacognitive Theory) เป็นแนวคิดหนึ่งของจิตวิทยาการเรียนรู้ หมายถึงการที่บุคคลมีความรู้ความเข้าใจเกี่ยวกับความคิดของตนเอง หรือการคิดเกี่ยวกับการคิดของตน (Thinking about thinking) รวมถึงวิธีการที่ทำให้บรรลุเป้าหมายในการเรียนรู้ เกี่ยวกับการที่บุคคลมีความรู้ความเข้าใจเกี่ยวกับความคิดของตนเองจะต้องประกอบด้วยความรู้ที่บุคคลใช้ในการควบคุมกระบวนการคิด เช่น ทักษะความสามารถ กิจกรรมการเรียนรู้ หรือกลวิธี ส่วนประกอบอีกประการเป็นการจัดระเบียบกระบวนการทางความคิด ทฤษฎีอภิปัญญาจะเกี่ยวข้องกับกลวิธีการสรุปความ การอนุมานการคาดเดาเหตุการณ์และกลวิธีอื่น ๆ ในการพัฒนาความเข้าใจในการอ่าน การที่ผู้อ่านรู้จักกลวิธีในการอ่านดังกล่าวข้างต้นในการพัฒนาความเข้าใจในการอ่านของตนนั้นยังไม่เพียงพอผู้อ่านเป็นต้องรู้จักเลือกกลวิธีในการอ่านได้อย่างถูกต้องและเหมาะสมกับบทอ่านด้วยเช่นกัน (Flavell, 1976, 1979, 1987; Cross & Paris, 1988 : 131; Vacca ,1995, อ้างถึงใน ณฐมน วงศ์ทาทอง, 2560, หน้า 71) การจัดระเบียบของความคิดตามแนวทางทฤษฎีอภิปัญญา เป็นกระบวนการทางความคิดที่กำกับตรวจสอบการเรียนรู้ส่วนการจัดระเบียบเมตาคอกนิชันของตนเอง ประกอบด้วย การกระทำสำคัญ 3 ประการ ได้แก่การวางแผนการกำกับตรวจสอบและการประเมิน ความคิดและการเรียนรู้ (Jacobs and Paris, 1987; Brown, 1987, อ้างถึงในฉัตรปวีณ อัมภา, 2555,หน้า 15-16)

3.3.1 องค์ประกอบสำคัญของอภิปัญญา (Metacognition) สำหรับการอ่านเพื่อความเข้าใจ มี 2 ประการ (Schraw and Dennison , 1994, อ้างถึงในฉัตรปวีณ อัมภา, 2555,หน้า 15-16) ดังนี้

1) ความรู้เกี่ยวกับความเข้าใจประกอบด้วยองค์ประกอบย่อยๆได้แก่ ผู้อ่านมีความรู้เกี่ยวกับกลวิธีการทำความเข้าใจ (Declarative knowledge) ผู้อ่านมีความรู้เกี่ยวกับขั้นตอนของกลวิธีเหล่านี้ (Procedural knowledge) ผู้อ่านรู้เกี่ยวกับว่าจะใช้กลวิธีเหล่านี้เมื่อไรและทำไม (Conditional knowledge)

2) การจัดระเบียบความเข้าใจบทอ่านประกอบด้วย องค์ประกอบย่อยๆ ดังนี้คือการวางแผนการใช้ยุทธศาสตร์การจัดการกับข้อมูลการกำกับ ตรวจสอบความเข้าใจการเลือกใช้ยุทธศาสตร์เมื่อความเข้าใจเกิดติดขัดและการประเมินผล

ฉลอง รัตนพงษ์ (2557, หน้า 38-39) นำเสนอไว้ว่า อภิปัญญา (Metacognition) คือหลักการการควบคุมและประเมินความคิดตนเองความสามารถของบุคคลที่ได้รับการพัฒนาเพื่อควบคุมกำกับกระบวนการทางปัญญาและกระบวนการคิด มีความตระหนักในงาน และสามารถใช้อยุทธวิธีทำงานจนสำเร็จอย่างสมบูรณ์

1) องค์ประกอบของอภิปัญญา แบ่งเป็น 2 องค์ประกอบ คือการตระหนักรู้ (Awareness) และความสามารถในการควบคุมตัวเอง (Self-Regulation)

2) ความรู้ในเชิงอภิปัญญา ประกอบด้วย ความรู้เกี่ยวกับบุคคล (Person) ความรู้เกี่ยวกับงาน (Task) และความรู้เกี่ยวกับกลวิธี (Strategy)

3) กระบวนการในอภิปัญญา ประกอบด้วย การวางแผน (Planning) การตรวจสอบ (Monitoring) และการประเมิน (Evaluating)

ออกซฟอร์ด (Oxford 1990, อ้างถึงใน ฉลอง รัตนพงษ์, 2557, หน้า 36-37) อธิบายว่า กลวิธีที่เกี่ยวกับกระบวนการคิดเกี่ยวข้องโดยตรงกับภาษาเป้าหมายและรวมถึงวิธีต่างๆ เช่น เหตุผลที่อนุมานได้ การพยากรณ์ การบันทึกประเด็นที่สำคัญ การใช้ความรู้ประสบการณ์เดิม การเดาความหมายจากเนื้อหา กลวิธีอภิปัญญา (Metacognitive strategy) เป็นกิจกรรมที่ผู้อ่านควบคุมการอ่านด้วยตนเองหรืออีกนัยหนึ่งก็คือเป็นกลวิธีที่อยู่บนฐาน “การคิดเกี่ยวกับการคิด” (Thinking about thinking) นั่นคือผู้อ่านรู้ว่าใช้กลวิธีวิธีนี้เมื่อไหร่ ใช้อย่างไรและดัดแปลงไปสู่เป้าหมายที่ในการอ่าน ยุทธวิธีอภิปัญญาประกอบด้วยวางแผน การวัดผลและกำหนดทักษะเป็นของตนเองสิ่งเหล่านี้ประกอบด้วยทักษะในการตัดสินใจเรื่องที่ย่าน การวัดผล การคาดเดา การเน้นข้อมูลที่สำคัญ การอ่านผ่านๆ คำใหม่ที่ไม่สำคัญ การตรวจสอบประสิทธิภาพการคาดเดาความหมาย การอ่านข้อความที่ยังไม่เข้าใจใหม่อีกครั้งและการตรวจสอบคุณภาพงานเกี่ยวกับการอ่านทั้งหมด

อภิปัญญา (Metacognition) มีบทบาทสำคัญในการอ่าน ในขณะที่ผู้อ่านกำลังอ่านโดยใช้กลวิธีต่าง ๆ ทั้งทักษะระดับพื้นฐานและในระดับสูง ผู้อ่านใช้ทักษะการอ่านโดยไม่รู้ตัว ผู้อ่านเลือกใช้กลวิธีต่างๆ เพื่อให้ได้ความหมาย ทักษะในขั้นตอนนี้คือการใช้ความรู้ความเข้าใจ (Cognition) แต่เมื่อใดผู้อ่านไม่สามารถทำความเข้าใจกับบทอ่านได้ จึงพยายามหาต้นเหตุของความล้มเหลว หากกลวิธีเพื่อเอาชนะความล้มเหลว ทักษะในการควบคุมการเรียนรู้ของตนเองโดยสามารถบอกได้ว่าใช้กลวิธีใดมาช่วยทำความเข้าใจกับบทอ่าน เรียกว่า เป็นทักษะอภิปัญญา (Metacognitive Skill) ทฤษฎีอภิปัญญาที่นำมาใช้กับความเข้าใจในการอ่านเกี่ยวข้องกับเกณฑ์ 4 เกณฑ์ (Gunning, 1996, อ้างถึงใน ฐมน วรงค์ทาทอง, 2560, หน้า 71) ดังต่อไปนี้

1. การตระหนักรู้ในตนเอง (Knowing Oneself) หมายถึง ผู้อ่านตระหนักรู้ว่าตนรู้อะไรและความสามารถในการอ่านของตน รวมทั้งสามารถใช้ความรู้เดิมในการเตรียมตัวในการอ่าน ผู้อ่านที่มีประสิทธิภาพจะรู้จักเลือก และใช้กลวิธีในการอ่านได้อย่างเหมาะสมกับบทอ่านและกับตนเอง

2. การควบคุมตนเอง (Self-regulation) หมายถึง ผู้เรียนรู้ว่าตนอ่านอะไร อ่านอย่างไรและสามารถใช้ความรู้ของตนมาใช้ในการอ่าน ผู้อ่านมีการสำรวจบทอ่าน มีการจัดเรียงข้อมูลต่าง ๆ ที่ปรากฏในบทอ่าน หากจุดมุ่งหมายในการอ่านหรือเลือกใช้กลวิธีในการอ่านที่มีประสิทธิภาพ

3. การตรวจสอบ (Checking) หมายถึง ผู้เรียนสามารถที่จะประเมินความสามารถในการอ่านของตนเองได้ ผู้อ่านรู้ปัญหาในการอ่าน สาเหตุและวิธีการแก้ไข ผู้อ่านสามารถประเมินการอ่านด้วยตนเองได้ โดยการตั้งคำถามให้กับตนเองว่ามีจุดประสงค์ในการอ่านเพื่ออะไรตรงประเด็นหรือไม่ เป็นต้น

4. การแก้ไข (Repairing) ผู้เรียนสามารถแก้ไขจุดบกพร่องหรือข้อผิดพลาด ในกรณีที่มีปัญหาเกิดขึ้นในการอ่านหรืออ่านไม่เข้าใจ

ฟลาวเวล (Flavell, J. H, 1979, pp. 906-911) เป็นนักคิดในกลุ่ม Metacognition เขาทำการศึกษาอภิปัญญาและการตรวจสอบเชิงพุทธิปัญญา เขาอธิบายว่าอภิปัญญา (Metacognition) และพุทธิปัญญา (Cognition) มีความแตกต่างกัน Metacognition หมายถึง ความรู้ส่วนตัวของแต่ละคนต่อสิ่งที่ได้เรียนรู้หรือสิ่งที่ตนรับรู้ (Knowing) แตกต่างจาก Cognition ซึ่งหมายถึง การรู้คิด หรือปัญญาที่เกิดจากจากเรียนรู้อะไรก็ตามด้วยความเข้าใจ (Comprehension) อย่างไรก็ตาม อภิปัญญาและพุทธิปัญญามีความสัมพันธ์เกี่ยวข้องกัน จำเป็นที่ผู้เรียนต้องมีทั้งสองส่วนในการเรียนรู้ ข้อค้นพบของฟลาวเวลระบุว่า กระบวนการเรียนรู้การคิด (Metacognition) ต้องอาศัยความสามารถทางปัญญา (Cognition) องค์ประกอบของความรู้ของพวกเราจึงมี 3 ประการ ได้แก่

1) บุคคล (Person Variable) หมายถึง ความรู้ความเข้าใจในความสามารถของผู้เรียนว่าตนเองมีคุณสมบัติและมีความสามารถอยู่ในระดับใด

2) ภาระงาน (Task Variable) หมายถึง ลักษณะของงานที่จะต้องเรียนรู้

3) กลวิธี (Strategy Variable) หมายถึง เทคนิคหรือวิธีการที่ผู้เรียนเลือกใช้ในการเรียนรู้ภาระงาน ในกลุ่มแนวความคิดการเรียนรู้โดยการประมวลสารสนเทศอภิปัญญาจะเปรียบได้กับระบบควบคุม (Control Process) แสดงผลว่าทำอะไร (What) ทำอย่างไร (How) และทำเมื่อไร (When)

บทบาทของอภิปัญญาในการอ่าน

บทบาทของอภิปัญญาในการอ่าน (Role of Metacognitive Strategies in Reading) การดัดแปลงความคิดไปสู่ความคิดที่เป็นของตนเองซึ่งเป็นกุญแจสำคัญในการ ไตร่ตรอง เพื่อการอ่านที่คล่องแคล่วและแสดงบทบาทสำคัญในการอ่านเพื่อความเข้าใจ Duffy (2002) กล่าวว่ายุทธวิธีอภิปัญญาเป็นยุทธวิธีหลักที่นำไปสู่การควบคุมการอ่านที่เป็นของตนเอง และไม่เพียงแต่ทำให้ผู้อ่านรู้และเข้าใจแต่เป็นตัวนำและควบคุมการใช้เหล่านั้นด้วย นั่นคือผู้อ่านรู้ว่า กลยุทธ์อะไรที่ใช้ ใช้เมื่อไหร่และทำไมจึงนำกลยุทธ์นี้ไปใช้ และดัดแปลงเพื่อให้เหมาะสมกับ ความมุ่งหมาย ผู้อ่านที่ใช้กลยุทธ์อภิปัญญา วางแผน ควบคุม วัดผลและจัดทักษะของตัวเอง นอกจากนี้ โคเฮน (Cohen, 1998, 1999; Pressey, 2002, อ้างถึงใน ฉลอง) กล่าวว่า ผู้อ่านใช้ความรู้เดิมและมีการปฏิสัมพันธ์กับเนื้อหาที่กำลังอ่าน ในการแก้ปัญหาและเรียนรู้ประสบการณ์ใหม่ๆ อย่างไรก็ตามแล้วแต่ผู้อ่านที่มีประสิทธิภาพจะพัฒนา วางแผนก่อนที่จะมีการอ่านเนื้อหาจริง โดยการจัดการว่าจะทำอะไรในขั้นก่อนการอ่าน ขั้นกำลังอ่าน และขั้นหลังการอ่าน ผู้อ่านทำการอ่านอย่างรู้ตัวและเพิ่มการตระหนักรู้เมื่อเผชิญกับปัญหาที่จำเป็นต้องทำเพื่อแก้ปัญหา เช่น การเลือกกลวิธีการอ่าน (Reading Strategy) ที่ถูกต้อง เหมาะสม และรู้ว่าเมื่อไหร่จะใช้กลยุทธ์นั้นและรู้ว่าใช้อย่างไรในขั้นตอนท้ายสุดผู้อ่านสามารถวัด ประสิทธิภาพ การวางแผน การตรวจสอบ ตัวอย่างเช่น กลยุทธ์การอ่านที่เลือกมาใช้สามารถแก้ปัญหาได้หรือไม่ หากแก้ไม่ได้ก็จำเป็นต้องใช้กลยุทธ์อื่นๆ การฝึกฝนยุทธวิธีอภิปัญญา (Metacognitive Strategy) เพื่อพัฒนาการอ่านเพื่อความเข้าใจนั้นมีประโยชน์อย่างยิ่งทำให้ผู้เรียนมีโอกาสในการวางแผนก่อนการอ่าน ควบคุมขั้นตอน การอ่าน มีการจัดการการอ่านด้วยตนเองและเป็นการวัดผลตนเอง นักเรียนกลายเป็นผู้อ่านอิสระซึ่งนั่นคือเป้าหมายของการอ่าน (Duffy, 2002; Palincsar and Brown 1980, 1984; Salataci and Akyel 2002; Allen, 2003; ฉลอง รัตนพงษ์, 2557, หน้า 36-37)

จากที่กล่าวมา ทฤษฎีอภิปัญญา (Metacognitive Theory) มุ่งหาคำตอบเกี่ยวกับกระบวนการคิดอย่างรู้ตัวของผู้อ่านในทุกขั้นตอนการอ่าน โดยใช้กลยุทธ์อภิปัญญาช่วยในการพัฒนาความสามารถการอ่านของผู้อ่านซึ่งเป็นวิธีที่เกี่ยวกับความสามารถในการนำเอากลวิธีการอ่านไปใช้แก้ปัญหาเมื่อผู้อ่านต้องเผชิญหน้ากับการอ่านเนื้อหาที่ยากต่อการทำความเข้าใจ เพื่อให้การอ่านของตนมีประสิทธิภาพ ดังนั้นทฤษฎีโครงสร้างความรู้ภายใต้แนวคิดคอนสตรัคติวิสต์เชิงปัญญา และทฤษฎีอภิปัญญามุ่งศึกษากระบวนการคิดที่เกี่ยวข้องโดยตรงกับความสามารถการอ่านเพื่อความเข้าใจ ภาษาเป้าหมายซึ่งเป็นกระบวนการปฏิสัมพันธ์ระหว่างผู้อ่านกับข้อความหรือตัวอักษรสร้างความหมายโดยอาศัยพื้นฐานความรู้เดิมเป็นกระบวนการที่เกิดขึ้นในสมอง รวมถึงกลวิธีที่ใช้แก้ปัญหา เป็นความสามารถของบุคคลที่ได้รับการพัฒนาเพื่อควบคุมกำกับกระบวนการทางปัญญามีความตระหนักว่าอ่านอะไรอยู่ และสามารถใช้อยู่ยุทธวิธีในการอ่านจนประสบผลสำเร็จ นอกจากประเภทและขอบเขตของทักษะการอ่านแล้ว นักวิจัยด้านการอ่านได้ตรวจสอบกระบวนการและกลยุทธ์การอ่านโดยให้

นักเรียนมีส่วนร่วมและครูในการคิดทบทวนด้วยวาจาและการสัมภาษณ์ (Afflerbach & Johnston, 1984; Anderson et al., 1991; Cohen, 1987; Cohen & Hosenfeld, 1981; Cohen & Upton, 2007; Faerch & Kasper, 1987; Nevo, 1989, cited in Jang 2009, p.35). โดยเฉพาะการวิจัย กระบวนการอ่านและทักษะในสถานการณ์ทดสอบเมื่อเทียบกับสถานการณ์ที่ไม่ผ่านการทดสอบ แนะนำว่ากลยุทธ์ที่ใช้ขึ้นอยู่กับประเภทของคำถามทดสอบ (Farr, Pritchard, & Smitten, 1990) และ นักเรียนในสถานการณ์ทดสอบใช้กลยุทธ์บ่อยกว่านักเรียนที่อ่านในสถานการณ์ที่ไม่ทำการทดสอบ (Cordon & Day, 1996)

กลวิธีตรวจสอบความเข้าใจ (Comprehension Monitoring Strategies)

ความสามารถทางภาษาของผู้เรียนมีความสัมพันธ์เชิงบวกกับความสามารถในการอ่าน กล่าวคือผู้อ่านที่มีความสามารถทางภาษาไม่ตึงเครียดทักษะการอ่านและขาดการใช้กลยุทธ์การอ่านที่เหมาะสม ความสามารถทางภาษามีความสัมพันธ์เชิงบวกกับความสามารถในการอ่านและการใช้กลยุทธ์การอ่าน กล่าวคือ ผู้เรียนที่มีความสามารถทางภาษาโดยรวมในระดับดีจะนำไปสู่การพัฒนาทักษะการอ่านที่ดีและสามารถใช้เทคนิคต่าง ๆ ที่เหมาะสมเพื่อช่วยให้เข้าใจเนื้อหาที่อ่านง่ายขึ้น เรียกว่ากลยุทธ์การอ่าน (Reading Strategies) ซึ่งมีเจตนาหรือมุ่งเป้าหมายความพยายามในการควบคุมและปรับปรุงความพยายามของผู้อ่านในการถอดรหัสข้อความ เข้าใจคำและโครงสร้าง ความหมาย เข้าใจคำและสร้างความหมายของข้อความ ทักษะการอ่านคือการกระทำโดยอัตโนมัติซึ่งส่งผลให้ในการถอดรหัสและทำความเข้าใจด้วยความรวดเร็ว ประสิทธิภาพและความคล่องแคล่วและมักจะเกิดขึ้นโดยปราศจากการรับรู้ขององค์ประกอบหรือการควบคุมที่เกี่ยวข้อง (Kletzien , 1991, อ้างถึงใน ชนกร สุวรรณพฤษี , 2563, หน้า 207; Sheorey & Mokhatari, 2001; Afflerbach et al., 2008, p. 368)

แอนเดอร์สัน (Anderson,1985: 82 – 83, อ้างถึงใน ณัฐมน วงษ์ทาทอง, 2560, หน้า 47-49) ได้กล่าวไว้ว่า กลวิธีหรือกลยุทธ์ (Strategy) เป็นเทคนิค วิธีการสำคัญที่ช่วยให้ผู้อ่านสามารถอ่านได้อย่างมีประสิทธิภาพ แบ่งออกได้ 3 ประเภท ได้แก่

1) กลวิธีอ่านทางพุทธิปัญญา (Cognitive Reading Strategies) ประกอบด้วย การคาดเดาหรือทำนายเรื่องราวของบทอ่าน และการเดาความหมายคำศัพท์ที่ไม่ทราบด้วยพื้นความรู้เดิม (Background knowledge)

2) กลวิธีอ่านทางอภิปัญญา (Metacognitive Reading Strategies) ประกอบด้วย การตั้งเป้าหมายเพื่อกำหนดขอบเขตในการอ่านและการประเมินความเข้าใจในการอ่าน

3) กลวิธีอ่านแบบชดเชยให้ (Compensating Reading Strategies) ประกอบด้วย การจดโน้ตช่วยบันทึกรายละเอียดหรือประเด็นสำคัญและการเคลื่อนไหวร่างกายเพื่อการจดจำข้อมูลที่อ่าน

โคเฮน (Cohen, 1998, อ้างถึงใน ธรรมนูญ วงษ์ทาทอง, 2560, หน้า 47-49) เสนอกลวิธีที่พัฒนาทักษะการอ่านไว้ ดังนี้

1) เข้าใจจุดประสงค์ของการอ่าน (Understanding the Purpose) ผู้อ่านจำเป็นต้องมีเหตุผลสำหรับการอ่าน ซึ่งช่วยในการตัดสินใจ ประเภทของการอ่านที่ต้องการใช้ เช่น การอ่านคร่าวๆ (Skimming) หรืออ่านเพื่อหาส่วนที่สำคัญ (Scanning) ในเนื้อเรื่อง เป็นต้น

2) ใช้ตัวรวบรวม (Advance Organizer) คือ องค์ประกอบของข้อความที่ช่วยให้ผู้อ่านสร้างความคาดหวังเกี่ยวกับเนื้อหา ประกอบด้วย การหาหัวเรื่องหลัก (Titles) และหัวเรื่องรอง (Subtitles) การหาภาพประกอบ และการหาโครงสร้างของเนื้อเรื่องที่อ่านว่าเป็นรูปแบบใด เช่น จดหมาย นวนิยาย

3) ใช้องค์ความรู้เดิม (Use Your Background Knowledge) ผู้อ่านควรจะนำความรู้เดิมมาใช้ในการอ่าน ผู้อ่านจะพบว่าง่ายและคุ้นเคยกับเรื่องที่อ่าน เพราะผู้อ่านสามารถเดาความหมายของคำศัพท์ที่ไม่รู้ได้

4) ใช้ความรู้ทางภาษาศาสตร์ (Use Your Linguistic Knowledge) ผู้อ่านจำเป็นต้องใช้ความรู้ทางภาษา เช่น การวิเคราะห์คำ การวิเคราะห์โครงสร้างประโยคและตัวเชื่อมความสัมพันธ์ของข้อความ

5) เดาคความหมายจากคำที่ไม่รู้มาก่อน เช่น การใช้บริบท (Context) หรือการมองหาหน้าที่ของคำ หรือส่วนประกอบของคำ เช่น คำอุปสรรค (Prefix) และปัจจัย (Suffix)

6) สร้างการคาดเดา (Make Prediction) ผู้อ่านควรจะเดาว่าผู้เขียนจะเขียนอะไรต่อไป รวมทั้งอาจจะมีการตั้งคำถามเกิดขึ้นในใจ หรืออาจจะสร้างคำถามในขณะที่อ่าน

7) ระบุใจความสำคัญ (Identify Main Idea)

8) อ่านวลีอย่างกว้างๆ (Read in Broad Phrase) วลีในที่นี้หมายถึง กลุ่มของคำที่แสดงความคิด ผู้อ่านควรพยายามทำเป็นนิสัยในการกวาดสายตา เพื่อหาข้อมูล เพราะข้อมูลเหล่านี้จะเป็นประโยชน์มากกว่าการอ่านคำต่อคำ

9) ใช้พจนานุกรมด้วยความระมัดระวัง (Use Dictionary with Care) ผู้อ่านไม่ควรใช้พจนานุกรมบ่อยครั้งเกินไป เพราะจะรบกวนการอ่านหรือทำให้หันเหไปจากเรื่องที่อ่าน

10) เขียนโครงร่าง (Write an Outline) โครงร่าง คือ แผนการหรือการรวบรวมของเนื้อหา ซึ่งแสดงความสัมพันธ์ระหว่างหัวข้อหลักและความคิด ในโครงร่างนั้น ความคิดโดยทั่วไปจะอยู่ทางซ้าย รายละเอียดเฉพาะจะถูกล้อมหน้าไว้ทางขวา

11) ทำบันทึกย่อแบบ ที (Take Note on T-List) เมื่อผู้อ่านกำลังเรียนรู้ข้อมูลใหม่ จำเป็นต้องทำบันทึกย่อเพื่อจะได้จดจำข้อมูลนั้น โดยเขียนใจความหลักหรือมโนทัศน์ไว้ทางซ้ายมือของ T-List แล้วเขียนรายละเอียดของใจความหลักไว้ทางขวา

12) ทำสรุปย่อ (Make Ongoing Summaries) คือ การจำแนกใจความหลักและใจความสนับสนุนจากรายละเอียดที่ไม่สำคัญ

13) เลือกกลวิธีที่เหมาะสม (Choose Appropriate Strategies) ผู้อ่านต้องตระหนักถึงกลวิธีที่หลากหลาย โดยที่ต้องวางแผน ตรวจสอบ ประเมินกลวิธีที่ใช้ และเลือกใช้กลวิธีที่มีประสิทธิภาพ

14) อ่านภาษาอังกฤษให้มากขึ้น (Read English More) อ่านภาษาอังกฤษทุกวัน โดยทำเป็นนิสัย โดยเลือกเนื้อหาที่เหมาะสมและตรงตามความสนใจ

15) เรียนรู้ที่จะประเมินตนเอง (Learn to Assess Yourself)

สำนักวิชาการและมาตรฐานการศึกษา (2551, อ้างถึงใน อนุมาน วงษ์ทาทอง, 2560, หน้า 47) อธิบายว่า กลวิธีที่ช่วยเพิ่มประสิทธิภาพในการอ่านเพื่อความเข้าใจมี 5 ประการ ดังนี้

1. การมองภาพรวม (Previewing) คือ การทบทวนชื่อเรื่อง หัวข้อหลัก และข้อความใต้ภาพ เพื่อที่จะสรุปโครงสร้างและเนื้อหาของเรื่องที่จะอ่าน

2. การคาดการณ์ (Predicting) คือ การใช้ความรู้ในเนื้อหาวิชา เพื่อที่จะคาดการณ์เกี่ยวกับเนื้อหาและคำศัพท์และตรวจสอบความเข้าใจ การใช้ความรู้ของรูปแบบของสารและจุดประสงค์เพื่อที่จะคาดการณ์เกี่ยวกับโครงสร้างปริเฉท การใช้ความรู้เกี่ยวกับผู้เรียน เพื่อที่จะคาดการณ์ลีลาการเขียน คำศัพท์และเนื้อหา

3. การอ่านคร่าว ๆ (Skimming) และการอ่านเร็ว (Scanning) คือ การสำรวจเนื้อหาเพื่อที่จะหาใจความสำคัญ ระบุโครงสร้างของเรื่อง การคาดการณ์ยืนยันและการคาดการณ์คำถาม

4. การเดาจากบริบท (Guessing from Context) คือ การใช้ความรู้เดิมในเรื่องของเนื้อหาวิชาและความคิดในเรื่องที่อ่านเป็นตัวชี้แนะต่อความหมายของคำที่ไม่คุ้นเคยแทนที่จะหาความหมายของคำจากพจนานุกรม

5. การถอดความ (Paraphrasing) คือ การทบทวนและตรวจสอบความเข้าใจในเรื่องที่อ่าน เมื่อถึงตอนสุดท้ายของแต่ละตอน โดยการเรียบเรียงข้อความใหม่

ฟลาวเวล (Flavell, J. H, 1979, pp. 906-911) นักวิชาการกลุ่มแนวคิดอภิปัญญา ได้นำเสนอข้อค้นพบเกี่ยวกับองค์ประกอบของการเรียนรู้ 4 ประการ ได้แก่ 1) คุณลักษณะของผู้เรียน (Learner Characteristics) คือสิ่งต่างๆ ที่เกี่ยวข้องกับผู้เรียน เช่น ความรู้เดิม ทักษะคิด แรงจูงใจ รูปแบบความคิด เป็นต้น 2) กิจกรรมของผู้เรียน (Learner Activities) ในส่วนนี้จะเกี่ยวกับขบวนการใช้สมองของผู้เรียน ในขณะที่เกิดการเรียนรู้ โดยพิจารณาว่า ผู้เรียนจะทำกิจกรรมอะไรบ้าง เพื่อให้เกิดการเรียนรู้ที่ดี เช่น นั่งฟังคำบรรยายด้วยความสนใจ จดโน้ตตาม ชีตเส้นใต้ข้อความที่สำคัญ เป็นต้น 3) ธรรมชาติของสิ่งที่เรียน (Nature of the Learning Material) คือข้อมูลนั้นเป็นข้อมูล ประเภทใด เนื้อหาของข้อมูลนั้น มีลักษณะเป็นรูปธรรมหรือนามธรรม มีการจัดเรียงลำดับเนื้อหาดีมากน้อย

เพียงใด และ 4) วิธีการประเมินความสามารถของผู้เรียน (Nature of the Criterion) คือลักษณะต่างๆ ที่ผู้เรียน แสดงออกมาเมื่อเรียนรู้แล้ว แสดงทักษะ (Skill) หรือคุณลักษณะ (Attribute) ให้ปรากฏ เช่น ทำแบบทดสอบ หรือเขียนเรียงความ

สรุปได้ว่า กลวิธีตรวจสอบความเข้าใจเป็นกระบวนการหรือวิธีการในการประเมินความสามารถการอ่านของผู้เรียนรวมทั้งเทคนิคหรือวิธีการที่ผู้เรียนใช้ในการทำความเข้าใจหรือช่วยให้เข้าใจเนื้อหาที่อ่านง่ายขึ้น

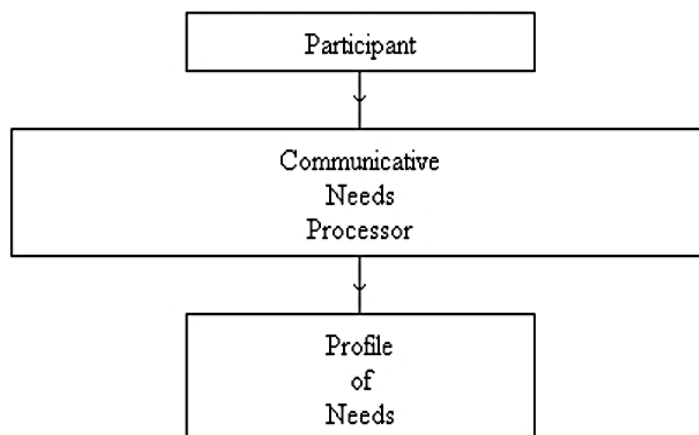
เป็นกลวิธีการอ่านทางพุทธิปัญญา ประกอบด้วย การคาดเดาหรือทำนายเรื่องราวของบทอ่าน และการเดาความหมายคำศัพท์ที่ไม่ทราบด้วยพื้นความรู้เดิม (background knowledge) หรือการคาดเดาจากบริบท (Context clues) การมองภาพรวมโดยทบทวนชื่อเรื่อง หัวข้อหลัก และข้อความใต้ภาพเพื่อสรุปโครงสร้างและเนื้อหาของเรื่องที่อ่าน การตีความและเชื่อมโยงเรื่องที่อ่านกับประสบการณ์หรือเหตุการณ์ทั่วไป โดยลักษณะต่างๆ ที่ผู้เรียนแสดงออกมาเมื่อเกิดกระบวนการเรียนรู้ ได้แก่ ทักษะ (Skill) หรือคุณลักษณะ (Attribute) ให้ปรากฏ โดยมีวิธีการประเมินความสามารถของผู้เรียน เช่น ทำแบบทดสอบ เป็นต้น

4. คุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจ

ความสามารถการอ่านเพื่อความเข้าใจที่ถูกอธิบายด้วยกระบวนการทางพุทธิปัญญาเป็นโมเดลแสดงโครงสร้างความรู้ในสมองคนเราอันประกอบไปด้วยความสามารถทางปัญญาของแต่ละบุคคล พื้นความรู้เดิมหรือประสบการณ์ที่สั่งสมมา รวมทั้งกลวิธีหรือกลยุทธ์ที่ช่วยในการทำความเข้าใจสิ่งที่อ่านเพื่อสร้างความหมายสรรสร้างเชื่อมโยงความสัมพันธ์จนเกิดเป็นความรู้ใหม่ ความสามารถการอ่านเพื่อความเข้าใจประกอบไปด้วยคุณลักษณะหรือทักษะย่อยซึ่งนักจิตวิทยาทางภาษาและนักวิชาการได้นำเสนอไว้ ดังนี้

4.1 Munby's taxonomy ,1978

มันบี้ (Munby, 1978, pp. 154-167) นักวิชาการทางภาษาศาสตร์ ได้เสนอโมเดลการประมวลผลความรู้ทางภาษาอังกฤษถูกเรียกว่าโมเดลมันบี้ (The Munby model) ซึ่งนับเป็นจุดเริ่มต้นในการศึกษาความสามารถทางภาษาของผู้เรียนภาษาอังกฤษ



ภาพ 7 โมเดลมันบี้
(The Munby model, 1978)

หัวใจของโมเดลนี้คือ Communication Needs Processor (CNP) หน่วยประมวลผลกลางความต้องการจำเป็นในการสื่อสาร ซึ่งประกอบด้วยหมวดหมู่ต่างๆ ได้แก่ ข้อมูลเกี่ยวกับผู้เรียนหรือผู้เข้าร่วม จะถูกป้อนเข้าสู่โมเดล CNP หลังจากทีกระบวนการจัดการหมวดหมู่เหล่านี้โดยหน่วยประมวลผลกลางเรียบร้อยแล้ว จะได้ไปไฟล์ของความต้องการจำเป็นที่ใช้ในการสื่อสารซึ่งเป็นคำอธิบายว่าเมื่อจบหลักสูตรผู้เรียนจะนำความรู้ทางภาษาไปใช้ประโยชน์อย่างไรได้บ้าง

คุณลักษณะของความสามารถการอ่านโดยนักวิชาการด้านภาษาศาสตร์ที่ถูกนำเสนอในวรรณกรรมส่วนใหญ่ใช้ออนุกรมวิธานทักษะย่อยด้านการอ่านของมันบี้ (Munby's taxonomy, 1978) อาทิ Li and Suen (2013) H. Chen and J. Chen (2515 pp. 6-7) มี 18 รายการ ดังต่อไปนี้

- ความเข้าใจต้นฉบับ, เนื้อหาของภาษาซึ่งเป็นงานเขียนโดยใช้ตัวอักษรเฉพาะ
- ทำความเข้าใจข้อมูลที่ระบุอย่างชัดเจน
- การคัดแยกจุดที่เกี่ยวข้องจากข้อความ
- สรุปความเกี่ยวกับความหมายคำศัพท์จากข้อความในย่อหน้าหรือบริบทและการใช้คำศัพท์ที่ไม่คุ้นเคย
- การทำความเข้าใจข้อมูลเมื่อไม่ได้ระบุไว้อย่างชัดเจน
- การตีความข้อความโดยเชื่อมโยงไปถึงข้อมูลภายนอกเป็นสิ่งที่ไม่ได้กล่าวไว้ในบทความ
- ความรู้ทางไวยากรณ์
- การทำความเข้าใจความสัมพันธ์ภายในประโยค

- การทำความเข้าใจความสัมพันธ์ระหว่างส่วนต่าง ๆ ของข้อความผ่านอุปกรณ์เชื่อมโยงคำศัพท์หรือไวยากรณ์

- เข้าใจความคิดรวบยอด concept ที่อยู่ในบทอ่าน
 - การระบุประเด็นหลักหรือข้อมูลสำคัญในวาทกรรม
 - แยกแนวคิดหลักจากรายละเอียดสนับสนุนนำเฉพาะประเด็นหลักที่สัมพันธ์เกี่ยวข้องมาเชื่อมต่อ

- นำประเด็นหลักที่รวบรวมได้มาและเชื่อมต่อกันโดยพิจารณาความเกี่ยวข้องกัน จัดลำดับข้อความและทำการสรุป

- การอ่านแบบข้ามคำ (Skim) โดยจะไม่อ่านต่อเนื่องทั้งประโยค หรือทั้งบทความ แต่จะอ่านข้ามเป็นคำ ๆ เป็นประโยค หรือบรรทัด ๆ ไปก็ได้ โดยมีจุดประสงค์เพื่อจับใจความสำคัญของบทความนั้น ๆ เพื่อให้เข้าใจรายละเอียดที่สำคัญเพียงบางสิ่งบางอย่าง

- การอ่านแบบ Scan กวาดสายตาคร่าว ๆ เร็ว ๆ เพื่อค้นหาข้อมูล หรือคำตอบเฉพาะเจาะจงสำหรับคำถาม โดยไม่จำเป็นต้องเข้าใจความหมายของคำศัพท์ทุกคำในประโยค

- การแปลงรหัสเปลี่ยนข้อมูลที่เป็นตัวอักษรให้เป็นการแสดงในรูปแบบแผนภาพ (Transcode) เพื่อเข้าใจความหมาย

4.2 Heaton's taxonomy

Heaton's taxonomy อนุกรมวิธานของฮีตัน (Heaton, 1991) สรุปคุณลักษณะของการอ่านเพื่อความเข้าใจจากอนุกรมวิธาน ดังนี้

- รับรู้ความหมายของคำศัพท์ และกลุ่มคำ
- เข้าใจข้อมูลที่ระบุไว้อย่างชัดเจน
- เข้าใจความสัมพันธ์ภายในประโยค.
- การรับรู้/เข้าใจในทันทีเกี่ยวกับความสัมพันธ์เชิงพื้นที่และการลำดับความคิด
- เข้าใจความสัมพันธ์ระหว่างส่วนของข้อความ
- เข้าใจความหมายทางความคิด
- ระบุแนวคิดหลักและลักษณะเด่นอื่นๆ ในข้อความ
- ระบุหัวข้อเรื่องและค้นหาข้อมูลเฉพาะ
- คาดเดาความหมายของคำศัพท์ที่ไม่รู้จัก
- เข้าใจข้อมูลซึ่งไม่ถูกระบุไว้อย่างชัดเจนโดยอาศัยการคาดคะเนเพื่อสรุปข้อมูล
- คาดคะเนสิ่งที่ไม่ปรากฏชัดเจนแต่บอกเป็นนัยโดยเชื่อมโยงไปถึงข้อมูลภายนอก
- ประเมินสิ่งที่อ่านและตัดสินใจว่าสิ่งที่ผู้เขียนนำเสนอ มีเหตุผลน่าเชื่อถือ

หรือไม่ เพียงใด

- นำวิธีการที่ยืดหยุ่นมาใช้

4.3 Jang's Attributes (Jang, 2005, 2009)

การศึกษาของจาง (Jang, 2005, 2009) มีวัตถุประสงค์เพื่อประเมินความถูกต้องอย่างยิ่งของ CDA กับการทดสอบความเข้าใจในการอ่านเป็นภาษาที่สอง L2 ที่ไม่ใช่การวินิจฉัยที่มีอยู่ และเพื่อให้ข้อมูลเกี่ยวกับความท้าทายและเงื่อนไขที่ความเชื่อถือได้ของโปรไฟล์ทักษะของ Fusion Model คุณลักษณะของโปรไฟล์ทักษะ L2 ที่ได้ และความสามารถในการวินิจฉัยของรายการทดสอบ LanguEdge™

คุณลักษณะที่ 1 ความหมายของคำในบริบท (Word meanings in context) เป็นความสามารถในการระบุคำหรือวลีที่มีความหมายใกล้เคียงที่สุดกับความหมายที่ผู้เขียนตั้งใจไว้ โดยสามารถค้นหาและวิเคราะห์ข้อความที่อยู่โดยรอบและใช้ประโยชน์จากคำบอกใบ้ที่ปรากฏในข้อความ

คุณลักษณะที่ 2 ความรู้คำศัพท์ (Vocabulary knowledge) เป็นความสามารถในการระบุความหมายของคำโดยใช้ความรู้เกี่ยวกับคำศัพท์ แทนการอาศัยเงื่อนไขจากข้อความเป็นหลัก

คุณลักษณะที่ 3 ความสัมพันธ์ทางไวยากรณ์ของคำหรือวลี (Grammatical Relationships of words/phrases) เป็นความสามารถในการเข้าใจความสัมพันธ์ทางไวยากรณ์ของคำหรือวลีเป็นความสัมพันธ์ที่กำหนดโครงสร้างระหว่างคำในวลีและอนุประโยค คำศัพท์ทั่วไปที่ใช้เพื่ออ้างถึงความสัมพันธ์ทางไวยากรณ์เฉพาะ ได้แก่ Subject (ประธานของประโยค), Direct object (กรรมตรง), Indirect object (กรรมรอง), Ergative (คำกริยาที่เป็นได้ทั้งกรรมกริยาและกรรมกริยา), Absolutive (เป็นได้ทั้งกรรมกริยาหรือกรรมกริยาที่มีความสมบูรณ์ในตัวของมันเอง เช่น kill, genitive (การแสดงความเป็นเจ้าของของคำนามคือการเติม 's และ การใช้ of) ข้ามประโยคในย่อหน้าเดียวกัน รวมทั้งภาระงานในการระบุคำหรือวลีที่คำสรรพนามอ้างถึงอย่างเฉพาะเจาะจง หรือการกำหนดตำแหน่งที่ประโยคใหม่สามารถแทรกได้โดยไม่มีปัญหาทางหลักเหตุและผลหรือไวยากรณ์ Transitive Verbs คือ กริยาที่มีกรรมในประโยค อาจจะมีกรรม 1 ส่วน หรือมากกว่า 1 ส่วน ก็ได้ Intransitive Verbs คือ กริยาที่ไม่มีกรรมในประโยค กริยาที่บอกว่าการกระทำนั้น เกิดขึ้นหรือส่งผลต่อใครหรือสิ่งใด กริยาที่บอกว่าการกระทำนั้น ไม่ได้มีผลต่อใครหรือสิ่งใดเป็นการกระทำที่มีความสมบูรณ์ในตัวเอง

คุณลักษณะที่ 4 การทำความเข้าใจข้อมูลที่ระบุไว้อย่างชัดเจนในข้อความ (Comprehending information explicitly stated in the text) เป็นความสามารถในการเข้าใจข้อมูลที่ผู้อ่านต้องการซึ่งถูกระบุไว้อย่างชัดเจนในข้อความ โดยการค้นหาข้ามประโยคภายในย่อหน้า

เดียวกัน และสามารถจับคู่คำหรือวลีซึ่งมีความหมายเสริมกันหรือสอดคล้องกันไปในทางเดียวกัน ที่ถูกระบุสถานะไว้อย่างชัดเจนในข้อความ

คุณลักษณะที่ 5 การทำความเข้าใจข้อมูลที่ไม่ได้ระบุไว้อย่างชัดเจนในข้อความ (Comprehending information NOT explicitly stated in the text) เป็นความสามารถในการเข้าใจข้อมูลที่ไม่ได้ระบุไว้อย่างชัดเจนในข้อความและสามารถระบุตัวเลือกแบบถอดความที่เป็นการใช้คำต่างๆในการแสดงความหมายของข้อความที่อ่านเพื่อให้เกิดความชัดเจนโดยไม่เปลี่ยนแปลงความหมายของข้อความที่ผู้เขียนตั้งใจสื่อสาร ได้อย่างถูกต้องแม่นยำ

คุณลักษณะที่ 6 การสรุปความข้อโต้แย้งและความมุ่งหมายของผู้เขียน (Inferencing arguments and author's purpose) เป็นความสามารถในคาดคะเนตามหลักเหตุผลเกี่ยวกับข้อโต้แย้งซึ่งมีจุดประสงค์เพื่อชักชวนผู้อื่นว่าความคิดนั้นๆถูกหรือผิด โดยเป็นข้อมูลที่ผู้เขียนไม่ได้ระบุไว้อย่างชัดเจนในข้อความ

คุณลักษณะที่ 7 การวิเคราะห์ข้อเท็จจริง (Comprehending negation) เป็นความสามารถในการแยกข้อมูลที่เป็นจริง ออกจากข้อมูลที่เป็นข้อคิดเห็น ทั้งที่ปรากฏในย่อหน้าเดียวกันและข้ามไปย่อหน้าอื่นอีก และระบุคำสำคัญหรือข้อความสำคัญที่จะนำไปสู่คำตอบ

คุณลักษณะที่ 8 สรุปแนวคิดหลัก (Summarizing major ideas) เป็นความสามารถในการระบุแนวคิดหลัก โดยการแยกแยะข้อความที่เป็นแนวคิดหลักออกจากข้อความทั่วไป ทั้งที่ปรากฏในย่อหน้าเดียวกันและข้ามไปย่อหน้าอื่นอีกด้วย

คุณลักษณะที่ 9 การวางแผนผังการเปรียบเทียบความต่างและแนวคิดหลัก (Mapping major contrasts and ideas) หมายถึง การเชื่อมโยงความสัมพันธ์ของข้อมูลโดยตระหนักถึงการเรียงลำดับข้อความซึ่งเป็นข้อความที่มีความสัมพันธ์กันในแต่ละส่วน และสามารถระบุข้อโต้แย้งที่สำคัญจากเรื่อง

4.4 Gao's attributes

เกา (Gao, 2006) หลังจากการวิเคราะห์วรรณกรรมและในส่วนของโครงสร้างแบบทดสอบความเข้าใจในการอ่าน MELAB มีคุณลักษณะ/ทักษะย่อย 10 รายการ ดังต่อไปนี้

- รับรู้และกำหนดความหมายของคำหรือวลีเฉพาะโดยใช้เบาะแสบริบทหรือความรู้ด้านเสียงหรืออักขรวิธีหรือคำศัพท์ (PC1)
- เข้าใจโครงสร้างประโยคและความหมายของประโยคโดยใช้ความรู้ทางวากยสัมพันธ์ (PC2)
- ทำความเข้าใจความสัมพันธ์ระหว่างประโยคและการจัดระเบียบของข้อความ โดยใช้ความรู้ความเข้าใจเกี่ยวกับองค์กรเชิงวาทศิลป์ (PC3)
- คาดการณ์ถึงสิ่งที่จะเกิดขึ้นต่อไปโดยอาศัยข้อความที่อ่าน (PC4)

- วิเคราะห์วัตถุประสงค์ของผู้เขียน (PC5)
- ระบุใจความสำคัญ ค้นหาหัวเรื่อง หรือแนวคิดหลัก และอ่านเนื้อหาคร่าวๆ (PC6)
- ค้นหาข้อมูลเฉพาะที่ร้องขอในคำถาม สแกนข้อความเพื่อดูรายละเอียด ซึ่งรวมถึง (ก) จับคู่คำศัพท์สำคัญในคำถามกับรายการคำศัพท์สำคัญในส่วนที่เกี่ยวข้องของข้อความ และ (ข) ระบุหรือกำหนดคำพ้องความหมายหรือการแปลความหมายตามตัวอักษรของคำ วลีหรือประโยค ในส่วนที่เกี่ยวข้องของข้อความ (PC7)
- ทำการตีความและข้อสรุปตามข้อมูลที่ระบุโดยปริยายในข้อความ (PC8)
- สังเคราะห์ข้อมูลที่น่าเสนอในประโยคหรือส่วนต่าง ๆ ของข้อความ (PC9)
- ประเมินทางเลือกอื่นเพื่อหาคำตอบโดยเชื่อมโยงความคิดหลักเข้าด้วยกัน (PC10)

4.5 DELTA's Attributes

เดลต้า (DELTA, 2012) หรือชื่อเต็มว่า The Diagnostic English Language Tracking Assessment (DELTA) เป็นระบบทดสอบและติดตามการวินิจฉัยภาษาอังกฤษทางเว็บที่ใช้ อยู่สามแห่งมหาวิทยาลัยในฮ่องกง DELTA อิงตามกรอบการทำงานของ Bachman and Palmer (1996) ของความสามารถทางภาษาและประกอบด้วยกรอบการฟังแบบปรนัยส่วนบุคคล คำศัพท์ การอ่าน และไวยากรณ์ ในการประเมินความสามารถการอ่าน เดลต้าระบุคุณลักษณะการ อ่านเพื่อความเข้าใจโดยอ้างอิงจากอนุกรมวิธานของนักภาษาศาสตร์ (Munby, 1978; Weir, 1983; Alderson & Lukmani, 1989; Jang, 2009) โครงสร้างที่รวมไว้ในองค์ประกอบการอ่านของ DELTA ประกอบด้วยทักษะย่อยโดยอิงตามทักษะแปดประการที่ระบุในการศึกษาโดย Alderson และ Lukmani (1989) เป็นหลัก และอ้างอิงถึง Jang (2009) ด้วยการประยุกต์และเรียบเรียงใหม่โดยทีม DELTA เพื่อให้นักเรียนเข้าถึงความตั้งใจได้มากที่สุด ทักษะย่อยในการอ่านของ DELTA แสดงไว้ ด้านล่าง:

- คุณลักษณะที่ 1 การระบุข้อมูลเฉพาะ
- คุณลักษณะที่ 2 การตีความคำหรือวลีตามที่คุณเขียนใช้
- คุณลักษณะที่ 3 การทำความเข้าใจแนวคิดหลักและแนวคิดสนับสนุน
- คุณลักษณะที่ 4 การทำความเข้าใจข้อมูลและการอนุมาน
- คุณลักษณะที่ 5 การตีความเหตุผลของผู้เขียน
- คุณลักษณะที่ 6 การอนุมานทัศนคติหรือความตั้งใจของผู้เขียน
- คุณลักษณะที่ 7 การทำความเข้าใจความสัมพันธ์ทางไวยากรณ์ของคำหรือวลีใน

ข้อความ

คุณลักษณะที่ 8 การระบุประเภทข้อความ

4.6 Li's Attributes

หลี่ และคณะ (Li et. al., 2015, p.8) ได้ทำการศึกษาเกี่ยวกับการตรวจสอบปัญหาในการเลือก CDM ที่เหมาะสมสำหรับการวิเคราะห์การวินิจฉัยทางปัญญาของการทดสอบการอ่านเพื่อทำความเข้าใจการประเมินความสามารถภาษาอังกฤษของมิชิแกน หรือ MELAB (Li, H., & Suen, H., 2013, cited in Li et. al., 2015, p.8) ประการแรก พวกเขาเสนอกรอบการอ่านเบื้องต้นตามรูปแบบของกระบวนการทางความคิดที่เป็นพื้นฐานของการทดสอบการอ่าน MELAB ที่พัฒนาโดยเกาและโรเจอร์ส (2010) และโมเดลการอ่าน TOEFL iBT ที่สร้างโดย Jang (2009) รายการทักษะการอ่านของ MELAB และ TOEFL แบ่งออกเป็น 5 หมวดหมู่ ได้แก่ คำศัพท์ ไวยากรณ์ การแยกเนื้อหาที่ชัดเจน ข้อมูล การเชื่อมต่อและการสังเคราะห์และการอนุมาน ประการที่สอง ด้วยการอ้างอิงถึงกรอบกระบวนการเริ่มต้นนี้ พวกเขาสร้าง Q-matrix เริ่มต้นตามหลักฐานจากการคิดออกเสียงของนักเรียนรายงานและพิจารณาการให้คะแนนของผู้เชี่ยวชาญ ในขั้นสุดท้ายเมตริกซ์คิว (Q-matrix) เริ่มต้นนี้ได้รับการตรวจสอบเพิ่มเติมโดยใช้หลักฐานเชิงประจักษ์เบื้องต้น เนื่องจากพบข้อสอบเพียง 2 ข้อที่ต้องการทักษะของการสรุปความ (inferencing) พวกเขาตัดสินใจที่จะรวม "การเชื่อมต่อและการสังเคราะห์" และ "การสรุปความ" เข้าด้วยกันได้ทักษะที่กว้างขึ้นโดยเรียกทักษะดังกล่าวว่า "การทำความเข้าใจข้อมูลที่เป็นนัยไม่ปรากฏชัดเจนในเรื่อง" จึงมีคุณลักษณะที่เกี่ยวข้องดังนี้

คุณลักษณะที่ 1 คำศัพท์ (Vocabulary) หมายถึง ทักษะการใช้คำศัพท์การรู้จำและกำหนดความหมายเฉพาะของคำหรือกลุ่มคำ ความรู้ด้านเสียง อักษรวิธี และ/หรือการเดาความหมายคำศัพท์จากบริบท

คุณลักษณะที่ 2 วากยสัมพันธ์ (Syntax) หมายถึง การทำความเข้าใจโครงสร้างประโยคและความหมายของประโยคโดยใช้ความรู้เกี่ยวกับไวยากรณ์ ไวยากรณ์ เครื่องหมายวรรคตอน ส่วนของคำพูด

คุณลักษณะที่ 3 การสกัดข้อมูลที่ปรากฏชัดเจนในเรื่อง (Extracting Explicit Information) หมายถึง การจับคู่คำศัพท์และ/หรือข้อมูลวากยสัมพันธ์ในคำถามกับข้อมูลในส่วนที่เกี่ยวข้องของข้อความและระบุหรือกำหนดคำพ้องความหมายหรือการถอดความความหมายตามตัวอักษรของคำ วลี หรือประโยคในส่วนที่เกี่ยวข้องของข้อความ

คุณลักษณะที่ 4 การทำความเข้าใจข้อมูลที่เป็นนัยไม่ปรากฏชัดเจนในเรื่อง (Understanding Implicit Information) หมายถึงการสรุปความที่ผู้เขียนไม่ได้บอกอย่างชัดเจนแต่ผู้อ่านต้องคาดเดาโดยอาศัยข้อความต่างๆ ใช้การเชื่อมต่อและการสังเคราะห์การบูรณาการ เชื่อมโยงหรือสรุปข้อมูลที่นำเสนอในประโยคหรือส่วนต่าง ๆ ของข้อความเพื่อสร้างความหมาย

4.7 Thip's Attributes

ทิพย์ ขำอยู่ (2517, หน้า 6-10) ทำการศึกษาวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษและทำการสรุปคุณลักษณะทักษะการอ่านมี 6 คุณลักษณะดังนี้

คุณลักษณะที่ 1 ความรู้เรื่องชนิดของคำ (Part of speech) หมายถึง พื้นความรู้ของผู้เรียนเกี่ยวกับประเภทหรือชนิดของคำในภาษาอังกฤษซึ่งมีหน้าที่และตำแหน่งที่แตกต่างกันไป ในประโยคที่จำเป็นต้องใช้เป็นทักษะพื้นฐานสำหรับการอ่านภาษาอังกฤษ ประกอบด้วย คำนาม (Noun) คำสรรพนาม (Pronoun) คำกริยา (Verb) คำคุณศัพท์ (Adjective) คำกริยาวิเศษณ์ (Adverb) คำบุพบท (Preposition) คำสันธาน (Conjunction) คำอุทาน (Interjection)

คุณลักษณะที่ 2 ความรู้เรื่องการใช้คำในบริบทซึ่งตีความหมายของคำศัพท์ยาก หมายถึง พื้นความรู้เกี่ยวกับตัวบอกในประโยคหรือข้อความต่างๆที่ช่วยบอกให้ผู้อ่านทราบถึงความหมายของคำศัพท์ยากหรือคำศัพท์ที่ไม่คุ้นเคย ที่จำเป็นต้องใช้เป็นทักษะพื้นฐานสำหรับการอ่านภาษาอังกฤษ ประกอบด้วย การนิยามความหมาย (Context clue given by definition) การกล่าวซ้ำ (Restatement or Explanation clue) การเปรียบเทียบความเหมือนหรือความขัดแย้ง (Comparison and Contrast clue) การยกตัวอย่าง (Illustration or Example clue)

คุณลักษณะที่ 3 การเรียนรู้ความหมายจากโครงสร้างของคำศัพท์ (Word formation) หมายถึง การวิเคราะห์คำโดยการเรียนรู้รากคำ ส่วนของคำที่เติมหน้าคำและส่วนของคำที่เติมท้ายคำช่วยให้หาความหมายของคำศัพท์ที่ไม่คุ้นเคยเพื่อใช้เป็นทักษะพื้นฐานสำหรับการอ่านภาษาอังกฤษ

คุณลักษณะที่ 4 การอ่านเพื่อหาหัวเรื่องและใจความสำคัญของเรื่อง (Identify topic and main idea) หมายถึง การใช้ทักษะการอ่านเพื่อหาหัวเรื่องและใจความสำคัญของเรื่อง เพื่อเป็นทักษะพื้นฐานสำหรับการอ่านภาษาอังกฤษซึ่งประกอบด้วย 1) การระบุคำที่บอกประเภทของคำต่างๆในกลุ่ม เริ่มจากการหาคำที่บอกประเภทหรือชนิดของคำในกลุ่มเดียวกันได้ 2) การระบุหัวเรื่องของข้อความ และ 3) การอ่านเพื่อหาใจความสำคัญ

คุณลักษณะที่ 5 การอ่านเพื่อหารายละเอียดของเรื่อง (Reading for details) หมายถึง การอ่านเพื่อค้นหาสิ่งที่ต้องการหลังจากที่ได้อ่านผ่านและอ่านแบบพินิจพิเคราะห์แล้วทำให้ทราบว่าสิ่งที่ต้องการทราบอยู่ที่ใดของเรื่อง ลักษณะคำถามเกี่ยวกับรายละเอียดที่สำคัญของเรื่องซึ่งเป็นทักษะการอ่านภาษาอังกฤษที่ผู้เรียนจำเป็นต้องมี

คุณลักษณะที่ 6 การอ่านเพื่อสรุปความ (Making inferences) หมายถึง การตรวจสอบเท็จจริงทั้งหมดของเรื่องที่อ่าน เพื่อพิจารณาหรือตัดสินใจที่อ่านทั้งหมดนั้นหมายความว่าถึงสิ่งใดซึ่งเป็นการพัฒนาทักษะการสรุปความหรือตีความในทักษะการอ่านภาษาอังกฤษที่จำเป็นสำหรับผู้เรียน

4.8 Javidanmehr's Attributes

จาวีแดนเมาะห์และซาหรับ (Javidanmehr & Sarab, 2019, p.6-8) ได้ทำการกำหนดคุณลักษณะการอ่านเพื่อความเข้าใจจากหลายแหล่ง ได้แก่ การทบทวนวรรณกรรมเพื่อกำหนดคุณลักษณะหรือทักษะย่อยของการอ่านจับใจความ โครงสร้างแบบทดสอบ เทคนิคการคิดออกเสียงของผู้สอบ และคณะผู้เชี่ยวชาญการตัดสินความตรงของเมตริกซ์คิว (Q-matrix) รวมถึงคุณลักษณะทั้ง 5 ได้รับการพัฒนาและตรวจสอบคุณลักษณะความสามารถการอ่านเพื่อความเข้าใจ 5 คุณลักษณะ ดังนี้

คุณลักษณะที่ 1 มีความรู้ด้านคำศัพท์ (Vocabulary Knowledge) หมายถึงความสามารถในการเข้าใจความหมายคำศัพท์ด้วยการเดาจากข้อความบริบทรอบๆ โดยใช้กระบวนการคิดหรืออาศัยความรู้หรือประสบการณ์เดิม

คุณลักษณะที่ 2 ความรู้ด้านไวยากรณ์ (Syntactic Knowledge) หมายถึงความสามารถในการใช้ความรู้ด้านไวยากรณ์ซึ่งเป็นลำดับคำที่เหมาะสมในวลีหรือประโยคที่ช่วยผู้อ่านในการทำความเข้าใจเนื้อหาที่ใช้แทนคำนามในบทอ่าน ช่วงเวลา หรือลำดับเหตุการณ์

คุณลักษณะที่ 3 การระบุข้อมูลที่ต้องการ (Explicit Information) หมายถึงความสามารถในการอ่านอย่างละเอียดเพื่อดึงข้อมูลเฉพาะหรือข้อมูลที่ผู้อ่านต้องการออกจากบทอ่าน บอกระยะสำคัญของเรื่อง (Passage) หรือสาระสำคัญของย่อหน้าหนึ่งๆ (Paragraph)

คุณลักษณะที่ 4 การสรุปความจากเรื่องที่อ่าน (Making Inferences) เป็นการใช้ความคิดเพื่อหาข้อสรุปของเรื่องที่อ่านตามหลักเหตุผล โดยคาดคะเนจากชื่อเรื่อง หรือจากความมุ่งหมายของผู้เขียน หรือเชื่อมโยงความรู้หรือประสบการณ์เดิมของผู้อ่านกับการเรียนรู้ใหม่

คุณลักษณะที่ 5 เชื่อมโยงและสังเคราะห์ข้อมูล (Connecting and Synthesizing) เป็นความสามารถในการรวบรวมและเชื่อมต่อส่วนของข้อความซึ่งแสดงความสัมพันธ์ระหว่างความคิดหรือใจความสำคัญโดยเป็นความเชื่อมโยงเกี่ยวข้อกันระหว่างประโยคใจความหลักและประโยคสนับสนุนในบทอ่าน และหาข้อสรุปโดยนำเสนอด้วยประโยคใหม่ซึ่งใช้ภาษาที่ง่ายขึ้น

4.9 Tonekaboni's Attributes

โตนีกาโบนี และราวัน (Tonekaboni & Ravand, 2021, p. 60) ได้ทำการระบุคุณลักษณะของความเข้าใจในการอ่านแบบทดสอบถูกกำหนดจากการรายงานด้วยวาจาย้อนหลังภายหลังจากทำแบบทดสอบและการตัดสินของผู้เชี่ยวชาญ 6 คุณลักษณะดังนี้

คุณลักษณะที่ 1 ความรู้คำศัพท์ (Vocabulary Knowledge) คือ ความสามารถในการสร้างความหมายของคำโดยใช้ความรู้เดิม เช่น ชื่อเรื่อง ความรู้ทั่วไป เจื่อนำทางภาษาและบริบท

คุณลักษณะที่ 2 การสรุปความ (Inferencing) เป็นความสามารถในการสมมติหรือสร้างการเชื่อมต่อ (ระหว่างข้อความและองค์ประกอบตามบริบท) ซึ่งไม่ได้ระบุไว้อย่างชัดเจน เช่นกันโดยอัตโนมัติโดยอาศัยความรู้เดิมมากรอกข้อมูลที่ไม่ได้อยู่ในข้อความหรือไม่ใช่โดยอัตโนมัติ/เชิงกลยุทธ์โดยข้อมูลที่ยึดในประโยคของข้อความ ความสามารถในการสมมติหรือทำให้การเชื่อมต่อ (ระหว่างองค์ประกอบที่เป็นข้อความและบริบท)

คุณลักษณะที่ 3 สร้างรูปแบบสถานการณ์ (Building Situation Model) เป็นความสามารถในการประมวลผลและบูรณาการแนวคิดจากข้อความและแนวคิดที่เกี่ยวข้องและความรู้ทั่วไปเพื่อสร้างตัวแทนความเข้าใจในจิตของเนื้อหาข้อความ

คุณลักษณะที่ 4 ความรู้เกี่ยวกับโครงสร้างข้อความ (Text Structure Knowledge) เป็นความสามารถในการแยกแยะว่าข้อความมีโครงสร้างและจัดระเบียบอย่างไรโดยใช้ความรู้เดิม ความรู้เกี่ยวกับวากยสัมพันธ์ และความรู้เกี่ยวกับความสัมพันธ์ระหว่างและระหว่างประโยคและย่อหน้าเพื่ออนุมานเกี่ยวกับข้อความ เพื่อจัดระเบียบเนื้อหาและสร้างภาพแทนใจของเนื้อหาข้อความ

คุณลักษณะที่ 5 ระบุส่วนที่เป็นประเด็นหลัก (Identifying and constructing the main idea) เป็นความสามารถในการระบุและสร้างส่วนสำคัญของย่อหน้าแนวคิดหลักหรือชื่อเรื่องของข้อความ

คุณลักษณะที่ 6 สังเคราะห์ข้อมูลจากบทอ่าน (Synthesizing information from text) เป็นความสามารถในการรวบรวมแนวคิดหลักและเรียบเรียงประโยคขึ้นใหม่ ซึ่งยังคงความหมายเดิมโดยมีเจตนาเพื่อเชื่อมโยงแนวคิดหลักในแต่ละย่อหน้าหรือแต่ละส่วนของงานเขียนเข้าไว้ด้วยกัน

จากการนำเสนอคุณลักษณะหรือทักษะย่อยของความสามารถการอ่านเพื่อความเข้าใจโดยนักวิชาการทางภาษาศาสตร์ที่กล่าวมาในข้างต้น ผู้วิจัยได้ทำการสังเคราะห์คุณลักษณะความสามารถการอ่านเพื่อความเข้าใจแสดงในตาราง 2 มีรายละเอียดดังนี้

ตาราง 1 ผลการสังเคราะห์คุณลักษณะความสามารถอ่านเพื่อความเข้าใจภาษาอังกฤษ

Munby's taxonomy	Heaton's taxonomy	Gao, 2006	Jang, 2009	Li et al., 2015	Thip, 2017	Javidanmehr 2019	Tonekaboni, 2021	ผลการสังเคราะห์
- ความเข้าใจ	- การรับรู้ความหมายของคำศัพท์	- รู้และกำหนด	- รู้	- คำศัพท์	- ความรู้เรื่องชนิดของคำ	- ความรู้คำศัพท์	- ความรู้คำศัพท์	- มีความรู้
ความหมายของตัวอักษรในแง่เขียนงาน	และกลุ่มคำ	ความหมายของคำหรือวลีเฉพาะโดย	ความหมายของคำจากบริบท	หมายถึงทักษะการใช้คำศัพท์การรับรู้จุดจำและกำหนด	หมายถึงพื้นฐานของผู้เรียนเกี่ยวกับชนิดของคำในภาษาอังกฤษซึ่งมีหน้าที่และตำแหน่งที่แตกต่างกันไป	a. การจดจำ	- เป็นความ	คำศัพท์
ประพันธ์	- คาดคะเนความหมายของคำที่ไม่รู้จัก	ใช้เบาะแสบริบทหรือความรู้ด้านเสียงหรืออักขรวิธีหรือคำศัพท์	บริบท	กำหนดความหมายคำศัพท์หรือกลุ่มคำ ความรู้ด้านเสียงอักขรวิธี และ/หรือการเดา	คำในภาษา	คำศัพท์โดยใช้	สามารถในการสร้าง	พิจารณาจากความเข้าใจเกี่ยวกับ
- คาดเดาคำศัพท์จากข้อความในย่อหน้าหรือบริบทและการใช้คำศัพท์ที่ไม่คุ้นเคย			- มีความรู้คำศัพท์	กลุ่มคำ ความรู้ด้านเสียงอักขรวิธี และ/หรือการเดา	อังกฤษที่ปรากฏในบริบทประจำวัน	เพื่อนำตามบริบท	ความหมายของคำ	ความหมาย
				คำศัพท์หรือกลุ่มคำ ความรู้ด้านเสียงอักขรวิธี และ/หรือการเดา	คำในภาษา	b. การรับรู้	โดยใช้ความรู้เดิม	คำศัพท์ และกลุ่มคำโดยอาศัยพื้นฐานเดิม หรือเข้าใจความหมายของคำศัพท์ยาก
				ความหมายคำศัพท์จากบริบท	บริบทประจำวัน	ความหมายของคำศัพท์	ความหมายของคำ	คำศัพท์ยาก หรือไม่คุ้นเคย โดยอาศัยการคาดเดา
				บริบทประจำวัน	บริบทประจำวัน	ความหมายของคำศัพท์	ความหมายของคำ	ความหมาย
				คำศัพท์จากบริบท	บริบทประจำวัน	ความหมายของคำศัพท์	ความหมายของคำ	คำศัพท์จากบริบทหรือ
				ความหมายของคำศัพท์จากบริบท	บริบทประจำวัน	ความหมายของคำศัพท์	ความหมายของคำ	บริบทหรือ

Munby's taxonomy 1978	Heaton's taxonomy 1991	Gao, 2006	Jang, 2009	Li et al., 2015	Thip, 2017	Javidanmehr 2019	Tonekaboni, 2021	ผลการตั้งเครื่องหมาย
(transcode) เพื่อเข้าใจ ความหมาย								คำบ่งชี้ที่พบในข้อความ ประกอบด้วยการนิยาม ความหมาย การกล่าวซ้ำ การเปรียบเทียบ ความเหมือน หรือความขัดแย้ง การยกตัวอย่าง
- ทำความเข้าใจ ข้อมูลที่ระบุ อย่างชัดเจน -การคัดแยกจุด ที่เกี่ยวข้องจาก ข้อความ - การอ่านแบบ ข้ามคำ	- เข้าใจข้อมูลที่ ระบุไว้อย่าง ชัดเจน - การรับรู้/เข้าใจ ในทันทีเกี่ยวกับ ความสัมพันธ์เชิง พื้นที่และการ ลำดับความคิด	- ค้นหาข้อมูลที่ เฉพาะที่ร้องขอใน คำถามและสแกน ข้อความเพื่อดู รายละเอียด	- เข้าใจข้อมูลที่ ที่ถูกระบุไว้ อย่างชัดเจน ในเรื่องที่อ่าน	- การสกัดข้อมูลที่ ปรากฏชัดเจนใน เรื่องการสกัด ข้อมูลที่ชัดเจน หมายถึงการจับคู่ คำศัพท์และ/หรือ ข้อมูล ความสัมพันธ์	- การอ่านเพื่อหา รายละเอียดของเรื่อง หมายถึง การอ่านเพื่อ ค้นหาสิ่งที่ต้องการ หลังจากที่ได้อ่านผ่าน และอ่านแบบพินิจ พิเคราะห์แล้วทำให้ ทราบว่าสิ่งที่ต้องการ	- ความเข้าใจใน เนื้อหาซึ่งเป็น ข้อมูลที่ถูก นำเสนออย่างชัด แจ้งหมายถึงการ ประมวลผลข้อมูล ที่ชัดเจนและการ จับคู่ข้อมูลในส่วน		
					ยากหรือคำศัพท์ที่ไม่ คุ้นเคย ที่จำเป็นต่อผู้ใช้ เป็นที่กะพินฐาน สำหรับการอ่าน ภาษาอังกฤษ ประกอบด้วยการนิยาม ความหมาย การกล่าวซ้ำ การเปรียบเทียบ ความเหมือนหรือความ ขัดแย้ง การยกตัวอย่าง			คำบ่งชี้ที่พบใน ข้อความ ประกอบด้วยการ นิยาม การ ความหมาย การ กล่าวซ้ำ การ เปรียบเทียบ ความเหมือน หรือความ ขัดแย้ง การ ยกตัวอย่าง

Munby's taxonomy 1978	Heaton's taxonomy 1991	Gao, 2006	Jang, 2009	Li et al., 2015	Thip, 2017	Javidanmehr 2019	Tonekaboni, 2021	ผลการสังเคราะห์
(skim)เพื่อให้เข้าใจ รายละเอียดที่สำคัญเพียงบางสิ่งบางอย่าง - การอ่านแบบ Scan กวาดสายตา ตคร่าว ๆ เร็ว ๆ เพื่อค้นหา ข้อมูล หรือ คำตอบเฉพาะ เจาะจงสำหรับ คำถาม โดยไม่ จำ เป็นต้องเข้าใจ ความหมายของ คำศัพท์ทุกคำ ในประโยค	คำถามกับข้อมูล ในส่วนที่เกี่ยวข้อง ของข้อความและ ระบุหรือกำหนด คำพ้องความหมาย หรือการถอดความ ความหมายตาม ตัวอักษรของคำ วลี หรือประโยค ในส่วนที่เกี่ยวข้อง ของข้อความ	คำถามกับข้อมูล ที่เป็นส่วนที่เกี่ยวข้อง ของข้อความและ ระบุหรือกำหนด คำพ้องความหมาย หรือการถอดความ ความหมายตาม ตัวอักษรของคำ วลี หรือประโยค ในส่วนที่เกี่ยวข้อง ของข้อความ	คำถามกับข้อมูล ที่เป็นส่วนที่เกี่ยวข้อง ของข้อความและ ระบุหรือกำหนด คำพ้องความหมาย หรือการถอดความ ความหมายตาม ตัวอักษรของคำ วลี หรือประโยค ในส่วนที่เกี่ยวข้อง ของข้อความ	คำถามกับข้อมูล ที่เป็นส่วนที่เกี่ยวข้อง ของข้อความและ ระบุหรือกำหนด คำพ้องความหมาย หรือการถอดความ ความหมายตาม ตัวอักษรของคำ วลี หรือประโยค ในส่วนที่เกี่ยวข้อง ของข้อความ	ทราบอยู่ที่ใดของเรื่อง ลักษณะคำถามเกี่ยวกับ รายละเอียดที่สำคัญของ เรื่องซึ่งเป็นทักษะการ อ่านภาษาอังกฤษที่ ผู้เรียนจำเป็นต้องมี	ที่เป็นคำถามกับ ส่วนที่เป็นคำตอบ		รายละเอียด หรือข้อมูล เฉพาะที่ ต้องการ หลังจากที่ได้ อ่านผ่านและ อ่านแบบพินิจ พิเคราะห์แล้ว ทำให้ทราบว่า ข้อมูลอยู่ที่ใด ของเรื่อง เช่น จำนวน ราคา วันเวลา สถานที่ ซึ่งข้อมูล ดังกล่าวจะ ปรากฏอย่าง ชัดเจนผู้อ่าน สามารถเข้าใจ ได้โดยมีต้อง

Munby's taxonomy 1978	Heaton's taxonomy 1991	Gao, 2006	Jang, 2009	Li et al., 2015	Thip, 2017	Javidanmehr 2019	Tonekaboni, 2021	ผลการสังเคราะห์
- การทำความเข้าใจ เข้าใจข้อมูลเมื่อ ไม่ได้ระบุไว้ อย่างชัดเจน - การตีความข้อ ความโดย เชื่อมโยง ไปถึงข้อมูลภายนอก ไม่ได้กล่าวไว้ใน บทอ่าน	- เข้าใจข้อมูลซึ่ง ไม่ถูกระบุไว้อย่าง ชัดเจนโดยอาศัย การคาดคะเนเพื่อ สรุปข้อมูล - คาดคะเนสิ่งที่ไม่ ปรากฏชัดเจนแต่ บอกเป็นนัยโดย เชื่อมโยงไปถึง ข้อมูลภายนอก ในข้อความ	- คาดการณ์ถึงสิ่ง ที่จะเกิดขึ้นต่อไป โดยอาศัยข้อความ ที่อ่าน - วิเคราะห์ วัตถุประสงค์ของ ผู้เขียน - ทำการตีความ และข้อสรุปตาม ข้อมูลโดยปริยาย ในข้อความ	- การทำความเข้าใจ เข้าใจข้อมูลที่ไม่ ได้ถูกระบุ ไว้อย่าง ชัดเจน - การสรุป ความข้อ โต้แย้งและ ความมุ่งหมาย ของผู้เขียน เป็นการ คาดคะเนตาม หลักฐานที่ เกี่ยวข้องกับ โต้แย้งซึ่งมี จุดประสงค์ เพื่อชักชวน ผู้อ่านว่าความ	- การสรุปความ โดยอนุมานหรือ คาดคะเนด้วย เหตุผลเกี่ยวกับ เจตคติจุดมุ่งหมาย ของผู้เขียนที่ไม่ แสดงชัดเจนใน ข้อความ ข้อความ ความมุ่งหมาย ของผู้เขียน เป็นการ คาดคะเนตาม หลักฐานที่ เกี่ยวข้องกับ โต้แย้งซึ่งมี จุดประสงค์ เพื่อชักชวน ผู้อ่านว่าความ	- การอ่านเพื่อสรุปความ หมายถึง การตรวจ ข้อเท็จจริงทั้งหมดของ เรื่องที่อ่าน เพื่อพิจารณา หรือตัดสินว่าสิ่งที่อ่าน ทั้งหมดนั้นหมายความว่า ถึงสิ่งใดซึ่งเป็น พัฒนาทักษะการสรุป ความหรือตีความใน ทักษะการอ่าน ภาษาอังกฤษที่จำเป็น สำหรับผู้เรียน	- การสรุปความ จากเรื่องก่อน หมายถึงการใช้ ความคิดเพื่อหา ข้อสรุปของเรื่องที่ อ่านตามหลัก เหตุผล โดยคาด คะเนจากข้อเรื่อง หรือจากความมุ่ง หมายของผู้เขียน หรือเชื่อมโยง ความรู้หรือ ประสบการณ์เดิม ของผู้ผู้อ่านกับการ เรียนรู้ใหม่	- การสรุปความ เป็น ความ สามารถในการสมมติ หรือสร้างการ เชื่อมต่อ ระหว่าง ข้อความและองค์ ประ กอบตามบริบท ซึ่ง ไม่ได้ระบุไว้อย่าง ชัดเจน จุดมุ่งหมายหรือ ทัศนคติของ ผู้เขียนหรือข้อ โต้แย้ง โดย อาศัยเบาะแส ร่องรอยใน เนื้อความซึ่ง	ตีความ

Munby's taxonomy 1978	Heaton's taxonomy 1991	Gao, 2006	Jang, 2009	Li et al., 2015	Thip, 2017	Javidanmehr 2019	Tonekaboni, 2021	ผลการ สังเคราะห์
			คิดนั้นถูกต้อง หรือผิด โดย เป็นข้อมูลที่ได้ ผู้เขียนไม่ได้ ระบุไว้อย่าง ชัดเจนใน ข้อความ					เป็นคำหรือ ประโยคที่แนะ หรือทำการ ตรวจสอบที่จริง ทั้ง หมดของเรื่องที่ อ่าน เพื่อ พิจารณาหรือ ตัดสินว่าสิ่งที่ อ่านทั้งหมดนั้น หมายความว่าถึง สิ่งใดและ สามารถ เชื่อมโยงไปถึง ข้อมูลภายนอก ซึ่งไม่ได้กล่าวไว้ ในตอนอ่าน
ความรู้ทาง ไวยากรณ์	-เข้าใจ ความสัมพันธ์	- เข้าใจโครงสร้าง ประโยคและ	- ความ สัมพันธ์เชิง	- การทำความเข้าใจ ใจความ	- การเรียนรู้ความหมาย จากโครงสร้างของ	- ความรู้ทาง ไวยากรณ์หมายถึง	- ความรู้เกี่ยวกับ โครงสร้างของ	-มีความรู้ทาง ไวยากรณ์

Munby's taxonomy	Heaton's taxonomy	Gao, 2006	Jang, 2009	Li et al., 2015	Thip, 2017	Javidanmehr 2019	Tonekaboni, 2021	ผลการสังเคราะห์
1978	1991							หรือกลุ่มคำ ประเภทต่าง ๆ ในประโยคได้ อย่างถูกต้อง ตามโครงสร้าง ทางไวยากรณ์
-แยกแวกคิด หลักจาก รายละเอียด สนับสนุน นำ เฉพาะประเด็น หลักที่สัมพันธ์ เกี่ยวข้องกันมา เชื่อมต่อกัน -นำประเด็น หลักที่รวบรวม ได้มาและ เชื่อมต่อกันโดย พิจารณาความ	-เข้าใจความ หมายของความคิด รวมยอดที่อยู่ใน บทอ่าน -ระบุแนวคิดหลัก และลักษณะเด่น อื่นๆ ในข้อความ	-เข้าใจความ สัมพันธ์ระหว่าง ประโยคกับการจัด ระเบียบข้อความ -ประเมินทางเลือก อื่นเพื่อหาคำตอบ โดยเชื่อมโยง ความคิดหลักเข้า ด้วยกัน -สังเคราะห์ข้อมูล ที่นำเสนอใน ประโยคหรือส่วน ต่าง ๆ ของ	-การวิเคราะห์ เพื่อแยกข้อ เท็จจริง ออกมา -การวาง แผนผังการ เปรียบเทียบ ความต่างและ แนวคิดหลัก หมายถึง การ เชื่อมโยง ความ สัมพันธ์ของ	-ระบุประเภทของ ข้อความ	-การจัดระเบียบ และสังเคราะห์ ข้อมูลในส่วน ต่างๆของเรื่องที่ อ่าน ทำความ เข้าใจข้อมูลหลัก และข้อมูล สนับสนุน ความ สัมพันธ์ระหว่าง ความคิดใน ข้อความโครงสร้าง วาทศิลป์ เข้าใจ การเชื่อมโยงส่วน	-การสังเคราะห์ ข้อมูลจากบท อ่าน พิจารณาได้จาก ความ สามารถในการ แยกประโยค แนวคิดหลัก ออกจาก ประโยคอื่นซึ่ง เป็นรายละเอียด เสริม นำ รวบรวมแนวคิดหลัก ประเด็นหลักที่	-สร้างรูปแบบ สถานการณ์เป็น ความสนใจในการ ประมวลผลและ เชื่อมโยงแนวคิดจาก ข้อความที่อ่านกับ ความรู้ทั่วไปเพื่อ สร้างความหมายขึ้น ในใจ -สังเคราะห์ข้อมูล จากบทอ่านเป็น ความสนใจในการ รวบรวมแนวคิดหลัก	

Munby's taxonomy 1978	Heaton's taxonomy 1991	Gao, 2006	Jang, 2009	Li et al., 2015	Thip, 2017	Javidanmehr 2019	Tonekaboni, 2021	ผลการสังเคราะห์
เกี่ยวข้องกัน	ข้อความ	ข้อความ	ข้อมูลโดยอิสระหนึ่งถึง			หนึ่งของข้อความ	และเรียงเรียง	รวบรวมได้มา
จัดลำดับ			การเรียงลำดับ			กับส่วนอื่นของข้อความเดียวกัน	ประโยคขึ้นใหม่ซึ่งยังคงความหมายเดิม	เชื่อมต่อกันโดยพิจารณาความเกี่ยวข้อง
ข้อความและทำการสรุป			ข้อความซึ่งเป็นข้อความที่มีความสัมพันธ์กันในแต่ละส่วน และสามารถระบุข้อได้แย้งที่สำคัญจากเรื่อง			อย่างสอดคล้องกัน	โดยมีเจตนาเพื่อเชื่อมโยงแนวคิดหลักในแต่ละย่อหน้าหรือแต่ละส่วนของงานเขียนเข้าไว้ด้วยกัน	จัดลำดับข้อความ และทำการสรุปหรือเรียบเรียงประโยคขึ้นใหม่ที่ยังคงความหมายเดิม

5. กรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรป

CEFR ย่อมาจาก Common European Framework of Reference for Languages หมายถึง กรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรป (สถาบันภาษาอังกฤษ สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2558) สร้างขึ้นจากความมุ่งมั่นของสภายุโรปตั้งแต่ ปี พ.ศ. 2507 ที่ต้องการจะพัฒนาการศึกษาด้านภาษาเพื่อเพิ่มความเข้าใจด้านการสื่อสารระหว่างประเทศ และเพื่อส่งเสริมการเรียนรู้ที่สามารถนำไปใช้ได้จริง กรอบอ้างอิง CEFR ได้รับการพัฒนามาต่อเนื่องโดยสภายุโรปด้านการศึกษา (Council of European, 2001) ในปี พ.ศ. 2533 กรอบอ้างอิง CEFR ได้ถูกจัดตั้งขึ้นโดยสภายุโรป เพื่อให้เป็นส่วนหนึ่งในการสนับสนุนการทำงานร่วมกันระหว่างผู้สอนภาษาทุกประเทศในสหภาพยุโรป และเป็นเกณฑ์มาตรฐานวัดระดับความสามารถทางด้านภาษา และในปี พ.ศ. 2539 – 2541 สภายุโรปได้นำกรอบอ้างอิง CEFR ฉบับร่างมาทดลองใช้ และได้ถูกเปิดตัวอย่างเป็นทางการ ในปี พ.ศ. 2544 ที่สามารถกล่าวได้ว่าเป็นจุดเปลี่ยนที่สำคัญเนื่องจากสามารถนำกรอบอ้างอิง CEFR ไปปรับใช้ได้หลายบริบท และครอบคลุมภาษามากถึง 40 ภาษา (Council of European, 2018)

สถาบันภาษาอังกฤษ สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (2558, หน้า 9-11) อธิบายว่า กรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรป (CEFR) คือ มาตรฐานการประเมินความสามารถทางภาษาที่สหภาพยุโรปจัดทำขึ้นโดยมีวัตถุประสงค์เพื่อใช้เป็นแนวทางในการจัดการเรียนการสอน และการประเมินภาษาที่สองหรือภาษาต่างประเทศ ในปี ค.ศ. 2002 สภาแห่งสหภาพยุโรปได้กำหนดให้ใช้กรอบอ้างอิงทางภาษาของสหภาพยุโรป (CEFR) ในการตรวจสอบความสามารถทางภาษา ปัจจุบันกรอบอ้างอิงนี้ (CEFR) ได้รับการยอมรับอย่างกว้างขวางว่าเป็นมาตรฐานในการจัดลำดับความสามารถทางภาษาของแต่ละบุคคล

การนำกรอบอ้างอิง The Common European Framework of Reference for Languages (CEFR) ซึ่งเป็นกรอบอ้างอิงความสามารถทางภาษาอังกฤษที่สหภาพยุโรปจัดทำขึ้นมาใช้เป็นแนวทางในการจัดการเรียนรู้อังกฤษในระดับการศึกษาขั้นพื้นฐานครอบคลุม 5 ด้านต่อไปนี้ (กระทรวงศึกษาธิการ, 2557, หน้า 5)

1. ด้านการกำหนดเป้าหมายการจัดการเรียนรู้ และการพัฒนา โดยจำแนกผู้เรียนออกเป็น 3 กลุ่มหลัก และแบ่งออกเป็น 6 ระดับความสามารถตามกรอบอ้างอิง CEFR
2. ด้านการพัฒนาหลักสูตรการเรียนการสอนภาษาอังกฤษกล่าวคือมีการนำกรอบอ้างอิง CEFR มาใช้ในการพัฒนาหลักสูตร และกำหนดเป้าหมายของหลักสูตร

3. ด้านการจัดการเรียนการสอน กล่าวคือ การนำกรอบอ้างอิง CEFR มาเป็นแนวทางในการจัดกระบวนการเรียนรู้ และการเรียนการสอนภาษาอังกฤษเพื่อให้ผู้เรียนมีองค์ความรู้ และทักษะทางภาษาอังกฤษตามที่ระบุไว้ในกรอบอ้างอิง CEFR ในแต่ละระดับ

4. ด้านการทดสอบ และการวัดผล โดยการนำคำอธิบายความสามารถทางภาษาใน ระดับต่าง ๆ ที่กำหนดตามกรอบอ้างอิง CEFR มาใช้เป็นแนวทางในการทำแบบทดสอบ และวัดระดับ ความสามารถทางภาษาของผู้เรียน เพื่อให้ผู้เรียนมีความสามารถตรงตามเป้าหมาย และเกณฑ์ที่ กำหนด ด้วยเหตุนี้ ในการประเมินหรือตรวจสอบผลการจัดการศึกษา ตลอดจนการทดสอบหรือวัดผล ความสามารถทางภาษาของผู้เรียนต้องมีการใช้แบบทดสอบมาตรฐานที่เทียบเคียงผลคะแนนกับระดับ ความสามารถทางภาษาตามกรอบอ้างอิง CEFR ได้

5.1 การวัดความรู้ภาษาต่างประเทศตามกรอบอ้างอิงทางภาษาของสหภาพยุโรป

วัตถุประสงค์สำคัญประการหนึ่งของกรอบอ้างอิงทางภาษาของสหภาพยุโรป (CEFR) คือการใช้วัดระดับความสามารถทางด้านภาษาและให้คำอธิบายถึงความสามารถของผู้เรียนใน แต่ละระดับ (Martyniuk, 2011) กรอบอ้างอิง CEFR ไม่ได้เชื่อมโยงหรือยึดตามแบบทดสอบใด แบบทดสอบหนึ่งเป็นพิเศษ ผู้สอนสามารถปรับเปลี่ยนแบบทดสอบให้เหมาะสมเพื่อใช้ในการประเมิน และออกแบบบทเรียน เพื่อค้นหาช่องโหว่ของความรู้ทางภาษาสำหรับผู้เรียน (EF SET, 2018) เพื่อ เป็นกรอบและแนวทางในการศึกษาขอบเขตของทักษะความรู้และความสามารถที่ผู้เรียนควรเรียนรู้ และได้รับเพื่อให้สามารถใช้ภาษาเพื่อการสื่อสารในทุกรูปแบบได้อย่างมีประสิทธิภาพ และเป็น แนวทางในการกำหนดเป้าหมายของการศึกษาและผลลัพธ์ของการศึกษา เพื่อให้ทราบถึง ความก้าวหน้าในการเรียนการสอนมากกว่ามุ่งเน้นไปที่คะแนนการสอบ (Cambridge English, 2016; Council of Europe, 2001, 2018)

การนำกรอบอ้างอิง CEFR มาใช้ในการทดสอบ และการวัดผล โดยใช้ แบบทดสอบหรือแบบวัดที่สามารถเทียบเคียงผลได้กับกรอบอ้างอิง CEFR นั้นส่งผลให้ครูผู้สอน สามารถนำข้อมูลระดับความสามารถของผู้เรียนหรือผู้เข้ารับการทดสอบ มาดำเนินการปรับปรุงการ เรียนการสอนที่ตอบสนองความสามารถที่ต่างกันของแต่ละบุคคล หรือเพื่อหาแนวทางพัฒนาผู้เรียนที่ พบว่ามีข้อบกพร่องในคุณลักษณะใดคุณลักษณะหนึ่ง ให้มีความสามารถตามเป้าหมายหรือเกณฑ์ที่ กำหนดกระทรวงศึกษาธิการได้กำหนดเป้าหมายการพัฒนาระดับความสามารถทางภาษาของผู้เรียน ในระดับการศึกษาขั้นพื้นฐานไว้ตามตาราง 3 ดังนี้

ตาราง 3 แสดงเป้าหมายการพัฒนาในระดับความสามารถทางภาษาของผู้เรียน

ระดับนักเรียน	ระดับความสามารถทางภาษา	ระดับความสามารถทางภาษาตามกรอบ CEFR
ผู้สำเร็จการศึกษาระดับประถมศึกษา (ป.6)	ผู้ใช้ภาษาขั้นพื้นฐาน	A1
ผู้สำเร็จการศึกษาภาคบังคับ (ม.3)	ผู้ใช้ภาษาขั้นพื้นฐาน	A2
ผู้สำเร็จการศึกษาขั้นพื้นฐาน (ม.6/ ปวช.)	ผู้ใช้ภาษาขั้นอิสระ	B1

จากตาราง 3 กรอบมาตรฐาน CEFR กำหนดให้ผู้สำเร็จการศึกษาภาคบังคับ (ม.3) เป็นผู้ใช้ภาษาขั้นพื้นฐานมีระดับความสามารถทางภาษาเท่ากับ A2 ดังนั้น ในการประเมินหรือตรวจสอบผลการจัดการศึกษา หรือผลการพัฒนาผู้เรียนในแต่ละระดับข้างต้น หน่วยงานที่รับผิดชอบควรได้มีการทดสอบหรือวัดผล โดยใช้แบบทดสอบมาตรฐานที่เทียบเคียงผลคะแนนกับระดับความสามารถทางภาษาตามกรอบอ้างอิง CEFR เพื่อตรวจสอบว่าผู้เรียนมีผลสัมฤทธิ์ทางการเรียนภาษาผ่านเกณฑ์ระดับความสามารถที่กำหนดหรือไม่ (สถาบันภาษาอังกฤษ สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2558, หน้า 8)

5.2 ระดับความสามารถทางภาษาตามกรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรป

การระบุถึงทักษะที่ผู้เรียนแต่ละระดับต้องรู้อย่างชัดเจน โดยอ้างอิงตามกรอบมาตรฐานสากล (Common European Framework of Reference for Languages : CEFR) สหภาพยุโรปได้กำหนดคำอธิบายเกี่ยวกับกรอบมาตรฐานความสามารถทางภาษาตามต้นฉบับเป็นภาษาอังกฤษ (Cambridge, 2022) มีรายละเอียด ดังนี้

“The Common European Framework of Reference for Languages (CEFR) is an international standard for describing language ability. It describes language ability on a six-point scale, from A1 for beginners, up to C2 for those who have mastered a language. This makes it easy for anyone involved in language teaching and testing, such as teachers or learners, to see the level of different qualifications. It also means that employers and educational institutions can easily compare our qualifications to other exams in their country.”

CEF A1 (ALTE Breakthrough)	CAN understand the general meaning of a simplified textbook or article, reading very slowly.
CEF A2 (ALTE 1)	CAN understand simple textbooks, articles etc., understanding most key points. CAN follow simple argumentation. CAN understand visuals if they are predictable and if understanding depends on simple keys that can be looked up in a dictionary.
CEF B1 (ALTE 2)	CAN understand simple visuals on familiar topics, e.g. a weather map, if not much explanatory text is involved. CAN understand, given sufficient time, most information of a factual nature that (s)he is likely to come across during the course of study.
CEF B2 (ALTE 3)	CAN understand most visuals that (s)he is likely to come across, but MAY sometimes have difficulty with textual commentary.
CEF C1 (ALTE 4)	CAN handle most textbooks, articles etc., within own area of expertise. CAN scan texts for relevant information, and grasp main topic of text.
CEF C2 (ALTE 5)	CAN read quickly enough to cope with the demands of an academic course. CAN understand abstract concepts and argumentation. CAN scan texts for relevant information, and grasp main topic of text, reading almost as quickly as a native speaker.

เพื่อให้สะดวกในการทำความเข้าใจสถาบันภาษาอังกฤษ กระทรวงศึกษาธิการ (สถาบันภาษาอังกฤษ, 2558 , หน้า 10) ได้เรียบเรียงเป็นภาษาไทย มีรายละเอียดในตาราง 4

ตาราง 4 แสดงกรอบเป้าหมายระดับความสามารถในการใช้ภาษา CEFR

ระดับ	คำอธิบาย
A1	ผู้เรียนสามารถใช้และเข้าใจประโยคง่าย ๆ ในชีวิตประจำวัน สามารถแนะนำตัวเองและผู้อื่น สามารถตั้งคำถามเกี่ยวกับบุคคลอื่น เช่น เขาอยู่ที่ไหน รู้จักใครบ้าง มีอะไรบ้าง และตอบคำถามเหล่านี้ได้ ทั้งยังสามารถเข้าใจบทสนทนาเมื่อคู่สนทนาพูดช้าและชัดเจน
A2	ผู้เรียนสามารถใช้และเข้าใจประโยคในชีวิตประจำวัน (ในระดับกลาง) เช่น ข้อมูล เกี่ยวกับ

ระดับ	คำอธิบาย
	ครอบครัว การจับจ่ายใช้สอย สถานที่ ภูมิศาสตร์ การทำงาน และสามารถสื่อสารแลกเปลี่ยนข้อมูลทั่วไปในการใช้ชีวิตประจำวันสามารถบอกความคิด ความฝัน ความคาดหวัง ประวัติส่วนตัว สิ่งแวดล้อม และสิ่งอื่น ๆ
B1	ผู้เรียนสามารถพูด เขียน และจับใจความสำคัญของข้อความทั่ว ๆ ไปถ้าเป็นหัวข้อที่คุ้นเคยหรือสนใจ เช่น การทำงาน โรงเรียน เวลาว่าง ฯลฯ สามารถจัดการกับสถานการณ์ต่าง ๆ ที่เกิดขึ้นระหว่างการเดินทางในประเทศที่ใช้ภาษาอังกฤษได้ สามารถบรรยายประสบการณ์ เหตุการณ์ ความคิด ความฝัน พร้อมให้เหตุผลสั้นๆได้
B2	ผู้เรียนมีความสามารถในการใช้ภาษาในระดับดี สามารถใช้ภาษาพูดและเขียนได้แทบทุกเรื่องอย่างถูกต้องและคล่องแคล่วรวมทั้งสามารถอ่านและทำความเข้าใจบทความที่มีเนื้อหายากขึ้นได้
C1	ผู้เรียนสามารถเข้าใจ ข้อความยาวๆที่ซับซ้อนในหัวข้อหลากหลาย และเข้าใจความหมายแฝงได้สามารถแสดงความคิด ความรู้สึกของตนได้อย่างเป็นธรรมชาติ โดยไม่ต้องหยุดคิดหาคำศัพท์ สามารถใช้ภาษาทั้งในด้านสังคม การทำงาน หรือด้านการศึกษาได้อย่างมีประสิทธิภาพ
C2	ผู้เรียนมีความสามารถในการใช้ภาษาได้อย่างดีเยี่ยมใกล้เคียงเจ้าของภาษาสามารถใช้ภาษามาตรฐานได้อย่างสละสลวย ถูกต้องตามจุดประสงค์ที่จะสื่อสารได้ดี สามารถอ่านบทความที่เป็นภาษาต้นฉบับ (โดยเฉพาะวรรณกรรม) ได้เข้าใจ สามารถ และเลือกใช้ภาษาสำหรับพูดและเขียนได้อย่างเหมาะสม ผู้เรียนสามารถเข้าใจข้อความยาว ๆ ที่ซับซ้อนในหัวข้อหลากหลาย และเข้าใจความหมายแฝงได้

สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานกำหนดแนวปฏิบัติในการปฏิรูปการเรียนรู้ การสอนภาษาอังกฤษตามนโยบายแต่ละด้าน (สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2557) รวมถึงการกำหนดเป้าหมายการเรียนรู้ CEFR ได้จำแนก ระดับความสามารถด้านภาษาออกเป็นสามระดับหลัก และในแต่ละระดับก็ยังแบ่งออกเป็นระดับย่อย ซึ่งรวมทั้งหมดมี 6 ระดับดังต่อไปนี้

1) ระดับผู้ใช้ภาษาเบื้องต้น (Basic Language Users-A1, A2)

A1- มีความสามารถทางภาษาที่อยู่ในขอบเขตของความต้องการส่วนตัว และภาษา ที่ใช้ก็จะ เป็นลักษณะของภาษาที่ท่องจำมาจากที่เรียนหรือภาษาที่เดาการใช้ได้จากสถานการณ์ที่อยู่ต่อหน้าผู้พูด

A2- เป็นระดับความสามารถทางภาษาที่สามารถสื่อสารได้ในสถานการณ์ปกติ ธรรมดาทั่วไปในชีวิตประจำวันหรือเป็นการสนทนาในหัวข้อที่ตนเองคุ้นเคย

2) ระดับผู้ใช้ภาษาเอกเทศ (Independent Language User-B1, B2)

B1- สามารถใช้ภาษาพูดคุยในหัวข้อที่เกี่ยวกับสถานการณ์ทั่วไปในชีวิตประจำวัน ได้อย่างมีประสิทธิภาพโดยใช้ภาษาที่ง่าย ๆ โดยเฉพาะหัวข้อที่เป็นที่รู้จักดีหรือมีสถานการณ์ประกอบ ในการสนทนา

B2- สามารถใช้ภาษาในการสนทนาให้ข้อมูลและสามารถใช้ภาษาในการสนทนาโต้ตอบ อย่างมีประสิทธิภาพรวมถึงการใช้ความรู้และความเข้าใจทางภาษาเพื่อเป็นเครื่องมือในการวางแผน และแก้ไขข้อผิดพลาดของตนเอง

3) ระดับผู้ใช้ภาษาระดับความสามารถขั้นสูง (Proficient Language User-C1, C2)

C1- สามารถใช้ภาษาในการสนทนาโต้ตอบได้อย่างคล่องแคล่วและรวดเร็วทันต่อ เหตุการณ์และภาษาที่ใช้จะนำไปสู่การทำงานที่ซับซ้อนให้สำเร็จคล่องได้เป็นอย่างดี

C2- สามารถใช้ภาษาได้อย่างคล่องแคล่วถูกต้องและเหมาะสมกับสถานการณ์

บริติชเคาน์ซิล (British council, 2022) ระบุคำจำกัดความของระดับความสามารถทางภาษาตามกรอบอ้างอิง CEFR ด้วยภาษาต้นฉบับคือภาษาอังกฤษ ในระดับ A2 ดังนี้
SKILLS AT CEFR LEVEL A2

What language competences characterize a person who can prove they have a level A2 in English? The CEFR specifies the following:

He/she can understand sentences and frequently used expressions related to the areas of experience most immediately relevant to him/her (e.g. very basic personal and family information, shopping, places of interest, employment, etc.).

He/she can communicate in simple, everyday tasks requiring no more than a simple and direct exchange of information on familiar and routine matters.

He/she can describe in simple terms aspects of his/her past, environment and matters related to his/her immediate needs.

กระทรวงศึกษาธิการได้กำหนดให้ผู้เรียนชั้นมัธยมศึกษาปีที่ 3 ต้องมีความรู้ความสามารถด้านการอ่านภาษาอังกฤษ เมื่อเทียบกับกรอบ CEFR อยู่ในระดับ A2 โดยระบุคำอธิบายเป็นภาษาไทยเพื่อให้เข้าใจตรงกัน มีรายละเอียดคำอธิบายในตาราง 5 ดังนี้

ตาราง 5 แสดงคำอธิบายความสามารถด้านการอ่านภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3

ระดับความสามารถทางภาษา CEFR	การอ่านโดยรวม (OVERALL READING)	การอ่านเพื่อหาข้อมูลเบื้องต้น (READ FOR ORIENTATION)	การอ่านข้อมูลและข้อโต้แย้ง (READ INFO & ARGUMENT)	การอ่านคำสั่ง คำชี้แจง ขั้นตอน (READ INSTRUCTIONS)	การอ่านงานประพันธ์ (READ LITERATURE)	การอ่านจดหมาย (READ CORRESPONDENCE)
A2	สามารถเข้าใจบทอ่านสั้นๆ และง่าย ๆ ที่ประกอบด้วยคำศัพท์ที่คุ้นเคย	สามารถค้นหาข้อมูลสำคัญที่ต้องการในโฆษณา แผ่นพับ ใบปลิว หน้าเว็บไซต์ ตาราง เวลา และเข้าใจใจความสำคัญของข้อมูลนั้น	สามารถเข้าใจประเด็นสำคัญของเรื่องสั้นหรือข่าวง่ายๆ เมื่อผู้อ่านมีความรู้ในเรื่องนั้นอยู่บ้าง เช่น ข่าวกีฬา หรือบุคคลที่มีชื่อเสียง	สามารถเข้าใจคำแนะนำที่ใช้ภาษาง่ายๆ และขั้นตอนการใช้ เช่น การใช้โทรศัพท์ วิธีใช้เครื่อง ATM หรือตู้ ATM หรือวิธีซื้อเครื่องดื่มจากตู้อัตโนมัติ	สามารถเข้าใจประเด็นสำคัญของงานเขียนที่ใช้ภาษาง่ายๆ เกี่ยวกับชีวิตประจำวันที่มีสื่อหรือภาพประกอบ	สามารถเข้าใจข้อความสั้นๆ และง่าย ๆ จากเพื่อน เช่น อีเมล การสนทนาผ่านแอปพลิเคชัน เป็นต้น

เพื่อให้เห็นความสอดคล้องกันระหว่างกรอบอ้างอิง CEFR กับหลักสูตรแกนกลาง กลุ่มสาระการเรียนรู้ภาษาต่างประเทศระดับชั้นมัธยมศึกษาปีที่ 3 ผู้วิจัยจึงดำเนินการวิเคราะห์ความ สอดคล้องและนำเสนอข้อมูล โดยมีรายละเอียดในตาราง 6 ดังนี้

ตาราง 6 ความสอดคล้องระหว่าง CEFR กับหลักสูตรฯ กลุ่มสาระการเรียนรู้ภาษาต่างประเทศชั้น ม.3

สาระ และมาตรฐานการเรียนรู้ หลักสูตรแกนกลางฯ	กรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรป (CEFR)
<p>>มาตรฐานการเรียนรู้ ต 1.1 เข้าใจและตีความเรื่องที่ฟัง และอ่านจากสื่อประเภทต่าง ๆ และแสดงความคิดเห็น อย่างมีเหตุผล</p> <p>ตัวชี้วัด 1.ปฏิบัติตามคำขอร้อง คำแนะนำ คำชี้แจงและ คำอธิบายที่ฟังและอ่าน</p> <p>สาระแกนกลาง : คำขอร้อง คำแนะนำ คำชี้แจง และ คำอธิบายในการประดิษฐ์ การบอกทิศทางป้ายประกาศ ต่างๆ ขั้นตอนการใช้อุปกรณ์</p>	<p>>เป้าหมายตามกรอบอ้างอิง CEFR ระดับ A2</p> <p>เข้าใจบทอ่านสั้นๆและง่าย ๆ ที่ประกอบด้วยคำศัพท์ที่ ผู้อ่านคุ้นเคย</p> <p>การอ่านคำสั่ง คำชี้แจงขั้นตอน (READ INSTRUCTIONS) สามารถเข้าใจคำแนะนำที่ใช้ภาษา ง่ายๆและขั้นตอนการใช้ เช่น วิธีใช้โทรศัพท์ วิธีกดเงิน จากตู้ ATM หรือวิธีซื้อเครื่องดื่มจากตู้อัตโนมัติ</p>
<p>ตัวชี้วัด 3. ระบุและเขียนสื่อที่ไม่ใช่ความเรียงรูปแบบ ต่างๆ ให้สัมพันธ์กับประโยคและข้อความที่ฟังหรืออ่าน</p> <p>สาระแกนกลาง : การตีความ/ถ่ายโอนข้อมูลให้สัมพันธ์ กับสื่อที่ไม่ใช่ความเรียง เช่น สัญลักษณ์ เครื่องหมาย กราฟ แผนภูมิ ตาราง ภาพสัตว์ สิ่งของ บุคคล สถานที่ ต่างๆ</p>	<p>การอ่านเพื่อหาข้อมูลเบื้องต้น (READ FOR ORIENTATION) สามารถหาข้อมูลที่สำคัญ ในบทโฆษณาแผ่นพับ ใบปลิว หน้าเว็บ (web page) แคตตาล็อก ตารางเวลา ฯลฯ และเข้าใจใจความสำคัญ (main idea)</p>
<p>ตัวชี้วัด 4. เลือก/ระบุหัวข้อเรื่องใจความสำคัญ รายละเอียดสนับสนุนและแสดงความคิดเห็นเกี่ยวกับ เรื่องที่ฟังและอ่านจากสื่อประเภทต่างๆพร้อมทั้งให้ เหตุผลและยกตัวอย่างประกอบ</p> <p>สาระแกนกลาง : การจับใจความสำคัญ เช่น หัวข้อเรื่อง ใจความสำคัญ รายละเอียดสนับสนุนจากสื่อสิ่งพิมพ์และ สื่ออิเล็กทรอนิกส์ เช่น หนังสือพิมพ์ วารสาร วิทยุ โทรทัศน์ เว็บไซต์บนอินเทอร์เน็ต</p>	<p>การอ่านข้อมูลและข้อโต้แย้ง (READ INFORMATIONS & ARGUMENT) เป็นความสามารถเข้าใจประเด็นสำคัญ (main idea) ของเรื่องสั้นหรือข่าวง่ายๆเมื่อผู้อ่านมี ความรู้ในเรื่องนั้นอยู่บ้าง เช่น ข่าวกีฬา หรือบุคคลที่มี ชื่อเสียง</p> <p>การอ่านงานประพันธ์ (READ LITERATURE) สามารถ เข้าใจประเด็นสำคัญ(main idea) ของงานเขียนที่ใช้ ภาษาง่ายเกี่ยวกับชีวิตประจำวันที่มีสื่อหรือ ภาพประกอบ</p> <p>การอ่านจดหมาย(READ CORRESPONDENCE) สามารถเข้าใจข้อความสั้นๆและง่ายจากเพื่อน เช่น อีเมล การสนทนาผ่านแอปพลิเคชัน เป็นต้น</p>

จากตาราง 6 กรอบอ้างอิงความสามารถทางภาษาของสหภาพยุโรป (CEFR) มีความสอดคล้องกับหลักสูตรแกนกลางกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ ทั้งในส่วนของสาระการเรียนรู้ มาตรฐานการเรียนรู้ และตัวชี้วัด ดังนั้นในการศึกษาครั้งนี้ ผู้วิจัยจึงกำหนดขอบเขตเนื้อหาอิงเป้าหมายตามกรอบความสามารถทางภาษา CEFR ในระดับ A2 (ผู้เรียนระดับชั้นมัธยมศึกษาปีที่ 3) เนื่องจากเป็นยึดเป้าหมายหลัก ในการกำหนดนโยบายปฏิบัติการเรียนการสอนวิชาภาษาอังกฤษของกระทรวงศึกษาธิการในการศึกษาครั้งนี้ผู้วิจัย กำหนดคำอธิบายความสามารถด้านการอ่านในระดับ A2 คือ สามารถเข้าใจเนื้อหาในเอกสาร ตำราเรียน บทความ ฯลฯ ง่าย ๆ เข้าใจประเด็นสำคัญหรือใจความสำคัญ สามารถค้นหาข้อโต้แย้งง่าย ๆ ที่ปรากฏในเรื่องที่อ่าน สามารถเข้าใจด้วยการคาดเดาจากรูปภาพประกอบได้ และสามารถรู้ความหมายของคำจากการพินิจความรู้เดิมหรือค้นหาในพจนานุกรมได้

ตอนที่ 2 มโนทัศน์เกี่ยวกับการประเมินเชิงวินิจฉัยทางพุทธิปัญญา

1. ความหมาย ภูมิหลัง และความสำคัญของการประเมินเชิงวินิจฉัยทางพุทธิปัญญา

1.1 ความหมาย

การวินิจฉัย (Diagnostic) เป็นคำที่ใช้ในทางการแพทย์และพฤติกรรมศาสตร์เพื่อศึกษาปัญหาและสาเหตุของความเจ็บไข้และพฤติกรรมเบี่ยงเบนของบุคคลทั่วไป ในด้านการศึกษา มีการนำการวินิจฉัยมาใช้กันอย่างแพร่หลายมากขึ้นในช่วงหลายทศวรรษที่ผ่านมา “การวินิจฉัย” ในทางการแพทย์ (Medicine) หมายถึง การกระทำ หรือกระบวนการในการระบุสาเหตุของโรคหรือการบาดเจ็บผ่านทาง การประเมินประวัติของผู้ป่วย การตรวจ หรือการทบทวนข้อมูลทางห้องปฏิบัติการ ในทางชีววิทยา (Biology) “การวินิจฉัย” หมายถึง ข้อสรุปคุณลักษณะของระบบการจัดชั้นและหมวดหมู่ของพืชและสัตว์ (Taxonomic classification) ในด้านการศึกษา มีความหมายว่า การวินิจฉัยเป็นการค้นหาและระบุข้อบกพร่องของพฤติกรรมหรือความคิดที่ผิดปกติในตัวบุคคลด้วยวิธีการทดสอบที่เหมาะสม เพื่อแก้ไขหรือปรับปรุงพฤติกรรมที่ผิดปกตินั้น (Cohen, Swerdik & Smith, 1992; Gregory, 1992; Gipp, 1995; ศุภามณ จันทร์สกุล, 2561, หน้า 23; โกวิท ประวาลพุกษ์ และสงบ ลักษณะ อ้างถึงใน อาลี ทองคำวงศ์, 2562, หน้า 9) แม้วานิยามความหมายของ “การวินิจฉัย” อาจแตกต่างกันไปบ้างตามศาสตร์ที่ใช้ แต่มีลักษณะที่เหมือนกัน 3 ประการ (ศุภามณ จันทร์สกุล, 2561) คือ 1) การวิเคราะห์รายละเอียดและจำแนกคุณลักษณะสิ่งต่าง ๆ หรือปรากฏการณ์ 2) การระบุหรือกำหนดธรรมชาติของสิ่งต่าง ๆ หรือสาเหตุของปรากฏการณ์นั้น และ 3) การตัดสินใจหรือการสรุปซึ่งกระทำหรือเข้าถึงได้โดยการได้รายละเอียดหรือการวิเคราะห์

ความหมายของการประเมินเชิงวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostic Assessment: CDA) คือ กระบวนการที่เป็นระบบซึ่งใช้สำหรับการค้นพบสารสนเทศ ซึ่งเป็นทักษะ

ความสามารถทางปัญญาเฉพาะด้าน เพื่อค้นหาว่าผู้เรียนได้เรียนรู้อะไรมาบ้างเกี่ยวกับสิ่งที่จะเรียน สิ่ง ที่รู้มาก่อนนี้ถูกต้องหรือไม่ จึงเป็นการใช้ในลักษณะประเมินก่อนเรียน นอกจากนี้ยังใช้เพื่อให้เหตุผล ควบคุม และทำการตัดสินใจอย่างเหมาะสมเกี่ยวกับการกระทำที่เป็นผลอันเกิดมาจากลักษณะทางจิต หรือหาสาเหตุของปัญหาหรืออุปสรรคต่อการเรียนรู้ของผู้เรียนเป็นรายบุคคลที่มักจะเป็นเฉพาะเรื่อง เช่น ปัญหาการออกเสียงไม่ชัดแล้วหาวิธีปรับปรุงเพื่อให้ผู้เรียนสามารถพัฒนาและเรียนรู้ขั้นต่อไป วิธีการประเมินใช้ได้ทั้งการสังเกต การพูดคุยสอบถาม หรือการใช้แบบทดสอบก็ได้โดยมีจุดประสงค์ เพื่อออกแบบเครื่องมือในการวัดโครงสร้างความรู้และทักษะที่เฉพาะและวัดคุณลักษณะทางพุทธิ ปัญญาของนักเรียนอย่าง เฉพาะเจาะจงเป็นรายบุคคลที่จะให้สารสนเทศเกี่ยวกับผลการเรียนรู้ ผล การประเมินการจัดกิจกรรม การเรียนรู้ จุดแข็งและจุดอ่อนทางพุทธิปัญญาของนักเรียนจากการคิด และการเรียนรู้ โดยการอนุมานจากคะแนนสอบที่นักเรียนทำได้และให้ข้อมูลย้อนกลับไปในลักษณะ ของโปรไฟล์ความรู้ในแต่ละ คุณลักษณะของนักเรียน การประเมินเชิงวินิจฉัยทางพุทธิปัญญาจึง จำเป็นที่จะต้องมียุทธศาสตร์ที่ชัดเจนเพื่ออธิบายกระบวนการที่ผู้เรียนดึงกระบวนการทางพุทธิปัญญา ออกมา (Gierl, Leighton & Hunka, 2000; Ketterlin, Jung, Geller & Yovanoff, 2008, อ้างถึง ใน อาลี ทองคำวงศ์, 2562, หน้า 11; Rupp, Templin และ Henson, 2010, p. 11, อ้างถึงใน สุมาลี มีสกุล, 2558, หน้า 23; De la Torre and Minchen, 2014, p.90, อ้างถึงใน วราจคณา ก้อม น้อย, 2563, หน้า 59; แคทรียา แสงดาวเทียน, 2560, หน้า 2-3)

1.2 ภูมิหลังและความสำคัญ

การประเมินเชิงวินิจฉัยทางพุทธิปัญญาเริ่มต้นขึ้นเมื่อ Embretson ได้ตีพิมพ์ผลงานใน วารสาร Psychological Bulletin ในปีค.ศ. 1983 แสดงให้เห็นถึงการรวมกันของจิตวิทยาทางพุทธิ ปัญญาและความตรงเชิงโครงสร้างของความคิดร่วมสมัยว่าโครงสร้างหรือคุณลักษณะ (Construct) เป็นตัวแทนที่อ้างถึงความสัมพันธ์ที่ขึ้นกับการตอบสนองของงานบนกระบวนการ กลยุทธ์ และคลัง ความรู้ที่เกี่ยวข้องกับความสามารถตามที่แสดงออกมาให้เห็น (Performance) ผลงานของ Embretson เป็นการประยุกต์ของการพัฒนาจากจิตวิทยาพุทธิปัญญาไปสู่ทฤษฎีการวัด และใน หนังสือ Educational Measurement ที่ตีพิมพ์ในปีเดียวกันนี้ Snow and Lohman ได้แสดงให้เห็น อย่างชัดเจนในการนำแนวคิดทฤษฎีพุทธิปัญญาเข้ามาใช้ในการวัดทางการศึกษาในหัวข้อนี้ของ จิตวิทยาพุทธิปัญญาสำหรับการวัดทางการศึกษา (Implications of Cognitive Psychology for Educational Measurement) โดยชี้ให้เห็นว่าแนวคิด ทฤษฎี และวิธีการทางจิตวิทยาพุทธิปัญญา เป็นประโยชน์สำหรับการวัดทางการศึกษาคือให้ข้อมูลการวิเคราะห์ข้อสอบที่อธิบายถึงโครงสร้างของ ความเข้าใจ สร้างจุดมุ่งหมายการสอบที่ชัดเจนชี้ให้เห็นถึงความรอบรู้ และเพิ่มความแกร่งทางทฤษฎีที่ เกี่ยวข้อง ตั้งแต่ปีค.ศ. 1989 เป็นต้นมา มีบทความ หนังสือ และวรรณกรรมที่เขียนถึงความเฉพาะ และความโดดเด่นของการประเมินเชิงวินิจฉัยทางพุทธิปัญญาจำนวนมาก อาทิ กรอบโครงสร้าง

สำหรับการพัฒนาการประเมินเชิงวินิจฉัยทางพุทธิปัญญา (A Framework for Developing Cognitively Diagnostic Assessment) Nichols เขียนขึ้นในปีค.ศ. 1994 และหนังสือเรื่อง การประเมินเชิงวินิจฉัยทางพุทธิปัญญา (Cognitively Diagnostic Assessment) ของ Nichols, Chipman and Brennan ซึ่งเป็นบรรณาธิการร่วมกัน (Leighton and Gierl, 2007, อ้างถึงใน ไพฑูรย์ เอื้อบุญประดิษฐ์, 2563, หน้า 29-30)

ในช่วงปี 1983 – 1995 นับได้ว่าเป็นห้วงแห่งการรอกงามของการวินิจฉัยทางพุทธิปัญญาอันเนื่องมาจากนักวิชาการหลายท่านได้นำเสนองานวิจัยทางการวินิจฉัยทางพุทธิปัญญาที่สำคัญ อาทิ ทัตสึโอกะ (Tatsuoka, 1990) ได้พัฒนาโมเดลการวินิจฉัยทางพุทธิปัญญาเรียกว่า Rule Space Model เพื่อบ่งชี้ระดับความผิดปกติของแบบแผนการตอบข้อสอบโดยคาดการณ์แนวคิดของนักเรียนจากคำตอบที่นักเรียนได้แสดงออกมา Tatsuoka ได้เขียนกฎแสดงวิธีการคิดของนักเรียนในแต่ละรูปแบบซึ่งมีทั้งรูปแบบที่ถูกต้อง และไม่ถูกต้องเพื่อใช้ตรวจสอบวิธีการคิดของนักเรียน โดยนำคำตอบที่ได้มาตรวจสอบว่าตรงกันหรือใกล้เคียงกับลักษณะของกฎใด นอกจากนี้ยังมี เอ็มเบรทสัน (Embretson, 1994) ซึ่งนำเสนอแนวคิดระบบการออกแบบพุทธิปัญญา (Cognitive Design System Approach: CDS) ทั้งนี้เพื่อตอบข้อวิพากษ์ที่ว่า การออกแบบข้อสอบที่ไม่อิงทฤษฎีทำให้มีจุดอ่อนในความตรงเชิงโครงสร้างของการแปลผลคะแนน ระบบการออกแบบพุทธิปัญญาจึงถูกออกแบบโดยใช้ทฤษฎีพุทธิปัญญา (Cognitive theory) ในการพัฒนาข้อสอบเพื่อแก้ไขจุดอ่อนดังกล่าวทำให้การแปลผลคะแนนความสามารถมีความหมายขึ้น ระบบการออกแบบพุทธิปัญญาอยู่บนพื้นฐานกระบวนการตอบสนองทางพุทธิปัญญาสำหรับการประเมินที่วัดความสามารถพื้นฐานทางพุทธิปัญญา (Basic cognitive abilities) ได้แก่ Spatial Rotation หรือเหตุผลโดยทั่วไป (General reasoning) (ศุภามณ จันทร์สกุล, 2561)

การวินิจฉัยเป็นกระบวนการสำคัญในการวัดและการประเมินผลทางการศึกษา เนื่องจากการเชื่อมโยงและสร้างความสอดคล้องในองค์ประกอบของการจัดการศึกษา ทั้งการบริหารจัดการการเรียนรู้และการวินิจฉัยจากการวัดและการประเมินผล (แคทรียา แสงดาวเทียน, 2560, หน้า 18) การวินิจฉัยพุทธิปัญญาของผู้เรียน มีวัตถุประสงค์หลักเพื่อค้นหารายละเอียดในคุณลักษณะ/ทักษะต่าง ๆ ของผู้เรียน ซึ่งจะสะท้อนให้เห็นถึงระดับของความรอบรู้ของผู้เรียน ทั้งนี้เพื่อนำไปสู่การให้ข้อมูลย้อนกลับในระหว่างดำเนินการเรียนการสอน (Formative diagnostic feedback) เกี่ยวกับความรอบรู้ในองค์ประกอบที่เป็นทักษะจำเป็นในการเรียนทำให้ผู้เรียนเห็นภาพ และสามารถปรับปรุงทักษะที่บกพร่องหรือไม่รอบรู้ได้ นอกจากนี้ผลการประเมินยังให้การรายงานข้อวินิจฉัยที่ได้สู่ผู้เรียน ครู ผู้ปกครอง เพื่อร่วมกัน ปรับปรุงแก้ไขการเรียนของผู้เรียน (Leighton & Gierl, 2007, อ้างถึงใน ศุภามณ จันทร์สกุล, 2561, หน้า 24) โดยมีเป้าหมายมุ่งเน้นลักษณะเชิงพุทธิ

ปัญญาอย่างน้อย 3 ด้าน (Yang & Embretson, 2007, อ้างถึงใน แคทรียา แสงดาวเทียน, 2562, หน้า 29-30) คือ

1) โพรไฟล์ทักษะ (Skill profile) หรือความรู้ที่จำเป็นในเนื้อหาเชิงพุทธิปัญญาที่ให้ชุดของทักษะและความรู้แทนทักษะและมโนทัศน์ที่สำคัญมากที่สุดของเนื้อหา และเป็นพื้นฐานในการสร้างกรอบสำหรับการพัฒนาสมรรถนะลำดับขั้นที่สูงขึ้น

2) โครงสร้างขั้นตอนหรือเครือข่ายความรู้ ทักษะความรู้ในกลุ่มเนื้อหาเป็นตัวแทนมวลเนื้อเรื่องของทักษะพื้นฐาน ความรู้ในกลุ่มเนื้อหา โครงสร้างหรือกลุ่มของทักษะและความรู้

3) คุณลักษณะ องค์ประกอบ หรือความสามารถ กระบวนทัศน์ของการศึกษาเชิงพุทธิปัญญาให้วิธีที่ได้ประโยชน์จากกระบวนการภายในของความคิด ดังนั้น โมเดลเชิงพุทธิปัญญาที่เฉพาะสามารถพัฒนาเพื่องานด้านพุทธิปัญญาได้ จึงสามารถอธิบายกระบวนการเชิงพุทธิปัญญาของผู้สอบได้

1.3 ประเภทของการวินิจฉัย

การวินิจฉัยสามารถจำแนกได้หลายประเภทตามเกณฑ์ที่นำมาใช้ในการจำแนก 2 เกณฑ์ ได้แก่ เกณฑ์การจำแนกตามระดับความละเอียดในการวินิจฉัย และเกณฑ์การจำแนกตามจุดมุ่งหมายของการวินิจฉัย

1) เกณฑ์การจำแนกตามระดับความละเอียดในการวินิจฉัย (Underhill, 1972, อ้างถึงใน อาลี ทองคำวงศ์, 2562, หน้า 9-10) สามารถจำแนกได้เป็น 3 ประเภท ได้แก่

1.1 การวินิจฉัยแบบทั่วไป (General diagnosis) เป็นขั้นการสำรวจเพื่อทราบระดับความสามารถทั่ว ๆ ไปของนักเรียนรายบุคคลและรายกลุ่ม เครื่องมือที่ใช้ส่วนใหญ่เป็นแบบทดสอบทั้งแบบทดสอบวัดผลสัมฤทธิ์ที่ครูสร้างขึ้นและแบบทดสอบมาตรฐาน อาจใช้การวินิจฉัยโดยการเปรียบเทียบกับเกณฑ์ปกติ (Norm) ซึ่งจะทำให้ทราบข้อบกพร่องต่าง ๆ ของนักเรียนอย่างกว้าง ๆ

1.2 การวินิจฉัยแบบวิเคราะห์ (Analytical diagnosis) เป็นการวินิจฉัยระดับเฉพาะ (Specific level) โดยการเก็บรวบรวมข้อมูลโดยละเอียดเกี่ยวกับความสามารถของ นักเรียนเพื่อทราบข้อบกพร่องของนักเรียนในเรื่องใดเรื่องหนึ่ง เครื่องมือที่ใช้ส่วนใหญ่เป็นแบบทดสอบที่มีเนื้อหาเฉพาะเจาะจงในเรื่องใดเรื่องหนึ่ง การวินิจฉัยแบบเฉพาะมักดำเนินการภายหลังจากการวินิจฉัยแบบทั่วไปเพื่อให้ได้ข้อมูลที่มีรายละเอียดที่ลึกซึ้งมากขึ้น

1.3 การวินิจฉัยแบบคลินิก (Clinical diagnosis) เป็นการวินิจฉัยระดับละเอียดลึกซึ้ง เพื่อให้เห็นถึงสาเหตุของปัญหาหรือข้อบกพร่องที่พบในตัวนักเรียนซึ่งมีความซับซ้อน การใช้

ข้อมูลจากแหล่งใดแหล่งหนึ่งเพียงแหล่งเดียวจึงไม่เพียงพอ จำเป็นต้องใช้ข้อมูลจากหลายแหล่งประกอบกัน เช่น การใช้ข้อมูลจากแบบทดสอบประกอบการสัมภาษณ์ และการศึกษาสภาพครอบครัว

2) เกณฑ์การจำแนกตามจุดมุ่งหมายของการวินิจฉัย สามารถจำแนกได้ 3 ประเภท (สุพัทธ์ พิบูลย์, 2553, อ้างถึงใน อาลี ทองคำวงศ์, 2562, หน้า 9-10) ได้แก่

2.1 การสอบวินิจฉัยเพื่อวัดความรู้ หรือทักษะพื้นฐานที่จำเป็นของผู้เรียนหรือวัดความรู้ในเรื่องใด ๆ หมายถึง การวินิจฉัยความรู้พื้นฐานในเรื่องที่เกี่ยวข้องที่จะเป็นฐานความรู้สำคัญในการเรียนเรื่องนั้น ๆ ให้ประสบความสำเร็จและเป็นการสอบเพื่อวินิจฉัยความรอบรู้ในเรื่องนั้น ๆ เช่น การทดสอบความรู้พื้นฐานเดิมโดยไม่ได้มุ่งเน้นในการวินิจฉัยข้อบกพร่อง ซึ่งมักปรากฏในการทดสอบก่อนเรียน (pre – test) เพื่อนำคะแนนไปเปรียบเทียบกับคะแนนทดสอบหลังเรียน (post – test) แล้วตัดสินระดับพัฒนาการเรียนรู้ของผู้เรียน

2.2 การสอบวินิจฉัยเพื่อวิเคราะห์สาเหตุของความบกพร่องในการเรียนที่เกิดขึ้น เพื่อจะได้กำหนดวิธีสอนที่เหมาะสมหรือแก้ปัญหาได้ตรงจุด เป็นการวินิจฉัยหลังจากเกิดปัญหาในการเรียนแล้ว ครูผู้สอนต้องการตรวจสอบว่าปัญหาของนักเรียนเกิดจากความบกพร่องในลักษณะใด

2.3 การสอบวินิจฉัยเพื่อการวิเคราะห์โครงสร้างการเรียนรู้สำหรับการออกแบบการเรียนรู้ในเรื่องต่าง ๆ เป็นการวินิจฉัยเพื่อกำหนดโครงสร้างการเรียนรู้ว่าเรื่องใดควรเรียนรู้ก่อนหรือหลัง เรื่องใดมีผลเชื่อมโยงต่อการเรียนรู้อีกเรื่องหนึ่งเช่น ความรู้คำศัพท์มีผลทำให้ผู้อ่านบทอ่านภาษาอังกฤษสามารถทำความเข้าใจและบอกสาระสำคัญของเรื่องที่อ่านได้

จากประเด็นเกี่ยวกับประเภทของการวินิจฉัยข้างต้น กล่าวได้ว่าการศึกษาค้นคว้าครั้งนี้เป็นการพัฒนาเครื่องมือที่ใช้ในการวินิจฉัยแบบวิเคราะห์ ซึ่งเป็นการวินิจฉัยระดับเฉพาะเกี่ยวกับคุณลักษณะความสามารถของนักเรียนเพื่อประเมินความรอบรู้ในเรื่องใดเรื่องหนึ่งโดยเฉพาะ และหากจำแนกตามจุดมุ่งหมายของการวินิจฉัยงานวิจัยฉบับนี้มีความมุ่งหมายเพื่อนำเครื่องมือที่ผู้วิจัยพัฒนาขึ้นทำการวินิจฉัยความสามารถผู้เรียนซึ่งจะได้รูปแบบความรอบรู้ที่ให้สารสนเทศนำมาใช้ในการกำหนดโครงสร้างการเรียนรู้ว่าเรื่องใดควรเรียนรู้ก่อนหรือหลัง เรื่องใดมีผลเชื่อมโยงต่อการเรียนรู้อีกเรื่องหนึ่งซึ่งจัดว่าเป็นการทดสอบวินิจฉัยทางการเรียน

ศิริชัย กาญจนวาสี (2552) ได้กล่าวถึง การประเมินเพื่อวินิจฉัยว่าเป็นการประเมินผลเพื่อวินิจฉัยสาเหตุของปัญหาการเรียนรู้ที่เกิดขึ้นระหว่างการเรียนการสอน เช่น ปัญหาเกี่ยวกับผู้เรียนสิ่งแวดล้อม เนื้อหาวิชาเป็นต้น เพื่อที่จะได้ทำการปรับปรุงแก้ไขปัญหาได้ถูกต้องตรงประเด็น เครื่องมือที่ใช้วัดอาจเป็น แบบสอบวินิจฉัยที่เผยแพร่ทั่วไป (Published Diagnostic Tests) แบบสอบวินิจฉัยที่ครูผู้สอนสร้างเอง (Teacher-made Diagnostic Tests) หรือเทคนิคการสังเกต (Observation Techniques) เป็นต้น การทดสอบวินิจฉัยทางการเรียนจึงเป็นการทดสอบที่มีวัตถุประสงค์ในการค้นหาเพื่อระบุจุดเด่นและจุดด้อยของผู้เรียน หรือประเด็นที่นักเรียนยังไม่รอบรู้

เพื่อให้ข้อเสนอแนะเกี่ยวกับสาเหตุ หรือเหตุผลที่เป็นไปได้ว่าเพราะเหตุใดนักเรียนถึงยังไม่รอบรู้ในประเด็นนั้นๆ โดยใช้เป็นข้อมูลป้อนกลับ (Feedback) นำไปใช้ประโยชน์ในการปรับปรุงการจัดการเรียนรู้เพื่อพัฒนานักเรียนได้ตรงจุดที่บกพร่อง (Nitko and Brookhart, 2014, p. 84, อ้างถึงใน วรางคณา ก้อมน้อย, 2563, หน้า 59) นับว่าเป็นอีกแนวทางที่มีความสำคัญและมีคุณค่าดังจะกล่าวเพิ่มเติมในเรื่องเกี่ยวกับการประเมินเชิงวินิจฉัยทางพุทธิปัญญา

ต่อมาเมื่อมีการศึกษาการนำโครงสร้างทางพุทธิปัญญาเข้ามาใช้ในการออกแบบข้อสอบ การให้คะแนน และการรายงานถึงการสอนและการเรียนรู้ที่คิดว่ามีการพัฒนาโมเดลการให้คะแนน โดยใช้ทักษะทางพุทธิปัญญาเป็นฐานเพื่อนำมาใช้ในการประเมินเชิงวินิจฉัยทางพุทธิปัญญา ที่ในปัจจุบันเรียกว่าโมเดลวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostic Models: CDMs) (Huff and Goodman, 2007, อ้างถึงใน ไพฑูริย์ เอื้อบุญประดิษฐ์, 2563)

2. โมเดลวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostics Models: CDMs)

การประเมินเชิงวินิจฉัยทางพุทธิปัญญาเป็นการรวมกันของศาสตร์ 2 ศาสตร์ คือ จิตวิทยาพุทธิปัญญา (Cognitive Psychology) และโมเดลการวัดทางจิตมิติ (Psychometric Modeling) โดยจิตวิทยาพุทธิปัญญาจะช่วยให้ทราบถึงโครงสร้างคุณลักษณะทางพุทธิปัญญาในการเรียนรู้ ในขณะที่โมเดลการวัดทางจิตมิติทำให้ได้สารสนเทศในการวินิจฉัยความรอบรู้ของนักเรียนในแต่ละคุณลักษณะทางพุทธิปัญญาที่ใช้ในการเรียนรู้ ดังนั้น การประเมินเชิงวินิจฉัยทางพุทธิปัญญาจึงเป็นรูปแบบหนึ่งของการทดสอบโดยใช้โมเดลทางพุทธิปัญญา (Cognitive Model) มาพัฒนาหรือกำหนดข้อสอบที่วัดองค์ความรู้และทักษะที่เฉพาะเจาะจง และใช้โมเดลนี้โดยตรงกับการวิเคราะห์ทางจิตของรูปแบบการตอบสนองต่อข้อสอบของผู้ตอบ (Item Response) เพื่อส่งเสริมการวินิจฉัยที่เฉพาะเจาะจง โดยใช้โมเดลทางสถิติในการวิเคราะห์ที่สามารถสกัดสารสนเทศเชิงวินิจฉัยจากข้อมูลได้ (Leighton and Gierl, 2007, อ้างถึงใน Gierl and Zhou, 2008, อ้างถึงใน วรางคณา ก้อมน้อย, 2563, หน้า 4-5)

Lee and Sawaki (2009, อ้างถึงใน กุลพร พูนสวัสดิ์, 2558, หน้า 50) อธิบายว่าโมเดลการวัดทางจิตวิทยาที่ใช้ในการวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostic Psychometric Models: CDPMs) ได้ถูกออกแบบสำหรับแบบสอบวินิจฉัยโดยเฉพาะ แต่ก็สามารถใช้ในการสกัดข้อมูลเชิงวินิจฉัยจากแบบสอบที่มีอยู่แล้วได้ด้วย ซึ่ง CDPMs ถือเป็น Latent Class Models ซึ่งแสดงกลุ่มของผู้เรียนที่มีความรอบรู้และไม่รอบรู้ในทักษะที่ถูกระบุไว้ในเมทริกซ์คิว (Q-matrix) และในข้อตกลงเบื้องต้นเกี่ยวกับคุณลักษณะ (Attributes) มี 2 แบบ

1) Conjunctive (non-compensatory) หมายถึง การที่จะตอบข้อคำถามได้นั้นจะต้องมีทักษะหรือคุณลักษณะทั้งหมดที่กำหนด

2) Disjunctive (Compensatory) หมายถึง หากขาดทักษะใดทักษะหนึ่งหรือคุณลักษณะใด ผู้สอบยังสามารถตอบข้อคำถามได้อย่างถูกต้อง เนื่องจากทักษะที่แข็งแกร่งจะไปชดเชยทักษะที่อ่อน

ในการศึกษาวรรณกรรมที่เกี่ยวข้องของนักวิจัยอาจพบว่า มีชื่อเรียกโมเดลที่ใช้ในการประเมินเชิงวินิจฉัยทางพุทธิปัญญาแตกต่างกันออกไป เนื่องจากโมเดลเชิงวินิจฉัยทางพุทธิปัญญามีชื่อเรียกต่าง ๆ จำแนกตามความหมายแฝง มีรายละเอียดแสดงในตาราง 7 (Rupp & Templin, 2007, pp. 5-7; Rupp, Templin & Henson, 2010, p.3, อ้างถึงใน วราจคณา ก้อมน้อย, 2563, หน้า 60-61) ดังนี้

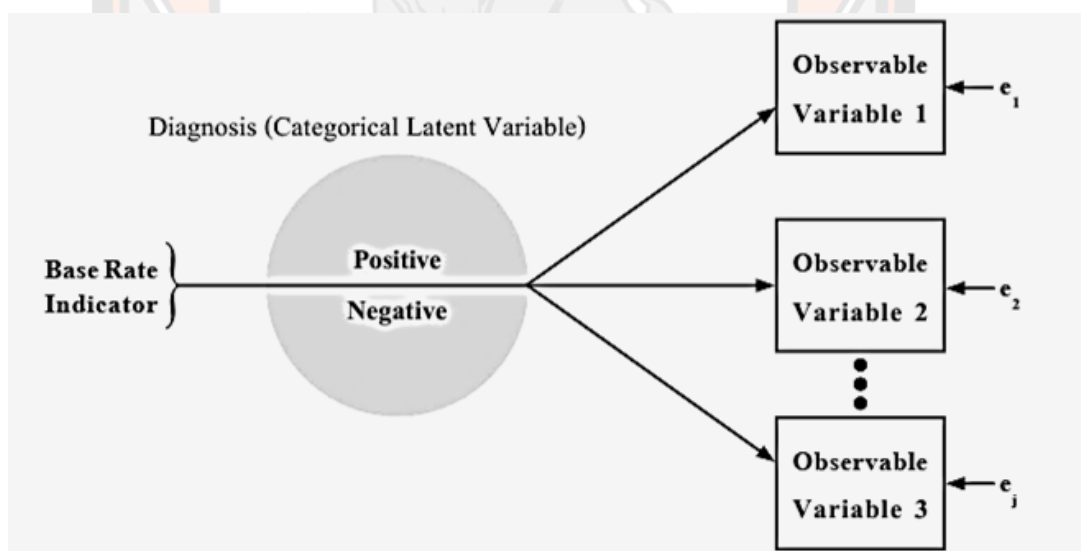
ตาราง 7 ชื่อเรียกโมเดลวินิจฉัยทางพุทธิปัญญาจำแนกตามความหมายแฝง

ความหมายแฝงที่สำคัญ	ชื่อที่ใช้เรียกโมเดล
รากฐาน คุณสมบัติ หรือจุดมุ่งหมายทางสถิติโดยเน้นว่าโมเดลกลุ่มนี้มีรากฐานแนวคิดในการจัดกลุ่มตัวแปรในลักษณะชั้นแฝงที่มีจำนวนจำกัด และมีวัตถุประสงค์การใช้โมเดลเหล่านี้ในการสร้างโปรไฟล์พุทธิปัญญา	-Restricted latent class models (Macready and Dayton, 1976; Haertel, 1980), -Latent response models (Maris, 1995) -Multiple classification (latent class) models (Maris, 1995, 1999)
คุณลักษณะแฝง (Traits) ของผู้ตอบ บนฐานของการจำแนกโดยขึ้นอยู่กับระดับของความรอบรู้หรือไม่รอบรู้ของคุณลักษณะแฝงนั้น ๆ	-Structured IRT models (Mislevy, 2007; Rupp & Mislevy, 2007) -Structured located latent class models (Xu & Von Davier, 2008)
ทฤษฎีที่อยู่เบื้องหลังของการนำไปใช้โดยเน้นว่าไม่ว่าจะนำโมเดลใดไปใช้ก็ตาม ทุกโมเดลล้วนต้องการรายละเอียดของทฤษฎีการตอบสนองซึ่งอยู่ในรากฐานของจิตวิทยาพุทธิปัญญา	-โมเดลจิตมิติทางพุทธิปัญญา (Cognitive Diagnostic Psychometric Models: CDPMs) (Rupp, 2007; Lee and Sawaki, 2009)
ให้ความสำคัญกับหลักการของจุดมุ่งหมายที่แท้จริงของการใช้โมเดล โดยเน้นว่าจุดมุ่งหมายที่แท้จริงของการใช้โมเดลคือการวินิจฉัยบนฐานทางพุทธิปัญญา(Cognitively-grounded Diagnosis) โดยมีใช้เป็นไปเพื่อการจัดอันดับแต่ให้ความสำคัญกับการวินิจฉัยเพื่อให้ข้อมูลสารสนเทศในระดับบุคคล	-โมเดลวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostic Models: CDM) (Nichols, Chipman, & Brennan, 1995; Templin & Henson, 2006; De La Torre, 2011)

อย่างไรก็ตามถึงแม้ว่าโมเดลวินิจฉัยทางพุทธิปัญญาที่พบในวรรณกรรมจะมีชื่อเรียกแตกต่างกันออกไป แต่คำเหล่านี้สามารถใช้แทนกันได้ (Lee & Sawaki, 2009) ดังนั้นในการศึกษาครั้งนี้ ผู้วิจัยนำเสนอในชื่อโมเดลวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostic Models: CDMs)

เพื่อเป็นที่เข้าใจตรงกัน และสื่อความหมายถึงแนวคิดที่ใช้ในการออกแบบการประเมินและการสร้างแบบสอบ โดยให้ความสำคัญกับหลักการของจุดมุ่งหมายที่แท้จริงของการใช้โมเดลคือการวินิจฉัยบนฐานทางพุทธิปัญญา (Cognitively-grounded Diagnosis) โดยมีใช่เป็นไปเพื่อการจัดอันดับ แต่ให้ความสำคัญกับการวินิจฉัยเพื่อให้ข้อมูลสารสนเทศในระดับบุคคล

ศุภามณ จันทรสกุล และสุกัญญา บุญศรี (2560, หน้า 57) กล่าวว่า โมเดลวินิจฉัยเชิงจำแนกหรือจัดประเภท (Diagnostic Classification Models: DCMs) เป็นวิธีการที่น่าสนใจในการประเมินวินิจฉัยความรอบรู้ เพื่อนำเสนอวิธีการประเมินวินิจฉัยโดยใช้โมเดลทางสถิติในการประเมินวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostic Assessment: CDA) เป็นการค้นหาแบบแผนการคิดคุณลักษณะ/ทักษะต่าง ๆ ของผู้เรียนซึ่งสะท้อนให้เห็นถึงระดับของความรอบรู้ที่แสดงถึง มโนทัศน์ความรู้และความเข้าใจที่ถูกต้อง (จุดแข็ง) และข้อบกพร่องที่แสดงถึงมโนทัศน์คลาดเคลื่อน และการพร่องความรู้ (จุดอ่อน) ในการเรียนของผู้เรียน (Leighton & Gierl, 2007; Rupp, Templin & Henson, 2010) สารสนเทศที่ได้จะเป็นข้อมูลย้อนกลับ (Feedback) สำคัญทำให้ผู้เรียนเห็นข้อบกพร่องในการเรียนของตน ตลอดจนเป็นประโยชน์ต่อผู้สอนในการปรับแก้และพัฒนาการเรียนให้เหมาะสมกับผู้เรียนแต่ละรายบุคคล



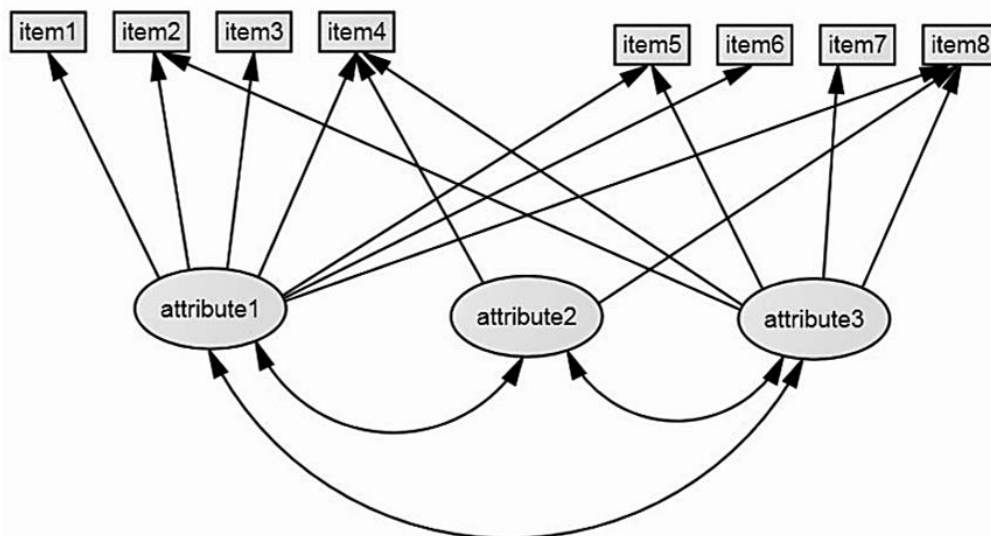
ภาพ 8 โครงสร้างโมเดลวินิจฉัยเชิงจำแนก
(ศุภามณ จันทรสกุล และสุกัญญา บุญศรี, 2560)

จากภาพ 8 ทางซ้ายมือแสดงถึงตัวแปรแฝงมีสัญลักษณ์เป็นวงกลมโดยมีเส้นแบ่งกลางวงกลม หมายถึง การจัดตัวแปรแฝงเป็น 2 ประเภทโดยมีตัวบ่งชี้พื้นฐาน (Base rate indicator) เป็นเกณฑ์การแบ่งการวินิจฉัยตำแหน่งเหนือเส้นตัวบ่งชี้พื้นฐานคือ “positive” หรือกลุ่มที่อยู่สูงกว่า

เกณฑ์การแบ่งส่วนตำแหน่งใต้เส้นตัวบ่งชี้พื้นฐานคือ “negative” หรือกลุ่มที่อยู่ตำแหน่งต่ำกว่าเกณฑ์การแบ่งสี่เหลี่ยมขวามือแสดงถึงตัวแปรสังเกตได้ซึ่งเป็นตัวบ่งชี้ของตัวแปรแฝง โดยมี e เป็นส่วนของความคลาดเคลื่อน (Error term) คล้ายโมเดลการวิเคราะห์องค์ประกอบ (Factor analysis)

เมื่อพิจารณาในขอบเขตของการวัดเชิงจิตมิติทางการศึกษา (Educational Psychometric Measurement: EPM) รูปแบบการวินิจฉัยโดย CDMs มีความแตกต่างกับทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) ที่มุ่งเน้นการวัดผลสัมฤทธิ์ของผู้เรียนโดยวัดผลจากคะแนนจากการทดสอบ และทฤษฎีการตอบสนองต่อข้อสอบ (Item Response Theory: IRT) ที่มุ่งเน้นการวัดความสามารถของผู้เรียน (Ability) ซึ่งแปรผันตามค่าพารามิเตอร์ของแบบทดสอบ ในขณะที่โมเดลวินิจฉัยทางพุทธิปัญญา CDM เป็นโมเดลการวัดผลทางการศึกษาที่ถูกพัฒนาขึ้น เพื่อวินิจฉัยความสามารถในการเรียนรู้ของผู้เรียนโดยมุ่งเน้นเพื่อการตรวจสอบข้อมูลความรู้ (Mastery) ของผู้เรียนต่อทักษะเฉพาะ (Specific Skills) ที่จำเป็นต่อการเรียนรู้ในรายวิชานั้น ๆ ส่วนในประเด็นอื่นทั้งสองทฤษฎีการทดสอบเป็นการกำหนดคะแนนให้ผู้ตอบเพียงคะแนนเดียว (Single Score) บนมาตราต่อเนื่อง (Continuous Scale) มักมีวัตถุประสงค์เพื่อเปรียบเทียบหรือจัดอันดับผู้ตอบ เมื่อเทียบกับกลุ่มหรือเกณฑ์มาตรฐาน แต่โมเดลวินิจฉัยทางพุทธิปัญญา (CDM) เป็นการกำหนดโปรไฟล์ทักษะหรือคุณลักษณะพหุมิติ (Multidimensional Skill Profiles) ให้กับผู้ตอบโดยจำแนกจากความรู้หรือไม่รู้ในทักษะหรือคุณลักษณะที่ใช้ตอบแบบทดสอบนั้น ๆ จึงสามารถให้สารสนเทศเพื่อการวินิจฉัยได้ ตัวแปรแฝง (Latent Variables) ใน CDMs เป็นตัวแปรแบบไม่ต่อเนื่อง เพราะเป็นตัวแปรแฝงที่แยกจากกันโดยสิ้นเชิง โดยบ่งชี้ถึงความรอบรู้และไม่รอบรู้เพียง 2 ค่า ในขณะที่การประมาณค่าความสามารถผู้สอบ หรือ ค่า θ ในโมเดลการตอบสนองข้อสอบพหุมิติ (MIRT) และโมเดลการวิเคราะห์องค์ประกอบ (FA) เป็นตัวแปรต่อเนื่อง (Thissen, & Steinberg, 1986, อ้างถึงใน เอกราช ดินาง, 2559, หน้า 3-4; de la Torre and Minchen, 2014, pp. 89-90 และ Ravand and Robitzsch, 2015, p. 2, อ้างถึงใน วรางคณา ก้อนน้อย, 2563, หน้า 60-61)

เมตริกซ์คิว หรือ Q-matrix (Tatsuka, 1985) เป็นโครงสร้างของโมเดลวินิจฉัยเชิงพุทธิปัญญา CDMs เป็นสมมติฐานเกี่ยวกับคุณลักษณะที่จำเป็นสำหรับการทำให้ข้อสอบแต่ละข้อมีความถูกต้อง (Li, 2011) คุณสมบัติอื่นของ CDMs คือความเป็นพหุมิติ CDMs เป็นโมเดลตัวแปรแฝงพหุมิติ เนื่องจากประมาณค่าความสามารถพหุประมาณค่าโปรไฟล์คุณลักษณะของผู้เรียนที่ไม่สามารถสังเกตได้มากกว่าหนึ่งคุณลักษณะ ตัวอย่างเช่น ในรูปที่ 9 ข้อสอบข้อที่ 2 และข้อที่ 5 วัดได้สองคุณลักษณะ และข้อสอบข้อที่ 4 และข้อที่ 8 วัดสามทักษะ ดังนั้นข้อสอบทั้ง 4 ข้อนี้มีโครงสร้างที่มีลักษณะพหุมิติ ดังภาพที่ 9



ภาพ 9 โครงสร้างที่มีลักษณะพหุมิติของ CDMs

จากการทบทวนวรรณกรรมอย่างครอบคลุม Rupp and Templin (2008, cited in Ravand, 2015, p.2) ได้กล่าวว่า CDMs มีลักษณะบางประการ เช่น ความเป็นพหุมิติ ลักษณะองค์ประกอบเชิงยืนยัน โครงสร้างน้ำหนักองค์ประกอบที่ซับซ้อน และลักษณะความน่าจะเป็น เช่นเดียวกับการวิเคราะห์องค์ประกอบ (Factor Analysis: FA) โมเดลการตอบสนองข้อสอบ (Item Response Theory: IRT) หรือโมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory: MIRT) หากกล่าวในเชิงสถิติโมเดลวินิจัยทางพุทธิปัญญา (CDMs) คือ ความน่าจะเป็นที่ได้จากการยืนยันโมเดลตัวแปรแฝงพหุมิติด้วยโครงสร้างน้ำหนักองค์ประกอบเชิงเดียวหรือน้ำหนักองค์ประกอบเชิงซ้อน ที่เหมาะสำหรับการสร้างโมเดลตัวแปรตอบสนองแบบแบ่งกลุ่ม (Categorical Response Variables) และมีตัวแปรพยากรณ์เป็นตัวแปรแฝงแบบแบ่งกลุ่ม (Categorical Latent Predictor Variables) CDMs ประกอบด้วยตัวแปรพยากรณ์แฝงหลายตัวที่นำไปสู่การสร้างชั้นแฝง

ใน CDMs จะมีการกำหนดโครงสร้างไว้อย่างจำกัด ข้อสอบแต่ละข้อสามารถวัดคุณลักษณะหรือทักษะย่อยได้มากกว่าหนึ่งทักษะ ซึ่งนับว่ามีลักษณะของข้อสอบแบบพหุมิติ (Adams, Wilson, & Wang, 1997; Lee and Sawaki, 2009; Baghaei, 2012) CDMs เป็นโมเดล Confirmatory ตรงที่มีกระบวนการ กลวิธี หรือทักษะย่อยที่จำเป็นในการดำเนินการให้สำเร็จในข้อสอบในแบบทดสอบที่กำหนดคือระบุไว้ใน Q-matrix (Tatsuoka, 1983) ตามทฤษฎีโครงสร้างความรู้ โดยใช้ในการเปรียบเทียบจากการวิเคราะห์องค์ประกอบเชิงยืนยัน Q-matrix คือ น้ำหนักองค์ประกอบของความสัมพันธ์ระหว่างข้อคำถามกับคุณลักษณะของ CDMs ถูกตั้งสมมติฐาน จากนั้นจึงทดสอบ Q-matrix ที่เทียบกับข้อมูลจริง CDMs เป็น confirmatory ในแง่อื่นซึ่งไม่ค่อยมีใครพูดถึง

สอดคล้องกับ Rupp และ Templin (2008) กล่าวไว้ว่า CDMs เป็น confirmatory ตรงที่ควรระบุว่า คุณลักษณะหรือทักษะย่อยมีปฏิสัมพันธ์ระหว่างกันอย่างไรในกระบวนการตอบสนองข้อสอบเป็น ขั้นตอนแรกในอันดับต้นๆนั้นคือไม่ว่าคุณลักษณะจะรวมกันในความสัมพันธ์แบบชดเชยได้เพื่อสร้าง คำตอบที่ถูกต้องก็ควรจะถูกระบุล่วงหน้าและเลือกโมเดลที่เหมาะสม โดยพิจารณาแนวคิดทฤษฎีที่เกี่ยวข้องเป็นตัวกำหนด (Ravand, 2015, p. 2)

สโนว์ และโลห์แมน (Snow and Lohman, 1989 อ้างถึงใน Ravand and Robitzsch, 2015, p. 2, อ้างถึงใน วรางคณา ก้อมน้อย, 2563, หน้า 63-64) กล่าวว่า การใช้โมเดลวินิจฉัยทาง พุทธิปัญญา (CDMs) สามารถเอาชนะข้อจำกัดบางอย่างของทฤษฎีการตอบสนองต่อข้อสอบ (IRT) ได้ ดังนี้ สามารถอธิบายการตอบสนองต่อข้อสอบของผู้ตอบด้วยทฤษฎีทางจิตอย่างแท้จริง สามารถ อธิบายกระบวนการทางจิตที่เกิดขึ้นจากการวัดตัวแปรเชิงโครงสร้างด้วยแบบทดสอบได้อย่างชัดเจน สามารถทำให้สมมุติฐานที่ใกล้เคียงความจริงมากขึ้น เกี่ยวกับตัวแปรที่ส่งผลต่อคุณภาพของข้อคำถาม ในแบบทดสอบได้ แตกต่างกับทฤษฎีการ ตอบสนองต่อข้อสอบ (IRT) เช่นกรณี โมเดล 3 พารามิเตอร์ (Three-parameter IRT model) ซึ่งมีสมมุติฐานว่ามีตัวแปรที่ส่งผลต่อคุณภาพของข้อคำถามใน แบบทดสอบเพียงสามตัวแปร และสิ่งที่สำคัญที่สุดหากสิ่งที่ต้องการทราบเกี่ยวกับผู้ตอบคือขั้นตอน กระบวนการคิดที่ผู้ตอบใช้ในการแก้ปัญหาในสถานการณ์การทดสอบนั้น ๆ ทฤษฎีการตอบสนองต่อ ข้อสอบ (IRT) จะให้เพียงการคาดหวังของผู้สำรวจแต่โมเดลวินิจฉัยทางพุทธิปัญญา (CDMs) มีพื้นฐาน บนหลักฐานเชิงประจักษ์ของกระบวนการและกลยุทธ์ที่ผู้ตอบต้องใช้ในสถานการณ์การทดสอบนั้น ๆ

ผู้วิจัยได้ทำการศึกษาเกี่ยวกับการจำแนกประเภทของโมเดลวินิจฉัยทางพุทธิปัญญา (CDMs) เพื่อเปรียบเทียบและพิจารณานำมาใช้ในการประเมินเชิงวินิจฉัยความสามารถการอ่านเพื่อ ความเข้าใจ พบว่า นักวิชาการจำแนกเป็น 2 ลักษณะ ได้แก่ แบบที่ 1 โมเดลชั้นแฝง (Latent Class Model) และโมเดลคุณลักษณะแฝง (Latent Trait Model) แบบที่ 2 แบบชดเชยได้ (Compensatory Model) และแบบไม่ชดเชย (Non-compensatory Model)

2.1 ประเภทของโมเดลวินิจฉัยทางพุทธิปัญญา

ไพทอร์ย์ เอื้อบุญประดิษฐ์ (2563, หน้า 33-41) อธิบายว่าประเภทของโมเดลวินิจฉัย ทางพุทธิปัญญา (CDMs) ที่ถูกพัฒนาขึ้นจากอดีตจนถึงปัจจุบันว่าถูกจำแนกเป็น 2 ประเภทหลัก คือ โมเดลชั้นแฝง (Latent Class Model) และโมเดลคุณลักษณะแฝง (Latent Trait Model) มี รายละเอียดดังนี้

1) โมเดลชั้นแฝง (Latent Class Model) เป็นโมเดลที่พัฒนาขึ้นเพื่อจำแนก ผู้สอบตามทักษะ (Skill) เป็นความรอบรู้ (Mastery) หรือความไม่รอบรู้ (Non-mastery) โมเดลนี้จึง ให้รูปแบบของความรอบรู้หรือความน่าจะเป็นของความรอบรู้เกี่ยวกับทักษะของผู้สอบ ตัวอย่างโมเดล วินิจฉัยทางพุทธิปัญญาที่เป็นโมเดลชั้นแฝงดังนี้

1.1 Rule-Space Model (RSM) (Tatsuoka, 1983, 1990, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) เป็นโมเดลทางคณิตศาสตร์ที่ Tatsuoka ได้พัฒนาขึ้นเพื่อใช้ จำแนกแบบแผนการตอบข้อสอบไปสู่แบบแผนความรู้ (Attribute-Mastery Pattern) จากการใช้ ทักษะทางพุทธิปัญญาที่ต่างกัน มีข้อตกลงเบื้องต้นว่าโมเดล ที่ได้จากข้อสอบอาจบรรยายได้ด้วย ทักษะทางพุทธิปัญญาเฉพาะหรือเรียกว่า คุณลักษณะ (Attribute) ซึ่งสามารถจำแนกวิธีการ ทักษะ หรือกระบวนการของผู้สอบได้ ดังนั้นคุณลักษณะจึงเปรียบเสมือนสถานะความรู้ของผู้สอบ (Latent Knowledge States) คุณลักษณะหรือความรู้ไม่สามารถสังเกตได้โดยตรง แต่แสดงให้เห็นได้จากการ ตอบข้อสอบในแบบทดสอบแนวคิดการประเมินกฎ (Rule-Space Method) ตรวจสอบระยะใกล้หรือ ไกลของสถานะความรู้ของผู้สอบกับแบบแผนความรู้ (Attribute-Mastery Pattern) พิจารณา จากแบบแผนการตอบข้อสอบ โดยให้ 0 คะแนนสำหรับข้อที่ตอบผิด หมายถึง ผู้สอบไม่มีความรอบรู้ ในสถานะความรู้นั้น และให้ 1 คะแนนสำหรับข้อที่ตอบถูก หมายถึง ผู้สอบมีความรอบรู้ในสถานะ ความรู้นั้น โดยนำหลักการของ Boolean Algebra มาใช้ในการกำหนดเมทริกซ์ของแบบแผนการ ตอบเพื่อใช้ในการคำนวณ อย่างไรก็ตาม Rule-Space Model มีข้อจำกัดคือ ผลการวินิจฉัยอาจไม่ สมบูรณ์ เนื่องจากดำเนินการวินิจฉัยหลังจากสร้างแบบสอบเรียบร้อยแล้ว จึงอาจทำให้การจำแนก คุณลักษณะที่ต้องใช้ในการตอบข้อสอบแต่ละข้อไม่ครอบคลุมคุณลักษณะที่ผู้สอบจะต้องมี นอกจากนี้ พบว่าเมื่อวิเคราะห์องค์ประกอบแล้วพบว่าคุณลักษณะไม่มีความสัมพันธ์เป็นลำดับขั้น ข้อจำกัด ดังกล่าวสวนทางกับงานวิจัยทางพุทธิปัญญาที่ชี้ให้เห็นว่าคุณลักษณะทางพุทธิปัญญาต้องสร้างให้เป็น ลำดับขั้น เนื่องจากทักษะทางพุทธิปัญญาไม่สามารถทำงานได้แบบเดี่ยว แต่เป็นเครือข่าย ความสัมพันธ์ของความสามารถ

1.2 Hybrid Model (Yamamoto, 1989, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญ ประดิษฐ์, 2563) เป็นโมเดลเกี่ยวกับการรวมระหว่างโมเดลการตอบสนองข้อสอบที่ใช้ทฤษฎีการ ตอบสนองข้อสอบเป็นฐาน (IRT Based) (ชั้นความผันแปรของความสามารถของแต่ละบุคคล) และ โมเดลการตอบสนองข้อสอบที่ใช้ชั้นแฝงเป็นฐาน (Latent Class Based) (Multinomial Independent Class) นอกจากนี้ Hybrid Model ยังใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional IRT Model) ในการอธิบายพฤติกรรมทั้งหมดของผู้สอบ กล่าวคือผู้สอบต้องใช้ หลายความสามารถในการทำข้อสอบแต่ละข้อ ต่อมาในปี ค.ศ. 1990 มีการขยาย Hybrid Model ให้ แสดงคุณลักษณะของผู้สอบแต่ละคน โดยเปลี่ยนจากกลยุทธ์การตอบที่ใช้ความสามารถเป็นฐาน (Ability-Based Response Ability) ไปเป็นกลยุทธ์การตอบอย่างสุ่ม

1.3 DINA Model (Deterministic-Input, Noisy-and-Gate Model)(Haertel, 1989; Junker and Sijtsma, 2001; De La Douglas, 2004, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563; วรางคณา ก้อนน้อย, 2563) เป็นโมเดลแบบ Conjunctive หรือ Non-Compensatory ใช้ในการ

จำแนกความรอบรู้ (Mastery) ของผู้สอบในข้อสอบแต่ละข้อเป็น 2 ชั้น ได้แก่ ชั้นที่ผู้สอบมีความรอบรู้ในทุกคุณลักษณะ และชั้นที่ผู้สอบไม่มีความรอบรู้อย่างน้อยหนึ่งคุณลักษณะ อธิบายด้วยโมเดลคณิตศาสตร์ที่อยู่ในรูปของฟังก์ชัน การตอบสนองข้อสอบ (Item Response Function) โดยพิจารณาโอกาสที่ผู้สอบจะตอบข้อสอบแต่ละข้อในชั้นแฝงแต่ละชั้น (Latent Class) ได้ถูกต้อง การวิเคราะห์ข้อมูลจะวิเคราะห์ความสอดคล้องระหว่างการออกแบบการประเมินเชิงวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostic Assessment) และข้อมูลการตอบข้อสอบ (Response Data) โดยจะประมาณค่าพารามิเตอร์ออกมา 2 พารามิเตอร์ ได้แก่ Slipping Parameter (s) และ Guessing Parameter (g) โดยข้อสอบแต่ละข้อจะมี Slipping Parameter และ Guessing Parameter อย่างละหนึ่งค่า

1.4 Unified Model (DiBello, Stout and Roussos, 1993) Unified Model เป็นโมเดลที่แสดงทั้งค่าพารามิเตอร์ของข้อสอบที่ขึ้นอยู่กับทักษะ (Skill-Based Item) และค่าพารามิเตอร์ของผู้สอบที่ขึ้นอยู่กับทักษะ (Skill-Based Examinee) เป็นโมเดลวินิจฉัยทักษะ (Skill-Diagnosis Model) โมเดลแรกที่น่าทักษะที่ไม่อยู่ใน Q-Matrix มารวมอยู่ในโมเดลด้วยข้อจำกัดของ Unified Model คือ บางพารามิเตอร์ในโมเดลไม่สามารถประมาณค่าด้วยวิธีการทางสถิติ จึงไม่เป็นที่นิยมนำมาวินิจฉัยแบบแผนการตอบข้อสอบของผู้เรียน (Hartz and Roussos, 2008)

1.5 NIDA Model (Noisy-Input, Deterministic-and-Gate Model) (Junker and Sijtsma, 2001, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) เนื่องจาก NIDA Model เป็นโมเดลที่กำหนดให้คุณลักษณะที่ต้องการวัดในข้อสอบ แต่ละข้อมี Slipping Parameter และ Guessing Parameter เท่ากัน แต่ Junker and Sijtsma มองว่าโอกาสที่จะตอบข้อสอบได้ถูกต้องนั้น ผู้สอบไม่จำเป็นต้องมีความรอบรู้ในคุณลักษณะที่ต้องการวัดเท่ากันเสมอไป จึงให้ความสำคัญกับคุณลักษณะที่ต้องการวัดในภาพรวมแทนการพิจารณาเป็นรายข้อ NIDA Model จะมีฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) เหมือนกับ DINA Model เนื่องจากเป็นโมเดลแบบ Non-Compensatory เหมือนกัน และพิจารณาถึงโอกาสที่ผู้สอบจะตอบข้อสอบ แต่ละข้อในชั้นแฝงแต่ละชั้นได้ถูกต้องเช่นเดียวกัน การวิเคราะห์ข้อมูลจะประมาณค่าพารามิเตอร์ออกมา 2 พารามิเตอร์ เช่นเดียวกับ DINA Model ได้แก่ Slipping Parameter (s) และ Guessing Parameter (g) แต่ต่างกันที่คุณลักษณะแต่ละคุณลักษณะจะมี Slipping Parameter และ Guessing Parameter อย่างละหนึ่งค่า

1.6 Fusion Model (FM) หรือ Reparameterized Unified Model (RUM) พัฒนาขึ้นโดย Hartz ในปี ค.ศ. 2002 เป็นโมเดลที่พัฒนามาจาก Unified Model โดยแก้ไขข้อจำกัดที่บางพารามิเตอร์ในโมเดลไม่สามารถประมาณค่าด้วยวิธีการทางสถิติ และเมื่อพิจารณาสมมติฐานของ NIDA Model ที่ว่าโอกาสที่ผู้สอบจะตอบข้อสอบได้ถูกต้องนั้น ผู้สอบจะต้องมีความรอบรู้ในทุกคุณลักษณะ Hartz จึงพัฒนา Fusion Model หรือ Reparameterized Unified Model ขึ้น เพื่อ

คล้ายข้อจำกัดของ DINA Model และ NIDA Model เป็นการรวมระหว่างการวินิจฉัยทักษะ (Skill Diagnosis) และการประมาณค่าพารามิเตอร์ในโมเดล การวิเคราะห์ข้อมูลจะประมาณค่าพารามิเตอร์ออกมา 4 พารามิเตอร์ ได้แก่ 1) Baseline Parameter หมายถึง โอกาสที่จะตอบข้อสอบ ได้ถูกต้อง เมื่อผู้สอบมีความรอบรู้ทุกคุณลักษณะที่ระบุใน Q-Matrix 2) Penalty Parameter หมายถึง ผลเสียของโอกาสที่จะตอบข้อสอบได้ถูกต้อง เมื่อผู้สอบไม่มีความรอบรู้ในคุณลักษณะของข้อสอบแต่ละข้อ 3) พารามิเตอร์ความยากของข้อสอบ (Difficulty Parameter) แต่ละข้อของคุณลักษณะทั้งหมดที่ไม่ได้ระบุใน Q-Matrix และ 4) พารามิเตอร์ความสามารถ (Ability Parameter) ในชั้นแฝงที่สัมพันธ์กับคุณลักษณะที่ไม่ได้ระบุใน Q-Matrix (Rupp, Templin and Henson, 2010, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563)

1.7 วิธีลำดับชั้นของคุณลักษณะ (Attribute Hierarchy Method: AHM) (Leighton, Gierl and Hunka, 2004, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) เป็นโมเดลทฤษฎีการตอบสนองข้อสอบทางพุทธิปัญญาที่พัฒนาขึ้นเพื่อแก้ข้อจำกัดของ Rule-Space Model โดยมีวัตถุประสงค์เพื่อให้การประเมินเชิงวินิจฉัยมีความชัดเจนเป็นประโยชน์ต่อการพัฒนาการจัดการเรียนรู้และพัฒนาผู้เรียนมากขึ้น เป็นวิธีการกำหนดจำนวนข้อสอบ ลักษณะข้อสอบ และแบบแผนการตอบข้อสอบให้ถูกต้องเรียกว่า คุณลักษณะ (Attribute) ซึ่งแสดงให้เห็นถึงประโยชน์ของการเชื่อมโยงระหว่างทฤษฎีทางพุทธิปัญญา (Cognitive Theory) กับการวัดทางจิตมิติ (Psychometric Practice) ปัจจุบันมีการใช้วิธีลำดับชั้นของคุณลักษณะในการสร้างแบบทดสอบเพื่อการประเมินเชิงวินิจฉัยทางพุทธิปัญญาโดยสร้างตามโครงสร้างของลำดับชั้น 4 รูปแบบ ดังนี้ 1) ลำดับชั้นเชิงเส้น (Linear Hierarchy) เป็นความสัมพันธ์ระหว่างคุณลักษณะ ที่มีลักษณะเป็นเส้นตรงที่มีจุดสิ้นสุดจุดเดียว 2) ลำดับชั้นเชิงลู่ออก (Hierarchy with a Divergent Branch) เป็นความสัมพันธ์ที่แยกออกเป็น 2 ทาง แต่มีจุดสิ้นสุดจุดเดียวกัน 3) ลำดับชั้นเชิงลู่ออก (Hierarchy with a Divergent Branch) เป็นความสัมพันธ์ ที่แยกเป็น 2 ทางและไม่ได้สิ้นสุดที่จุดเดียวกัน 4) ลำดับชั้นแบบไม่มีโครงสร้าง (Unstructured Hierarchy) เป็นความสัมพันธ์ที่ไม่ได้มีความสัมพันธ์จากจุดเริ่มต้นเดียว และไม่ได้มีจุดสิ้นสุดจุดเดียวกัน ลักษณะสำคัญที่เป็นจุดเด่นของวิธีลำดับชั้นของคุณลักษณะ คือ การให้คำจำกัดความของค่าว่าคุณลักษณะที่ชัดเจนและกำหนดให้จำแนกคุณลักษณะก่อนการพัฒนาแบบสอบ รวมทั้งการกำหนดลักษณะความสัมพันธ์ที่เป็นลำดับชั้นของคุณลักษณะซึ่งเป็นขั้นตอนสำคัญในการช่วยกำหนดลักษณะของข้อสอบแต่ละข้อ ทำให้การวินิจฉัยจากแบบแผนการตอบข้อสอบนั้นให้สารสนเทศที่เป็นประโยชน์ และชัดเจนมากยิ่งขึ้น อย่างไรก็ตาม วิธีลำดับชั้นของคุณลักษณะมีข้อจำกัดในกรณีที่แบบแผนการตอบ ข้อสอบที่สังเกตได้ไม่ตรงกับแบบแผนการตอบข้อสอบที่คาดหวังใดเลย อาจทำให้การจำแนกแบบแผนการตอบที่สังเกตได้ขาดความชัดเจน

1.8 Deterministic-Input, Noisy-or-Gate Model (DINO Model) (Templin and Henson, 2006, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) DINO Model มีฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) คล้ายกับ DINA Model แต่ต่างกันที่ DINO Model เป็นโมเดลแบบชดเชยได้ (Compensatory Model) นิยมใช้ในการวินิจฉัยผู้ป่วยที่มีอาการผิดปกติทางจิต (Psychological Disorder Diagnosis) ตั้งอยู่บนสมมติฐานที่ว่าโอกาสที่ผู้ป่วยจะมีอาการผิดปกติทางจิตนั้น ผู้ป่วยจะต้องมีคุณลักษณะใดคุณลักษณะหนึ่งอย่างน้อย 1 คุณลักษณะ หรือในแง่โมเดลคณิตศาสตร์หมายถึง โอกาสที่ผู้สอบจะตอบข้อสอบในทางบวก (Positive Response) หรือผู้สอบจะตอบข้อสอบแต่ละข้อในชั้นแฝงแต่ละชั้นได้ถูกต้อง การวิเคราะห์ข้อมูลจะประมาณค่าพารามิเตอร์ออกมา 2 พารามิเตอร์ เช่นเดียวกับ DINA Model ได้แก่ Slipping Parameter (s) ซึ่งหมายถึง โอกาสที่ผู้สอบจะตอบข้อสอบในทางลบ (Negative Response) หรือตอบผิด เมื่อมีคุณลักษณะปรากฏขึ้นอย่างน้อย 1 คุณลักษณะ หรือผู้ป่วยมีความผิดปกติทางจิตแต่เมื่อทำแบบทดสอบแล้วพบว่าไม่ผิดปกติและ Guessing Parameter (g) ซึ่งหมายถึง โอกาสที่ผู้สอบจะตอบข้อสอบในทางบวก (Positive Response) หรือตอบถูก เมื่อไม่มีคุณลักษณะใดปรากฏขึ้นหรือผู้ป่วยไม่มีความผิดปกติทางจิตแต่เมื่อทำแบบทดสอบแล้วพบว่าผิดปกติ โดยข้อสอบแต่ละข้อจะมี Slipping Parameter และ Guessing Parameter อย่างละหนึ่งค่า เหมือนกับ DINA Model

1.9 Log-Linear Cognitive Diagnostic Model (LCDM) (Henson, Templin and Willse, 2009, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) LCDM เป็นโมเดลเชิงสรุปอ้างอิง (Generalized Model) ที่ขยายแนวคิดมาจากโมเดลแบบชดเชยได้ (Compensatory) และโมเดลแบบไม่ชดเชย (Non-Compensatory) ด้วยการนำโมเดลที่อยู่ในกลุ่ม Compensatory และ Non-Compensatory มาวิเคราะห์ร่วมกัน เช่น นำ DINA Model มาวิเคราะห์ร่วมกับ DINO Model ในแบบทดสอบฉบับเดียวกัน เนื่องจากข้อสอบบางข้ออาจเหมาะกับ Deterministic-Input, Noisy-and-Gate Model ในขณะที่ข้อสอบบางข้ออาจเหมาะกับ DINO Model ด้วยเหตุนี้ Log-Linear Cognitive Diagnostic Model จึงเหมาะสำหรับนำมาใช้ในการสรุปอ้างอิงโดยนัยทั่วไปสำหรับโมเดลเชิงวินิจฉัยทางพุทธิปัญญา ข้อจำกัดของโมเดลนี้คือ ค่าพารามิเตอร์ในโมเดลมีหลายค่า และมีความซับซ้อน รวมทั้งการวิเคราะห์ข้อมูลมีหลายขั้นตอนการนำโมเดลเชิงวินิจฉัยทางพุทธิปัญญา หลายโมเดลมาวิเคราะห์ร่วมกัน ทำให้ไม่สะดวกในการนำไปประยุกต์ใช้

2) โมเดลคุณลักษณะแฝง (Latent Trait Model) เป็นการขยายแนวคิดหรือสรุปอ้างอิงจากโมเดลการตอบสนองข้อสอบแบบเอกมิติ (Unidimensional IRT Model) เกี่ยวกับการจัดตำแหน่งการประมาณค่าความสามารถของผู้สอบในแต่ละทักษะ โมเดลวินิจฉัยทางพุทธิปัญญาที่เป็นโมเดลคุณลักษณะแฝง มีดังนี้

2.1 Linear Logistic Test Model (LLTM) (Fisher, 1973, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) Linear Logistic Test Model เป็นโมเดลทางจิตมิติโมเดลแรกๆ ที่เชื่อมโยงระหว่างจิตมิติ ทางพุทธิปัญญา (Cognitive Psychometric) และการออกแบบข้อสอบ (Item Design) เป็นโมเดล ที่พัฒนามาจาก Rasch Model ที่มีแต่พารามิเตอร์ความยากของข้อสอบ Linear Logistic Test Model ถูกพัฒนาโดยการนำเนื้อหาข้อสอบ (Item Content) เข้าไปอยู่ในการ ทำนายความสำเร็จของข้อสอบ (Item Success) ดังนั้น ถ้าองค์ประกอบด้านเนื้อหาที่ถูกระบุมีความ เหมาะสม พารามิเตอร์จะสะท้อนผลไปที่ค่าความยากของข้อสอบที่ถูกประมาณค่าได้โดยตรง โดย องค์ประกอบด้านเนื้อหาแต่ละองค์ประกอบจะถูกกำหนดค่าเป็น 0-1 เมื่อ 0 หมายถึง มีองค์ประกอบ ที่ต้องการวัดอยู่ในข้อสอบและ 1 หมายถึง ไม่มีองค์ประกอบที่ต้องการวัดอยู่ในข้อสอบ การประมาณ ค่าความยากของข้อสอบจะประมาณจากผลรวมน้ำหนักของคะแนนคุณลักษณะในโมเดลความ ซ้ำซ้อนทางพุทธิปัญญา ซึ่งต่างจากค่าความยากของข้อสอบใน Rasch Model ที่ไม่ได้คำนึงถึงตัวแปร ทางพุทธิปัญญา

2.2 Multicomponent Latent Trait Model (MLTM) (Embretson, 1980, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) MLTM เป็นโมเดลที่ใช้วัดองค์ประกอบของการ ดำเนินการหลายองค์ประกอบที่อยู่ภายใต้การแก้ปัญหาข้อสอบ โดยจะประมาณค่าระดับของ คุณลักษณะแฝงในองค์ประกอบ (Component Trait Level) และค่าความยากของข้อสอบใน องค์ประกอบนั้น (Component Item Difficulty) ตั้งอยู่บนสมมติฐานว่าความสำเร็จข้อสอบ (Item Success) ขึ้นอยู่กับความสำเร็จของหลายองค์ประกอบ กล่าวคือ ถ้าองค์ประกอบใดองค์ประกอบหนึ่ง ล้มเหลวจะทำให้ไม่สามารถแก้ปัญหา ข้อสอบข้อนั้นได้ Multicomponent Latent Trait Model อยู่ในรูปของผลคุณความน่าจะเป็นขององค์ประกอบที่ให้ความน่าจะเป็นในภาพรวมของความสำเร็จใน การทำข้อสอบ โมเดลนี้ได้นำพารามิเตอร์การเดาข้อสอบได้ถูก (Guessing Parameter) มาใช้ในการ ประมาณค่าในโมเดล ทั้งนี้ MLTM เหมาะสำหรับคุณลักษณะที่ต้องมีมาก่อน เนื่องจากโมเดลทางพุทธิ ปัญญาของข้อสอบทั่วไปขึ้นอยู่กับความสลับไหลของข้อมูลจากสถานะหนึ่งไปยังอีกสถานะหนึ่ง นอกจากนี้ยังเป็นโมเดลที่เหมาะสมสำหรับการวัดแบบพหุมิติซึ่งขึ้นอยู่กับผลคุณที่ต่อเนื่องของความน่าจะเป็นของการดำเนินการนั้นคือ ผลคุณของโอกาสที่ผู้สอบจะตอบข้อสอบได้ถูกต้องในแต่ละ องค์ประกอบ โดยผู้ตอบจะต้องตอบองค์ประกอบที่มีมาก่อนให้ถูกต้อง MLTM ได้รวมโมเดลผลคุณที่ ต่อเนื่องของกระบวนการ ตอบข้อสอบกับ IRT Model เข้าด้วยกันเพื่อสะท้อนความแตกต่าง รายบุคคลในระดับของคุณลักษณะในองค์ประกอบนั้น ดังนั้น MLTM จึงประมาณค่าทั้งระดับของ คุณลักษณะในองค์ประกอบและความยากของข้อสอบในองค์ประกอบ ซึ่งโอกาสของการตอบข้อสอบ ในแต่ละองค์ประกอบขึ้นอยู่กับความยากขององค์ประกอบในข้อสอบและระดับของคุณลักษณะแฝงใน องค์ประกอบของบุคคล

2.3 Compensatory Multidimensional IRT Model (McKinley and Reckase, 1982, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) เป็นโมเดลการตอบสนองข้อสอบแบบพหุมิติแบบชดเชยได้ (Compensatory) มีจุดเด่น คือ ในการตอบข้อสอบแต่ละข้อใช้คุณลักษณะแฝงหรือความสามารถ (Ability) มากกว่า 1 คุณลักษณะ ถ้าความสามารถของผู้สอบในมิติใดมิติหนึ่งมีค่าลดลงแล้วความสามารถที่มีค่าสูงกว่าจะเข้ามาแทนที่ความสามารถที่มีค่าต่ำกว่า การวิเคราะห์ข้อมูลจะประมาณค่าพารามิเตอร์ออกมา 3 พารามิเตอร์ ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าโอกาสการเดาข้อสอบได้ถูก โดยโอกาสที่ผู้สอบจะตอบข้อสอบได้ถูกต้องขึ้นอยู่กับค่าความยากของข้อสอบซึ่งเป็นผลรวมความสามารถในแต่ละมิติ

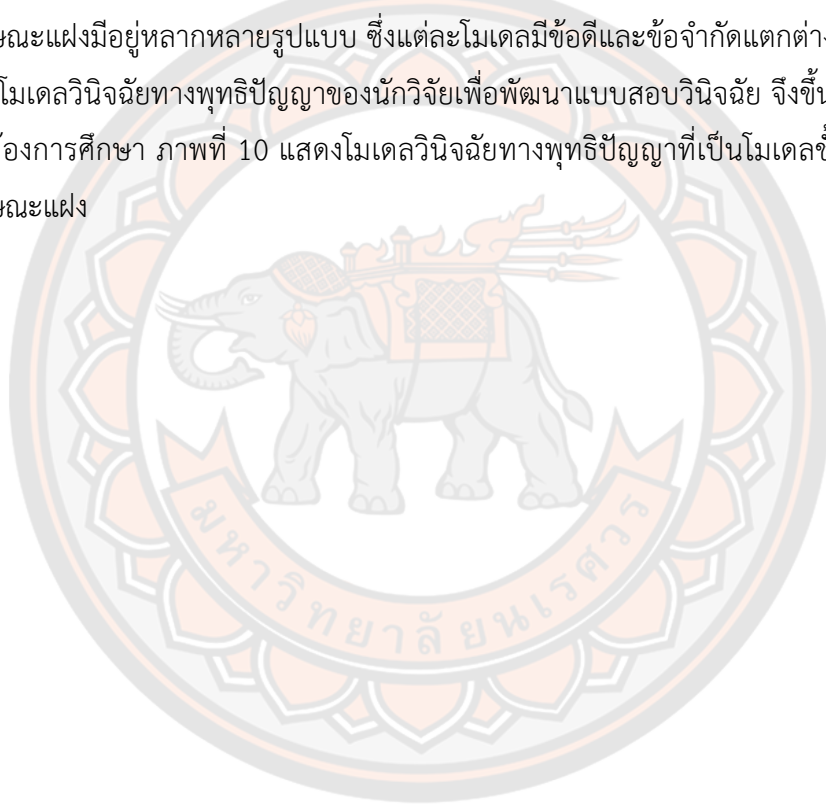
2.4 General Component Latent Trait Model (GLTM) (Embretson, 1984, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) GLTM เป็นการนำ Linear Logistic Test Model (LLTM) และ Multicomponent Latent Trait Model (MLTM) มารวมกัน ใช้ทดสอบสมมติฐานเกี่ยวกับตัวแปรทางทฤษฎีซึ่งอยู่ภายใต้ความยากของการตอบข้อสอบ (Response Difficulty) และการประมาณค่าพารามิเตอร์ ของข้อสอบ รวมทั้งนำโมเดลคณิตศาสตร์มาสร้างความสัมพันธ์ระหว่างค่าความยากของข้อสอบในองค์ประกอบนั้น ๆ กับคุณลักษณะด้านเนื้อหาของข้อสอบ ดังนั้น ค่าความยากของข้อสอบในโมเดลจึงเหมือนกับ LLTM และมีโมเดลการประมาณค่าเหมือนกับ MLTM

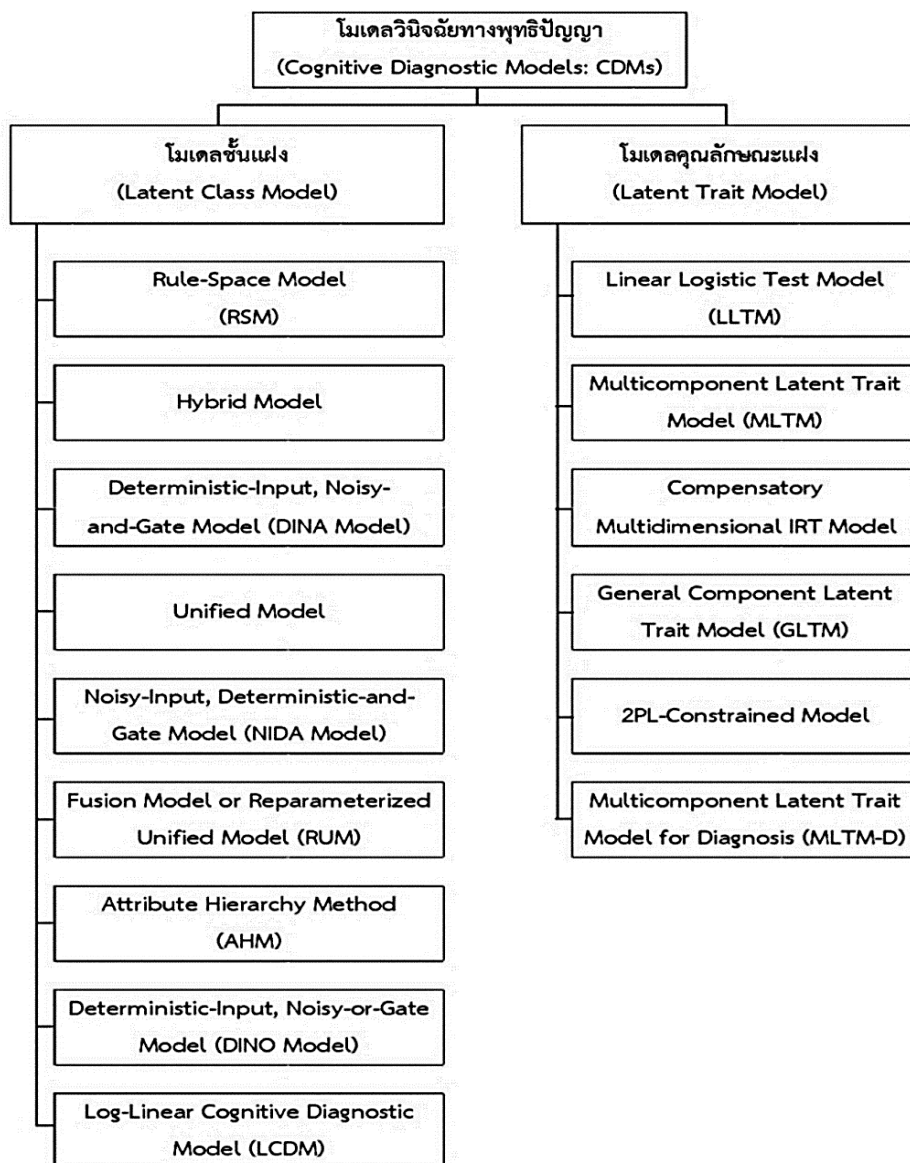
2.5 2PL-Constrained Model (Embretson, 1999, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) Embretson ได้พัฒนาโมเดลการสรุปอ้างอิงของโมเดลโลจิสติกแบบ 2 พารามิเตอร์ เรียกว่า 2PL-Constrained Model เป็นการประมาณค่าพารามิเตอร์ความยากของข้อสอบ (Item Difficulty) และค่าอำนาจจำแนกของข้อสอบ (Item Discrimination) โดยค่าความยากของข้อสอบจะเหมือนกับค่าความยากของข้อสอบใน Linear Logistic Test Model (LLTM) และค่าอำนาจจำแนกของข้อสอบในโมเดลนี้จะถูกกำหนดให้เป็นผลรวมเชิงเส้นของคะแนนคุณลักษณะกับน้ำหนักของคุณลักษณะนั้นที่มีผลต่อค่าอำนาจจำแนกของข้อสอบ

2.6 Multicomponent Latent Trait Model for Diagnosis (MLTM-D) (Embretson and Yang, 2008, อ้างถึงใน ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563) Multicomponent Latent Trait Model for Diagnosis เป็นโมเดลที่ออกแบบมาเพื่อใช้วินิจฉัยทักษะในลำดับชั้นของเนื้อหาที่จัดเรียงในข้อสอบโดยใช้วิธีลดหลั่น (Hierarchical Approach) ในการวินิจฉัยเอ็มเบรทสัน (Embretson & Yang, 2013) อธิบายว่า MLTM-D เป็นโมเดลคุณลักษณะแฝงที่ไม่มีการชดเชย (Non-compensatory) ซึ่งเป็นโมเดลคุณลักษณะแฝงแบบหลายองค์ประกอบสำหรับการวินิจฉัยสำหรับการวินิจฉัยทางพุทธิปัญญาใน MLTM-D ความสัมพันธ์แบบลำดับชั้นระหว่างส่วนประกอบและคุณลักษณะได้รับการระบุเพื่อให้สามารถนำไปใช้ในการวินิจฉัยได้ในสองระดับ MLTM-D เป็นลักษณะ

ทั่วไปของแบบจำลองลักษณะแฝงที่มีองค์ประกอบหลายองค์ประกอบเพื่อนำไปใช้กับการวัดลักษณะกว้าง ๆ เช่น การทดสอบวัดผลสัมฤทธิ์ ซึ่งโครงสร้างส่วนประกอบจะแตกต่างกันไปในแต่ละรายการ มีการอธิบายเงื่อนไขสำหรับการระบุแบบจำลองและนำเสนอตัวประมาณความน่าจะเป็นสูงสุดส่วนเพิ่มพร้อมด้วยข้อมูลการจำลองเพื่อแสดงการกู้คืนพารามิเตอร์เพื่อแสดงให้เห็นว่า MLTM-D สามารถใช้ในการวินิจฉัยได้อย่างไร จึงได้นำเสนอแอปพลิเคชันสำหรับการทดสอบความสำเร็จทางคณิตศาสตร์ในวงกว้าง ข้อดีของ MLTM-D สำหรับการวินิจฉัยคืออาจใช้ได้กับการประเมินขนาดใหญ่ที่มีรายการต่างกันมากกว่าโมเดลชั้นแฝง

จากที่กล่าวมาโมเดลวินิจฉัยทางพุทธิปัญญาทั้งประเภทโมเดลชั้นแฝง และโมเดลคุณลักษณะแฝงมีอยู่หลากหลายรูปแบบ ซึ่งแต่ละโมเดลมีข้อดีและข้อจำกัดแตกต่างกัน ดังนั้นในการเลือกใช้โมเดลวินิจฉัยทางพุทธิปัญญาของนักวิจัยเพื่อพัฒนาแบบสอบวินิจฉัย จึงขึ้นอยู่กับบริบทของเรื่องที่ต้องการศึกษา ภาพที่ 10 แสดงโมเดลวินิจฉัยทางพุทธิปัญญาที่เป็นโมเดลชั้นแฝงและโมเดลคุณลักษณะแฝง





ภาพ 10 โมเดลวินิจฉัยทางพุทธิปัญญา
(ไพทอร์ย์ เอื้ออื้อบุญประดิษฐ์, 2563, หน้า 41)

โมเดลวินิจฉัยทางพุทธิปัญญา (CDMs) เป็นวิธีการแบบสหวิทยาการ (Interdisciplinary Approach) ใช้ในการประเมินเชิงวินิจฉัย โดยเชื่อมโยงระหว่างจิตวิทยาพุทธิปัญญาและการวิเคราะห์ทางสถิติไว้ด้วยกัน เพื่อใช้สำรวจความสัมพันธ์ระหว่างกระบวนการทางจิตและกลวิธีที่ใช้ในการแก้ปัญหาข้อหนึ่ง ๆ ในแบบทดสอบและการตอบสนองของผู้ตอบ ผ่านการวิเคราะห์ทางสถิติที่ซับซ้อนมีลักษณะเป็นโมเดลชั้นแฝง (Latent Class Model) ซึ่งจำแนกผู้ตอบแบบทดสอบเป็นชั้นแฝง (Latent Classes) ขึ้นอยู่กับการตอบสนองต่อข้อสอบของผู้ตอบโดยชั้นแฝง

นั้นเป็นชั้นแฝงที่มีจำนวนจำกัด (Restricted Latent Classes) ขึ้นอยู่กับจำนวนของคุณลักษณะ (Attributes) ที่ใช้ในการตอบข้อสอบของแบบทดสอบ หากแบบทดสอบนั้นมีคุณลักษณะที่ต้องใช้จำนวน K คุณลักษณะ ผู้ตอบจะถูกจำแนกเป็น 2^K ชั้นแฝง โดยเลข 2 หมายถึง จำนวนของผลลัพธ์ที่เป็นไปได้สำหรับแต่ละคุณลักษณะซึ่งหมายถึง การรอบรู้ (Mastery) หรือไม่รอบรู้ (Non-mastery) ในคุณลักษณะนั้น ๆ (Ravand and Robitzsch, 2015, pp. 1-2, อ้างถึงใน วราจคณา ก้อมน้อย, 2563, หน้า 62) การคำนวณชั้นแฝงจะทำให้ทราบจำนวนโปรไฟล์แบบแผนการตอบที่เป็นไปได้ทั้งหมด ตัวอย่างเช่น แบบสอบชุดหนึ่งต้องใช้คุณลักษณะในการตอบแบบทดสอบจำนวน 3 คุณลักษณะ 2^K แทนค่า ($K=3$) จะได้ $2^3 = 8$ นั่นคือ ผู้ตอบจะถูกจำแนกเป็น 8 ชั้นแฝง โดยที่ 1 หมายถึง รอบรู้ และ 0 หมายถึง ไม่รอบรู้ แสดงในภาพ 11

$$\alpha = \{(0,0,0) (0,0,1) (0,1,0) (1,0,0) (0,1,1) (1,0,1) (1,1,0) (1,1,1)\}$$

ภาพ 11 ชั้นแฝงสำหรับการวินิจฉัย 3 คุณลักษณะ

จากภาพ 11 แสดงให้เห็นจำนวนชั้นแฝง (Latent class) ของรูปแบบการเรียนรู้ที่อาจเกิดขึ้นจากการวินิจฉัยซึ่งผู้สอบจะถูกจัดเข้าชั้นแฝงชุดใดชั้นหนึ่ง ขึ้นอยู่กับความรอบรู้ของผู้สอบว่ามีความสามารถตามคุณลักษณะที่วัดจากแบบสอบชุดนั้นๆหรือไม่ (เอกราช ดินาง, 2559, หน้า 4) ดังนั้น จึงสามารถนำมาใช้เพื่อประเมินจุดแข็ง (Strength) และจุดอ่อน (Weakness) ของนักเรียนในแต่ละทักษะ (Skills) หรือคุณลักษณะ (Attributes) ที่จำเป็นต้องใช้ในการตอบแบบสอบได้ สารสนเทศจากการสอบจึงให้ทั้งข้อมูลการวัดการเรียนรู้ของผู้เรียนและพัฒนาการของผู้เรียนใช้ปรับปรุงการเรียนการสอน และพัฒนาโครงการเพื่อแก้ไขปัญหาของนักเรียนบางคนและบางกลุ่ม (สังวรณั ้งดกระโทก, 2559) กรณีแบบสอบมีคุณลักษณะที่ต้องใช้ในการตอบแบบสอบจำนวน 3 คุณลักษณะ ตาราง 8 แสดงชั้นแฝง (Latent Class) ที่ผู้สอบจะถูกจำแนกเข้าไปในรูปแบบที่คาดหวัง

ตาราง 8 แสดงชั้นแฝงที่ผู้สอบจะถูกจำแนกในกรณีมี 3 คุณลักษณะที่ต้องใช้ในการตอบแบบสอบ

ชั้นแฝงที่ (รูปแบบ)	คุณลักษณะที่ 1	คุณลักษณะที่ 2	คุณลักษณะที่ 3	ความหมาย
1 (0,0,0)	ไม่รอบรู้	ไม่รอบรู้	ไม่รอบรู้	ผู้สอบไม่รอบรู้ทุกคุณลักษณะ
2 (1,0,0)	รอบรู้	ไม่รอบรู้	ไม่รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะที่ 1
3 (0,1,0)	ไม่รอบรู้	รอบรู้	ไม่รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะที่ 2
4 (0,0,1)	ไม่รอบรู้	ไม่รอบรู้	รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะที่ 3
5 (1,1,0)	รอบรู้	รอบรู้	ไม่รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะที่ 1 และ

ชั้นแฝงที่ (รูปแบบ)	คุณลักษณะที่ 1	คุณลักษณะที่ 2	คุณลักษณะที่ 3	ความหมาย
				2
6 (1,0,1)	รอบรู้	ไม่รอบรู้	รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะที่ 1 และ 3
7 (0,1,1)	ไม่รอบรู้	รอบรู้	รอบรู้	ผู้สอบรอบรู้ในคุณลักษณะที่ 2 และ 3
8 (1,1,1)	รอบรู้	รอบรู้	รอบรู้	ผู้สอบรอบรู้ในทุกคุณลักษณะ

จากตาราง 8 โมเดลวินิจฉัยทางพุทธิปัญญา (CDM) คือ โมเดลการวัดทางจิตวิทยาที่ใช้ประเมินจุดแข็งและจุดอ่อนของนักเรียน ซึ่งถูกพัฒนาขึ้นเพื่อใช้ระบุความรู้หรือไม่รอบรู้ในทักษะ (skills) หรือคุณลักษณะ (attributes) ที่จำเป็นต้องใช้ในการตอบแบบสอบของนักเรียน ข้อมูลจากการสอบให้ทั้งข้อมูลการวัดการเรียนรู้ของผู้เรียนและพัฒนาการของผู้เรียนใช้ปรับปรุงการเรียนการสอน และพัฒนาโครงการเพื่อแก้ไขปัญหาของนักเรียนบางคนและบางกลุ่ม (สังวรณั ังตกระโทก, 2556, หน้า 36)

การจัดกลุ่มของโมเดลการวินิจฉัยเชิงจำแนก (A Taxonomy of DCMs)

Rupp, Templin และ Henson (2010, 97, อ้างถึงใน สุมาลี มีสกุล, 2558, หน้า 39-40) ได้จัดกลุ่มของโมเดลการวินิจฉัยเชิงจำแนก (DCMs) โดยใช้เกณฑ์การพิจารณา 3 ข้อ คือ (1) ระดับการวัดตัวแปรสังเกตได้ (การตรวจให้คะแนนแบบสองค่ากับการตรวจการตรวจให้คะแนนแบบหลายค่า) (2) ระดับการวัดตัวแปรแฝงที่ต้องการจำแนก (การจำแนกตัวแปรแฝงแบบสองค่ากับการจำแนกตัวแปรแฝงแบบหลายค่า) และ (3) ข้อตกลงสำหรับการประกอบกันของคุณลักษณะตัวแปรแฝงที่ต้องการทำนาย สามารถจำแนกประเภทได้ดังตาราง 9

ตาราง 9 การจัดกลุ่มของโมเดลการวินิจฉัยเชิงจำแนก (DCMs) (Rupp and Templin, 2007, p.56)

ระดับการวัดตัวแปร สังเกตได้	ระดับของตัวแปรแฝงที่ต้องการจำแนก		ประเภทของโมเดล
	แบบสองค่า	แบบหลายค่า	
แบบสองค่า	RSM		Non-compensatory
	AHM		เป็นคุณลักษณะที่
	DINA		ไม่สามารถชดเชยได้ด้วยคุณลักษณะอื่น
	HO-DINA		
	MS-DINA		
	NIDA		
	RERUM		

ระดับการวัดตัวแปร	ระดับของตัวแปรแฝงที่ต้องการจำแนก		ประเภทของโมเดล
	BIN	BIN	
	MCLCM	MCLCM	
	Full NC-RUM	Full NC-RUM	
	Reduced NC-RUM	Reduced NC-RUM	

จากตาราง 9 ชื่อเต็มของโมเดลมีรายละเอียดดังนี้

RSM	: Rule-space method
AHM	: Attribute hierarchy method
DINA	: Deterministic input, noisy "and" gate
HO-DINA	: Higher order DINA
MS-DINA	: Multi-strategy DINA
NIDA	: Noisy input, deterministic "and" gate
RERUM	: Random-effects RUM
BIN	: Bayesian inference network
MCLCM	: Multiple classification latent class model
Full NC-RUM	: NC-RUM with continuous latent interaction term
Reduced NC-RUM	: NC-RUM without latent interaction term

ตาราง 9 การจัดกลุ่มของโมเดลการวินิจฉัยเชิงจำแนก (DCMs) (ต่อ)

ระดับการวัดตัวแปร	ระดับของตัวแปรแฝงที่ต้องการจำแนก		ประเภทของโมเดล
	แบบสองค่า	แบบหลายค่า	
สังเกตได้	DINO		Compensatory
	NIDO		เป็นคุณลักษณะที่สามารถชดเชยได้ด้วย
	BIN	BIN	คุณลักษณะอื่น
	MCLCM	MCLCM	
	C-RUM	C-RUM	
	GDM	GDM	
	H-GDM	H-GDM	
	LDCM	LDCM	
	G-DINA	G-DINA	
แบบหลายค่า	RSM		Non-compensatory
	AHM		เป็นคุณลักษณะที่ไม่สามารถชดเชยได้ด้วย
	BIN	BIN	

ระดับการวัดตัวแปร	ระดับของตัวแปรแฝงที่ต้องการจำแนก		ประเภทของโมเดล
	MCLCM	MCLCM	
	Full NC-RUM	Full NC-RUM	คุณลักษณะอื่น
	Reduced NC-RUM	Reduced NC-RUM	
	BIN	BIN	Compensatory
	MCLCM	MCLCM	เป็นคุณลักษณะที่
	C-RUM	C-RUM	สามารถชดเชยได้ด้วย
	GDM	GDM	คุณลักษณะอื่น
	H-GDM	H-GDM	
	LDCM	LDCM	
	G-DINA	G-DINA	

จากตาราง 9 ชื่อเต็มของโมเดลมีรายละเอียดดังนี้

DINO	: Deterministic input, noisy "or" gate
NIDO	: Noisy input, deterministic "or" gate
BIN	: Bayesian inference network
MCLCM	: Multiple classification latent class model
C-RUM	: Compensatory RUM
GDM	: General diagnostic model
H-GDM	: Hierarchical GDM
LDCM	: Log-linear cognitive diagnosis model
G-DINA	: Generalized DINA
RUM	: Reparametrized unified model/Fusion model
NC-RUM	: Non-compensatory RUM
RERUM	: Random-effects RUM
Full NC-RUM	: NC-RUM with continuous latent interaction term
Reduced NC-RUM	: NC-RUM without latent interaction term

การจัดกลุ่มของโมเดลการวินิจฉัยเชิงจำแนก ตามที่แสดงในตาราง 9 มีประเด็นสำคัญ 5 ประเด็น (Rupp, Templin & Henson, 2010, pp. 97, 99-101, อ้างถึงใน สุมาลี มีสกุล, 2558, หน้า 41) ดังนี้

1) โมเดลการวินิจฉัยเชิงจำแนก (DCMs) ทุกโมเดลสามารถนำไปใช้กับ ข้อมูลที่นำมาวิเคราะห์ซึ่งมีการตรวจให้คะแนนแบบสองค่า (เช่น การตรวจให้คะแนนแบบ ถูก-ผิด) และการจำแนก

ตัวแปรแฝงที่ได้จากการวินิจฉัยแบบสองค่า (เช่น การจำแนกผู้สอบออกเป็น ผู้รอบรู้ ผู้ไม่รอบรู้) ทุกโมเดล

2) โมเดลบางโมเดลในตาราง เช่น Multiple Classification Latent Class Models (MCLCMs) กับ Bayesian Inference Networks (BINs) จะปรากฏอยู่ร่วมกันในทุกประเภทของกลุ่มโมเดลที่เป็น non-compensatory model เนื่องจากทั้งสองโมเดลนี้เป็นโมเดลที่มีข้อจำกัดในการประมาณค่าน้อยกว่าโมเดลอื่น กล่าวคือสามารถใช้ได้โดยทั่วไปกับทั้งข้อมูลที่น่ามาวิเคราะห์ ซึ่งมีการตรวจให้คะแนนแบบสองค่าและแบบหลายค่า และสามารถจำแนกตัวแปรแฝงที่ได้จากการวินิจฉัยได้ทั้งแบบสองค่าและแบบหลายค่าในกลุ่มโมเดลที่เป็น Non-compensatory model

3) ข้อมูลที่นำมาใช้ในการวิเคราะห์และตัวแปรแฝงที่ได้จากการวินิจฉัยจะไม่ใช่ข้อมูลที่อยู่ในรูปของตัวแปรต่อเนื่อง เนื่องจากตัวแปรแฝงแบบต่อเนื่องจะไม่นำพาไปสู่การจำแนกทางสถิติโดยตรง ถึงแม้ว่าจะมีโมเดลการวินิจฉัยเชิงจำแนกบางโมเดลที่เอื้อต่อการรวมกัน ระหว่างข้อมูลตัวแปรแฝงแบบจัดประเภทและข้อมูลแบบตัวแปรแฝงต่อเนื่อง เช่น Ful NC-RUM

4) การจัดประเภทของโมเดลในตารางไม่เพียงพอต่อการที่จะแสดงให้เห็นความแตกต่างระหว่างรายละเอียดของโมเดลย่อยแต่ละโมเดลภายในกลุ่ม ได้รายละเอียดความแตกต่างของแต่ละโมเดลย่อยในกลุ่มนี้จะกล่าวถึงในหัวข้อถัดไป

5) มีโมเดลในตาราง 2 โมเดล คือ Rule-space methodology (RSM) และ Attribute hierarchy method (AHM) ที่แตกต่างจากโมเดลอื่นในเรื่องของวิธีที่ใช้ในการประมาณค่าเนื่องจากทั้งสองโมเดลมีวิธีการทางสถิติที่ใช้ในการจำแนกซึ่งไม่ได้ใช้โครงสร้างโมเดลความน่าจะเป็นที่สมบูรณ์เหมือนกับโมเดลอื่น

เมื่อจำแนกโมเดลวินิจฉัยทางพุทธิปัญญาตามประเภทของโมเดลแบบชดเชยได้ (Compensatory Models) และโมเดลแบบไม่ชดเชย (Non-compensatory Models) สามารถแบ่งได้เป็น 3 กลุ่ม มีรายละเอียดแสดงในตาราง 10 (Ravand and Robitzsch, 2015, p. 3, อ้างถึงใน วรางคณา ก้อมน้อย, 2563, หน้า 66) ดังนี้

ตาราง 10 ประเภทของโมเดลเชิงวินิจฉัยทางพุทธิปัญญาแบบจำแนกเป็น 3 กลุ่ม

ประเภทของโมเดลวินิจฉัยทางพุทธิปัญญา	ตัวอย่างโมเดล	อ้างอิง
Specific 1) โมเดลแบบชดเชยได้ (Compensatory Models) มีลักษณะคือ ผู้สอบไม่จำเป็นต้องมีความรอบรู้ครบทุกคุณลักษณะ (attributes) ที่จำเป็นต้องใช้ในข้อนั้น ๆ สามารถทดแทนหรือ	1. Deterministic input, noisy-or-gate model (DINO)	Templin & Henson (2006)
	2. Compensatory	Hartz

ประเภทของโมเดลวินิจฉัยทางพุทธิปัญญา	ตัวอย่างโมเดล	อ้างอิง
ชดเชยด้วยคุณลักษณะอื่นได้ จึงจะสามารถทำคะแนนได้ในข้อนั้น ๆ	Reparameterized Unified Model (C-RUM)	(2002)
2) โมเดลแบบไม่ชดเชย (Non-compensatory Models) มีลักษณะ คือ ผู้สอบต้องมีความรอบรู้ครบทุก คุณลักษณะ (attributes) ที่จำเป็นต้องใช้ในข้อนั้น ๆ ไม่สามารถทดแทนหรือชดเชยด้วยคุณลักษณะอื่น ๆ ได้จึงจะสามารถทำคะแนนได้ในข้อนั้น ๆ	1. Deterministic input, noisy-and-gate model (DINA) 2. Non-compensatory Reparameterized Unified Model (NC-RUM)	Junker & Sijtsma (2001) DiBello et al. (1995); Hartz (2002)
General 3) โมเดลวินิจฉัยทางพุทธิปัญญาแบบทั่วไป (General CDM) มีความอ่อนปรนมากขึ้น กล่าวคือโมเดลวินิจฉัยทางพุทธิปัญญาแบบทั่วไปอนุญาตให้ ในแบบทดสอบเดียวกันมีข้อสอบแต่ละข้อที่ใช้โมเดลต่างแบบกัน	1. General diagnostic model (GDM) 2. Log-linear CDM (LCDM) 3. Generalized deterministic-input, noisy-and-gate model (G-DINA)	von Davier (2005) Henson et al. (2009) de la Torre (2011)

จากตาราง 10 โมเดลการวินิจฉัยเชิงจำแนกในกลุ่ม Non-compensatory Model เป็นโมเดลแบบไม่ชดเชยคุณลักษณะที่ทำการวินิจฉัยซึ่งใช้ในการแก้ปัญหาของโมเดลในกลุ่มนี้ อยู่ร่วมกันภายใต้กฎการรวมกันเรียกว่า Conjunctive Condensation Rule กล่าวคือ ผู้สอบจำเป็นต้องใช้คุณลักษณะที่ทำการวัดทุกคุณลักษณะร่วมกันในการหาคำตอบของปัญหาที่กำหนด หรือมีความรอบรู้ครบทุกคุณลักษณะที่จำเป็นต้องใช้ในข้อสอบนั้น ๆ จึงจะสามารถหาคำตอบได้อย่างถูกต้องและมีค่าความน่าจะเป็นในการหาคำตอบได้อย่างถูกต้องสูงสุด ผู้สอบที่มีทักษะที่ทำการวัดด้วยข้อคำถามเดียวกันเหมือนกันจะมีค่าความน่าจะเป็นในการตอบข้อคำถามได้อย่างถูกต้องเท่ากัน โดยไม่สามารถทดแทนหรือชดเชยด้วยคุณลักษณะอื่น ๆ ที่แม้มีความเป็นลำดับขั้นและอยู่ในลำดับสูงกว่าได้ จึงเป็นข้อจำกัดในกรณีที่คุณลักษณะมีความเป็นลำดับขั้นซึ่งคุณลักษณะในลำดับที่สูงกว่าสามารถชดเชยคุณลักษณะในลำดับต่ำกว่าได้ การรวมของชุดตัวแปรแฝงตามข้อตกลงนี้จะรวมในรูปแบบของผลผลิต (product) ซึ่งหมายความว่าตัวแปรแฝงแต่ละตัวมีส่วนร่วมกันโดยใช้หลักการคูณ (multiplicative) ซึ่งเป็นโมเดลที่มีลักษณะของโมเดลเชื่อมต่อนี้ (conjunctive model) ผู้สอบต้องมีทุกทักษะที่ได้รับ ในการทำข้อสอบแต่ละข้อ การตอบแฝงจึงมีค่าเท่ากับ 1 แต่ถ้าผู้สอบขาดแม้แต่ทักษะเดียวแล้ว การตอบแฝงจะมีค่าเท่ากับ 0 ตัวอย่างโมเดลการวินิจฉัยเชิงจำแนกในกลุ่มนี้ได้แก่

โมเดล Rule space (วิธีการประเมินกฎ) AHM (วิธีลำดับชั้นคุณลักษณะ) DINA, NIDA และโมเดล Fusion (สุปราณี บุระ, 2554, หน้า 55; ศุภามณ จันทรสกุล และสุกัญญา บุญศรี, 2560, หน้า 57; วราภรณ์ ก้อมน้อย, 2563, หน้า 67; เกศรินทร์ สมราช, 2563, หน้า 64)

โมเดลการวินิจฉัยเชิงจำแนกในกลุ่ม Compensatory Model เป็นโมเดลแบบชดเชยคุณลักษณะที่ทำการวินิจฉัยซึ่งใช้ในการแก้ปัญหของโมเดลในกลุ่มนี้ไม่ได้อยู่ร่วมกันภายใต้กฎการรวมกันที่เรียกว่า Conjunctive Condensation Rule เช่นเดียวกับโมเดลแบบไม่ชดเชย (Non-compensatory Model) สำหรับโมเดลในกลุ่มนี้มีความอ่อนผัน ผู้สอบไม่จำเป็นที่จะต้องใช้คุณลักษณะที่ทำการวัดทุกคุณลักษณะในการหาคำตอบของปัญหาแต่มีเพียงอย่างน้อย 1 คุณลักษณะที่จำเป็นต้องใช้ในข้อคำถามนั้น ๆ ก็เพียงพอให้มีโอกาสตอบถูกได้ ซึ่งหมายถึงลำดับชั้นของคุณลักษณะไม่มีผลต่อความน่าจะเป็นในการตอบถูก เนื่องจากผู้สอบอาจใช้คุณลักษณะอื่นที่ทำการวัดในข้อเดียวกันแทนคุณลักษณะที่ไม่มีในการหาคำตอบของปัญหาที่กำหนดได้ กล่าวคือ ผู้สอบที่มีแบบแผนทักษะความสามารถที่ทำการวัดแตกต่างกันจะมีค่าความน่าจะเป็นในการตอบข้อคำถามได้อย่างถูกต้องไม่เท่ากันการตอบแฉงมีค่าเท่ากับ 1 และถ้าผู้สอบไม่มีทุกทักษะการตอบแฉงจะมีค่าเท่ากับ 0 เช่น ความสามารถในการคำนวณคณิตศาสตร์ที่สูงสามารถทดแทนความบกพร่องความสามารถสื่อสารในวิชาคณิตศาสตร์ได้ เป็นต้น นั่นหมายถึงทักษะบางทักษะสามารถถูกทดแทนได้ด้วยทักษะที่สูงกว่า ทำให้การตอบข้อสอบในข้อนั้นถูกต้องตามข้อตกลงนี้การรวมของชุดตัวแปรแฉงรวมในรูปแบบของผลรวม (sum) หมายความว่าตัวแปรแฉงแต่ละตัวมีส่วนร่วมทำให้เกิดการตอบโดยใช้หลักการบวก (additive) โมเดลการวินิจฉัยเชิงจำแนกในกลุ่มนี้ เช่น DINO Model, NIDO Model และ C-RUM Model

ส่วนโมเดลวินิจฉัยทางพุทธิปัญญาแบบทั่วไป (General CDM) มีความอ่อนปรนมากขึ้น กล่าวคือโมเดลวินิจฉัยทางพุทธิปัญญาแบบทั่วไปอนุญาตให้ในแบบทดสอบเดียวกันมีข้อสอบแต่ละข้อที่ใช้โมเดลต่างแบบกัน ลำดับชั้นของคุณลักษณะมีผลต่อความน่าจะเป็นในการตอบถูก โมเดลวินิจฉัยทางพุทธิปัญญาแบบทั่วไปในกลุ่มนี้ เช่น GDM, LCDM และ G-DINA

การเปรียบเทียบโมเดลวินิจฉัยทางพุทธิปัญญาแบบชดเชยได้ และแบบไม่ชดเชย มีนักการศึกษาได้สรุปไว้อย่างน่าสนใจ ดังนี้

ประเด็นแรก โมเดลวินิจฉัยทางพุทธิปัญญาแบบไม่ชดเชย (non-compensatory models) เหมาะกับการประเมินเชิงวินิจฉัยทางการศึกษา ในลักษณะของการวัดความรู้ความสามารถ ซึ่งได้แก่ โมเดล Rule Space วิธีลำดับชั้นคุณลักษณะ โมเดล DINA โมเดล NIDA และโมเดล Fusion ในขณะที่โมเดลแบบชดเชยได้ (compensatory models) เหมาะกับการประเมินทางจิตวิทยาในลักษณะของการวินิจฉัยความผิดปกติทางจิต เช่น โมเดล DINO

ประการที่สอง เมื่อพิจารณาการนำโมเดลไปประยุกต์ใช้กับข้อมูลเชิงประจักษ์ โมเดลวินิจฉัยทางพุทธิปัญญาแบบไม่ชดเชย (non-compensatory models) ที่ได้จากการพิจารณาในประเด็นแรก โมเดลมีความเหมาะสมกับข้อมูล ซึ่งได้แก่ โมเดล Rule Space วิธีลำดับชั้นคุณลักษณะ โมเดล DINA และโมเดล Fusion

ประเด็นต่อมา เมื่อพิจารณาถึงวิธีในการประมาณค่าพารามิเตอร์และ software ที่ใช้ในการวิเคราะห์ข้อมูล พบว่า โมเดล GDM และโมเดล Fusion ต้องใช้ software ที่มีความเฉพาะในการประมาณค่าพารามิเตอร์ ดังนั้นเพื่อความสะดวก ประหยัด และง่ายในการนำ software มาใช้งาน จึงควรเลือกโมเดลที่สามารถประมาณค่าได้จาก software ที่มีอยู่ทั่วไป อาทิ โมเดล DINA / G-DINA

ประเด็นสุดท้าย ในการวินิจฉัยความสามารถในการอ่านเพื่อความเข้าใจพบว่า มีงานวิจัยปรากฏทั้งโมเดลแบบชดเชยและโมเดลแบบไม่ชดเชยของความสัมพันธ์ระหว่างทักษะย่อยการอ่าน ซึ่งยังเป็นที่ยกเถียงกันในการหาข้อสรุปเกี่ยวกับการเลือกใช้โมเดลที่เหมาะสม เมื่อความสัมพันธ์ในทักษะย่อยของทักษะทางปัญญายังไม่เป็นที่รับรู้อย่างสมบูรณ์ การใช้ CDMs มีความยืดหยุ่นในการรองรับความสัมพันธ์ประเภทต่าง ๆ ระหว่างทักษะย่อย ๆ โมเดล DINA ถูกนำมาใช้ในการวินิจฉัยความสามารถในการอ่านเพื่อความเข้าใจอย่างแพร่หลาย ซึ่ง DINA เป็นตัวอย่างของ CDMs ที่นำไปใช้กับการศึกษาวิจัยของ Chen and Chen (2015) และ Ravand's (2012) (Templin and Henson, 2006; de la Torre, 2008 ; Z. Javidanmehr, 2019, p. 1; สุปราณี บุระ, 2554, หน้า 57-58)

โมเดลวินิจฉัยเชิงจำแนก (งานวิจัยบางฉบับใช้ชื่อว่าโมเดลจัดประเภท) ใช้วิธีการทางสถิติประมาณค่าคุณลักษณะ/ทักษะต่าง ๆ ที่ต้องการวัด เพื่อไปวินิจฉัยจัดประเภทผู้สอบโดยใช้หลักการรวมชุดของตัวแปรแฝงแบบชดเชยได้ และแบบไม่ชดเชยนำไปสู่โมเดลทางสถิติในการวิเคราะห์หลายโมเดลโดยโมเดลที่นิยมได้แก่ G-DINA Model เป็นโมเดลแบบทั่วไป (General) และ DINA Model เป็นโมเดลแบบไม่ชดเชย การนำโมเดลไปใช้ผู้ทดสอบจึงต้องพิจารณาคคุณลักษณะหรือทักษะที่ต้องการวัดเพื่อตัดสินใจเลือกใช้โมเดลในการนำโมเดลไปใช้

2.2 การตัดสินใจเลือกใช้โมเดล

ผู้วิจัยได้ทำการศึกษาและนำเสนอโมเดลวินิจฉัยทางพุทธิปัญญามาข้างต้น โดยพิจารณาถึงประเภทของโมเดลวินิจฉัยทางพุทธิปัญญาที่เหมาะสมกับการพัฒนาแบบสอบวินิจฉัยทางพุทธิปัญญาความสามารถการอ่าน เพื่อความเข้าใจภาษาอังกฤษซึ่งใช้ในการตัดสินใจความรู้ในแต่ละคุณลักษณะแล้วพบว่าโมเดลชั้นแฝง (Latent Class Model) เป็นโมเดลที่เหมาะสม อีกทั้งในการนำโมเดลวินิจฉัยทางพุทธิปัญญาไปประยุกต์ใช้กับข้อมูลเชิงประจักษ์นั้น นักวิจัยควรเลือกใช้โมเดลชั้นแฝงที่ได้รับการยืนยันด้วยผลการวิจัยจากวรรณกรรมที่เกี่ยวข้องว่ามีความสอดคล้องกับข้อมูลหรือเรียกว่า Model Fit ซึ่งได้แก่ Rule-Space Model, DINA Model, Fusion Model และ Attribute Hierarchy Method

เมื่อพิจารณาถึงข้อจำกัดของโมเดลชั้นแฝง อาทิ Rule-Space Model คือ อาจให้ผลวินิจฉัยไม่สมบูรณ์ เนื่องจากการกำหนดคุณลักษณะดำเนินการหลังจากสร้างแบบทดสอบเรียบร้อยแล้ว การจำแนกคุณลักษณะที่ต้องใช้ในการตอบข้อสอบแต่ละข้อจึงอาจไม่ครอบคลุมคุณลักษณะหรือความรู้ที่ผู้สอบต้องมี ส่วนโมเดล Fusion ต้องใช้ Software ที่มีความเฉพาะในการประมาณค่าพารามิเตอร์ สำหรับโมเดลลำดับชั้นของคุณลักษณะ (Attribute Hierarchy Model) ถึงแม้ว่า AHM จะมีจุดเด่นอยู่ที่การให้คำจำกัดความของค่าคุณลักษณะ (Attribute) ชัดเจนขึ้น กำหนดให้มีการจำแนกคุณลักษณะก่อนการพัฒนาแบบสอบ และกำหนดให้คุณลักษณะมีความสัมพันธ์เป็นลำดับชั้น ซึ่งสิ่งนี้เป็นขั้นตอนสำคัญในการช่วยกำหนดลักษณะของข้อสอบแต่ละข้อให้เป็นไปตามลำดับชั้นทางพุทธิปัญญาอย่างชัดเจน ทำให้การวินิจฉัยจากแบบแผนการตอบข้อสอบให้สารสนเทศที่เป็นประโยชน์ และมีความชัดเจน แต่โมเดลยังมีข้อจำกัดในประเด็นของการจำแนกแบบแผนการตอบข้อสอบที่สังเกตได้ อาจไม่ชัดเจน หากแบบแผนการตอบข้อสอบที่สังเกตได้นั้นไม่ตรงกับแบบแผนการตอบข้อสอบที่คาดหวังได้เลย (เกศรินทร์ สมราช, 2563; ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์, 2563)

การตัดสินใจเลือกใช้โมเดลวินิจฉัยทางพุทธิปัญญาในการศึกษาครั้งนี้ ผู้วิจัยพิจารณาจากการศึกษางานวิจัยความสามารถทางภาษาอังกฤษที่ใช้การประเมินเชิงวินิจฉัยทางพุทธิปัญญา พบว่า นักวิชาการได้นำโมเดลวินิจฉัยทางพุทธิปัญญาเชิงจำแนก 2 โมเดล ได้แก่ DINA และ G-DINA มาใช้อย่างแพร่หลายใกล้เคียงกัน แต่มีข้อสังเกตที่น่าสนใจว่าในระยะหลังจนถึงปัจจุบันพบว่าส่วนใหญ่มีเพียงโมเดล G-DINA เท่านั้นที่ถูกนำมาใช้ในงานวิจัยทางภาษา อาทิ การศึกษาของ Chen & Chen (2015) สัมภาษณ์การอ่านเพื่อความเข้าใจความสัมพันธ์ผ่านโมเดล G-DINA และการศึกษาของ Javidanmehr & Salab (2019) เกี่ยวกับการนำผลการทดสอบที่มีอยู่ โดยมีได้มาจากแบบสอบที่สร้างขึ้นเพื่อวินิจฉัย นำมาประเมินความสามารถในการอ่านเพื่อความเข้าใจ: การประยุกต์ใช้โมเดล G-DINA สู่การประเมินความสามารถการอ่านที่มีความสำคัญต่อชีวิตโดยมีจุดมุ่งหมายเพื่อเข้าเรียนต่อหรือจบการศึกษาและการศึกษาของ Tonekaboni & Ravand (2021) เกี่ยวกับการสร้างและการตรวจสอบความถูกต้องของ Q-matrix สำหรับการประเมินความสามารถการอ่านเพื่อความเข้าใจที่มีความสำคัญต่อชีวิต: การศึกษาด้วยโมเดล G-DINA และการศึกษาของเกศรินทร์ สมราช (2563) ศึกษาเกี่ยวกับการพัฒนาแบบสอบวินิจฉัยความสามารถในการใช้ภาษาอังกฤษ เพื่อการสื่อสารของนิสิตครูโดยประยุกต์ใช้ G-DINA MODEL

เพื่อให้มีความชัดเจนมากขึ้น ผู้วิจัยนำเสนอการเปรียบเทียบรายละเอียดของโมเดลวินิจฉัยทางพุทธิปัญญาเชิงจำแนก 2 โมเดลที่ใช้้อย่างแพร่หลายในงานวิจัยทางภาษาอังกฤษ ได้แก่ โมเดล DINA และโมเดล G-DINA โดยผู้วิจัยอิงข้อมูลจากตารางเปรียบเทียบโมเดลวินิจฉัยเชิงพุทธิปัญญา ของแคทรียา แสงดาวเทียน (2562, หน้า 43) มีรายละเอียดดังตารางที่ 11

ตาราง 11 การเปรียบเทียบโมเดลวินิจฉัยทางพุทธิปัญญาเชิงจำแนก

โมเดล	ประเภท CDM		วิธี ประมาณค่า พารามิเตอร์	ซอฟต์แวร์	ระดับตัว แปรแฝงที่ จำแนก/ สังเกตได้	โมเดล CDM	จุดเด่น
	ไม่ ชัดเจน ได้	ชัดเจน					
DINA	✓		JML, MML, MCMC	R OX M-plus	แบบสอง ค่า	โมเดลชั้น แฝง	ประหยัด/แปลผลได้ จาก 2 พารามิเตอร์/มี ความยืดหยุ่นในการ รองรับความสัมพันธ์ ประเภทต่างๆระหว่าง ทักษะย่อยๆ /ซอฟต์แวร์ ที่ใช้ประมาณค่าหาได้ ง่ายสะดวก
G- DINA		✓	JML, MML, MCMC	R M-plus	แบบสอง ค่า, แบบหลาย ค่า	โมเดลชั้น แฝง	มีการประมาณค่าเพื่อ วินิจฉัยคุณลักษณะที่ ผ่อนคลายกว่า/ยอมให้ ใช้ต่างโมเดลประมาณ ค่าความสามารถในแบบ สอบเดียวกัน

จากตาราง 11 ทั้ง DINA Model และ G-DINA Model จัดเป็นโมเดลชั้นแฝงที่มีวิธีประมาณค่าพารามิเตอร์ที่ไม่ต่างกัน ระดับตัวแปรแฝงที่จำแนกและระดับตัวแปรแฝงที่สังเกตได้ของ DINA เป็นแบบสองค่าเท่านั้น (รอบรู้/ไม่รอบรู้) G-DINA แบบสองค่าและแบบหลายค่า เนื่องจากยินยอมให้มีการใช้ต่างโมเดลกันในการประมาณค่าความสามารถในแบบสอบฉบับเดียวกัน ซอฟต์แวร์ที่ใช้ประมาณค่าหาได้ง่ายสะดวกโดยใช้โปรแกรม R โปรแกรม M-plus ส่วน DINA สามารถใช้โปรแกรม OX ได้ด้วยซึ่งไม่ต่างกันมาก

2.3 ข้อดีและข้อจำกัดของ DINA Model และ G-DINA Model

ข้อดีและข้อจำกัดของ DINA Model และ G-DINA Model จากทัศนะของนักวิชาการสรุปได้ ดังนี้

สุมาลี มีสกุล และคณะ (2558, หน้า 30) สรุปว่า ในการพิจารณาประสิทธิภาพของโมเดลจากค่าดัชนีความสอดคล้องของโมเดลเชิงสัมบูรณ์ (Absolute Fit Index) และค่าดัชนีความสอดคล้องของโมเดลเชิงสัมพัทธ์ (Relative Fit Index) ระหว่างผลที่ได้จากการวิเคราะห์ด้วย

G-DINA Model กับ DINA Model กล่าวไว้ว่า DINA Model เป็นโมเดลที่ซอกตกลงเบื้องต้นของ ข้อมูลที่ใช้วิเคราะห์เช่นเดียวกับ G-DINA Model แต่มีการประมาณค่าความน่าจะเป็นในการหาคำตอบของปัญหาของผู้สอบได้อย่างถูกต้อง โดยไม่คำนึงถึงอิทธิพลซึ่งเป็นความแตกต่างของความสามารถในแบบแผนทักษะความสามารถทางพุทธิปัญญาที่นำมาใช้ในการหาคำตอบของผู้สอบ DINA Model จึงมีการประมาณค่าพารามิเตอร์ในการคำนวณเป็นจำนวนน้อยกว่า G-DINA Model ซึ่งอาจกล่าวได้ว่า DINA Model เป็นโมเดลที่มีความประหยัดในทางสถิติมากกว่า

แคทรียา แสงดาวเทียน (2560) G-DINA model เป็นโมเดลให้ประสิทธิภาพในการประมาณค่าดีกว่า DINA model ทั้งนี้เนื่องจากฐานคิดในการประมาณ ค่าพารามิเตอร์ของ G-DINA model ที่มีความยืดหยุ่นเหมาะสมกับแบบแผนความสามารถของผู้สอบที่ต่างกัน โดย G-DINA model จะกำหนดให้ผู้สอบซึ่งมีแบบแผนทักษะความสามารถต่างกัน มีความน่าจะเป็นในการทำข้อสอบได้อย่างถูกต้องไม่เท่ากัน นักเรียนที่มีแบบแผนทักษะความสามารถที่ใช้แก้ปัญหาสอดคล้องกับทักษะที่ใช้ในการหาคำตอบของปัญหามากกว่า จะมีความน่าจะเป็นในการหาคำตอบของปัญหาได้อย่างถูกต้องมากกว่านักเรียนที่มีแบบแผนทักษะความสามารถที่ใช้แก้ปัญหาสอดคล้องกับทักษะที่ใช้ในการหาคำตอบของปัญหาน้อยกว่าและจะมีความน่าจะเป็นในการหาคำตอบของปัญหาได้อย่างถูกต้องมากที่สุดเมื่อมีแบบแผนทักษะความสามารถที่ใช้แก้ปัญหาสอดคล้องกับทักษะที่ใช้ในการหาคำตอบของปัญหาทุกทักษะ

ศุภามณ จันทรสกุล และสุกัญญา บุญศรี (2560) กล่าวว่า G-DINA model มีความยืดหยุ่นกว่า DINA model โดยทักษะบางทักษะสามารถถูกขจัดเซย์ได้ด้วยทักษะที่สูงกว่า แต่ในขณะที่ DINA model มีการประมาณค่าพารามิเตอร์น้อยกว่า G-DINA Model จึงเป็นโมเดลที่มีความประหยัด (Parsimonious model) มากกว่า G-DINA model นอกจากนี้ DINA model เป็นโมเดลแบบไม่ขจัดเซย์จึงมีซอกตกลงเบื้องต้นว่าผู้สอบจะต้องมีทักษะทุกทักษะซึ่งขาดทักษะใดทักษะหนึ่งไปไม่ได้ทำให้การจัดประเภททักษะ/ความสามารถของผู้สอบมีความเฉพาะกับแบบแผนความสามารถของผู้สอบ มากกว่า G-DINA model

Rupp et al. (2010, อ้างถึงใน ฌัฐกานต์ ประจันบาน, 2563, หน้า 29) กล่าวว่า G-DINA นี้ พัฒนามาจาก DINA MODEL ให้มีการนำมาใช้ในการวิเคราะห์ได้ง่ายขึ้น DCMs เหมาะที่จะใช้กับตัวแปรสังเกตได้และตัวแปรแฝงแบบ ใ้คะแนน 2 ค่า คือ การให้คะแนนแบบถูก-ผิด และการจัดกลุ่มตัวแปรแฝงแบบรอบรู้-ไม่รอบรู้ เพราะมีความยืดหยุ่นและสามารถประยุกต์ใช้ได้ในทุกบริบท

เดอลาทอร์ (de la Torre, 2011, p. 196, อ้างถึงใน วรางคณา ก้อมน้อย, 2563, หน้า 67) สรุปไว้ว่า G-DINA model มีการผ่อนปรนซอกตกลงเบื้องต้นเกี่ยวกับความน่าจะเป็นในการทำข้อสอบได้ โดยอนุญาตให้ในแบบทดสอบเดียวกัน มีการใช้ต่างโมเดลในการประมาณค่าแต่ละข้อ

(item-by-item model estimation) และทดสอบสมมุติฐานได้ ขึ้นอยู่กับการออกแบบและเมตริกซ์ ความสัมพันธ์ระหว่างข้อคำถามและคุณลักษณะหรือเมตริกซ์คิว (Q-matrix) และมีการประมาณค่าพารามิเตอร์โดยใช้อัลกอริทึม marginalized maximum likelihood estimation (MMLE) ซึ่งในการประมาณค่าวิธีนี้สามารถประมาณค่าพารามิเตอร์ได้รวดเร็วและมีประสิทธิภาพกว่าการประมาณค่าโดยอัลกอริทึม Markov chain Monte Carlo อีกทั้งการใช้โมเดลจีดีไอเอ็นเอ (G-DINA model) ยังสามารถช่วยลดรูปโมเดลจากโมเดลเต็มรูป และใช้การทดสอบวาลด์ (Wald test) ในการเปรียบเทียบโมเดลเต็มรูปและลดรูปสำหรับข้อที่มีพหุคุณลักษณะได้ทำให้ในการประมาณค่าพารามิเตอร์สามารถประมาณค่าจากโมเดลลดรูปได้ (de la Torre, 2011, p. 196, อ้างถึงใน วรางคณา ก้อมน้อย, 2563, หน้า 67)

Aryadoust (2018) สรุปว่า G-DINA จะมีข้อจำกัดน้อยกว่าโมเดล DINA และ DINO แต่โมเดลที่มีความยืดหยุ่นสูงก็จะมีความซับซ้อนมากกว่า กล่าวคือ จะใช้ข้อสอบมากกว่าและใช้การประมาณค่าพารามิเตอร์แบบไม่จำกัดโดยเป็นผลการศึกษากการประเมินผลการวินิจฉัยเชิงพุทธิปัญญาของการทดสอบการฟังในการสอบ GCE O-Level ของสิงคโปร์-เคมบริดจ์ โดยประยุกต์ใช้ Model DINA, DINO, G-DINA, HO-DINA และ RRUM

จากที่กล่าวมาสรุปได้ว่า G-DINA model ต่างจาก DINA model ในแง่ความแตกต่างในการประมาณค่าพารามิเตอร์ของ โดย G-DINA จะไม่สร้างพารามิเตอร์เดี่ยว (single parameter) ต่อข้อสอบรายข้อแต่จะใช้ EM algorithm คำนวณ Posterior probabilities ของทักษะและจำแนกชั้นแห่งของผู้สอบ (De La Torre, 2011; Rupp & Templin, 2008; Rupp, Templin & Henson, 2010; Basokcu, Ogretmen & Kelecioğlu, 2013)

2.4 การประเมินความสอดคล้องของโมเดลการวินิจฉัยเชิงจำแนก (Evaluating the Model Fit of Diagnostic Classification Models)

1) ค่าสถิติสำหรับใช้พิจารณาความสอดคล้องกลมกลืนของโมเดลทางสถิติกับข้อมูลเชิงประจักษ์ในเชิงสัมบูรณ์ (Absolute Fit Index)

ค่าสถิติพื้นฐานสำหรับนำไปใช้ในการพิจารณาความสอดคล้องกลมกลืนของโมเดลทางสถิติกับข้อมูลเชิงประจักษ์ที่ใช้ในโมเดลการวินิจฉัยเชิงจำแนกนี้มี 3 ตัว คือ ค่าสัดส่วนการตอบถูก (Proportion Correct) ค่าการแปลงสหสัมพันธ์ (Transformed Correlation) และค่า Log-Odds Ratio ค่าสถิติแต่ละตัวมีสูตรในการคำนวณดังนี้

$$\text{Prop} = \left[\frac{p_j(1-p_j)}{I} \right]^{\frac{1}{2}}$$

$$Z(\text{Corr}) = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right)$$

$$\text{และ } \text{Log (OR)} = [I_{00}^{-1} + I_{10}^{-1} + I_{01}^{-1} + I_{11}^{-1}]^{1/2}$$

เมื่อ Prop แทน ค่าสัดส่วนการตอบถูก (Proportion Correct)

p_j แทน ค่าคาดหวังของสัดส่วนการตอบถูกสำหรับข้อคำถามที่ j

r แทน ค่าสหสัมพันธ์ของข้อคำถามรายข้อ

Z (Corr) แทน ค่าสหสัมพันธ์การแปลงมาตรฐาน (Z-Transformed Correlation)

Log (OR) แทน ค่าอัตราส่วน Log-Odds (Log-Odds Ratio)

$I_{yy'}$ แทน จำนวนของผู้สอบที่ได้คะแนน y ในข้อคำถามที่ j และ y' ในข้อคำถามที่ j'

2) ค่าสถิติสำหรับใช้พิจารณาเปรียบเทียบความสอดคล้องของ โมเดลทางสถิติในเชิงสัมพัทธ์ (Relative Fit Index)

ค่าสถิติที่ใช้พิจารณาเปรียบเทียบความสอดคล้องของโมเดลทางสถิติในเชิงสัมพัทธ์ของโมเดลการวินิจฉัยเชิงจำแนกมีอยู่ 3 ตัว คือ -2Loglikelihood (-2LL), Akaike's information criterion (AIC) และ Bayesian information criterion (BIC) โดยค่าสถิติทั้งสามตัวมีสมการที่ใช้ในการประมาณค่าดังนี้

$$-2LL = -2\ln(\text{ML})$$

$$\text{โดยที่ } \text{ML} = \prod_i \sum_l p(\alpha_l) L(X_i | \beta, \alpha_l)$$

เมื่อ $p(\alpha_l)$ แทน ค่าความน่าจะเป็นในการตอบถูกของผู้สอบที่มีแบบแผนทักษะความสามารถ α_l

$L(X_i | \beta, \alpha_l)$ แทน ค่าโลลิฮูดฟังก์ชันความสามารถของผู้สอบที่มีแบบแผนทักษะความสามารถ α_l

$$\text{และ Information criterion} = -2\ln(L) + kp$$

เมื่อ L แทน ค่า log-likelihood ที่คำนวณได้จากการประมาณค่าในโมเดลการวินิจฉัยเชิงจำแนก

P แทน จำนวนข้อคำถามทั้งหมดและจำนวนพารามิเตอร์ที่ประมาณค่าในโมเดลการวินิจฉัยเชิงจำแนก

K แทน ค่าคงที่ซึ่งมีค่าเท่ากับ 2 สำหรับดัชนี AIC และมีค่าเท่ากับค่าลอการิทึมของขนาดกลุ่มตัวอย่าง สำหรับดัชนี BIC

ทั้งดัชนี -2LL, AIC และ BIC แสดงการรวมตัวกันระหว่างค่าความสอดคล้องของโมเดลและความประหยัดของโมเดล ซึ่งหมายความว่าโมเดลที่มีความซับซ้อนมากจะนำไปสู่การพัฒนาความสอดคล้องของโมเดลเพียงเล็กน้อยเมื่อเปรียบเทียบกับโมเดลที่มีความซับซ้อนน้อยกว่าซึ่งจะมีอุปสรรคมากกว่า สำหรับค่าสถิติทั้งสามตัวค่าที่ต่ำกว่าจะเป็นค่าที่ยอมรับว่ามีประสิทธิภาพดีกว่า นอกจากนี้ยังมีดัชนีค่าความสอดคล้องของโมเดลในเชิงสัมพัทธ์ในกลุ่ม Information criterion ที่ถูกดัดแปลงไปโดยมีพื้นฐานอยู่บนดัชนี AIC และ BIC เช่น Sample-size adjusted AIC , DIC และ Bayes' factors

2.5 โมเดลจี-ดีไอเอ็นเอ (G-DINA model)

โมเดลจีดีไอเอ็นเอ (generalized deterministic input, noisy-and-gate: G-DINA model) แนวคิดนี้ได้รับการพัฒนาขึ้นโดย de la Torre, J. (2011) G-DINA model คือ ลักษณะทั่วไปของโมเดลดีไอเอ็นเอ (DINA model) ที่มีการผ่อนคลายข้อตกลงเบื้องต้น โดยอนุญาตให้ในแบบทดสอบเดียวกัน ใช้โมเดลแตกต่างกันในแต่ละข้อได้ (item-by-item model estimation) ขึ้นอยู่กับการออกแบบและ Q-matrix โมเดลเฉพาะต่าง ๆ ทั้ง DINA DINO NC-RUM C-RUM และ ACDM สามารถวิเคราะห์ได้ใน GDINA เช่น ในแบบทดสอบเดียวกัน บางข้อ DINA model คือตัวเลือกที่ดีที่สุดแต่ในบางข้อในแบบทดสอบนั้น DINO model หรือ C-RUM อาจเหมาะสมที่สุด เป็นต้น (de la Torre, 2011 และ Ravand, Hamdollah & Robitzsch, Alexander , 2015) สมการความน่าจะเป็นที่นักเรียนคนที่ i จะตอบข้อสอบข้อที่ j ใน G-DINA model กรณีที่ข้อนั้น ๆ ต้องการคุณลักษณะ (attributes) หรือทักษะ (skills) α จำนวน K คุณลักษณะ คือ (George A.C., Robitzsch A, Kiefer T, GroB J, Ünlü A, 2016, อ้างถึงใน วรารคณา ก้อมน้อย, 2563, หน้า 68-69)

$$P(\alpha_{ij}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{ik} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{k'-1} \delta_{jkk'} \alpha_{ik} \alpha_{ik'} + \dots + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{ik},$$

เมื่อ δ_{j0} คือ ความน่าจะเป็นเริ่มต้นหรือความน่าจะเป็นในการตอบข้อสอบถูกเมื่อไม่มีคุณลักษณะ

ใดเลยหรือเรียกว่าค่าการเดา

δ_k คือ อิทธิพลหลัก (Main effect) ของการบรรลุคุณลักษณะเดียว

$\delta_{kk'}$ คือ ปฏิสัมพันธ์ของการบรรลุคุณลักษณะสองอย่าง

$\delta_{j1,2,\dots,K_j^*}$ คือ ปฏิสัมพันธ์ที่ซับซ้อนที่เกิดจากการบรรลุคุณลักษณะทุกอย่าง

เช่น ตัวอย่างที่ 1 สมการความน่าจะเป็นที่นักเรียนคนที่ i จะตอบข้อสอบข้อที่ j ใน G-DINA model กรณีที่ข้อนั้น ๆ ต้องการคุณลักษณะ (attributes) α_1 และ α_2 สองคุณลักษณะ คือ

$$P(X_{ij} = 1 | \alpha_1, \alpha_2) = \delta_{j0} + \delta_{j1}\alpha_1 + \delta_{j2}\alpha_2 + \delta_{j12}\alpha_1\alpha_2$$

โดยที่ พารามิเตอร์ δ_{j0} คือ ความน่าจะเป็นในการตอบข้อนั้น ๆ ถูก กรณีผู้ตอบข้อสอบไม่มีคุณลักษณะที่ต้องการในข้อนั้น ๆ

พารามิเตอร์ δ_{j1} and δ_{j2} คือ ค่าอิทธิพลหลักของคุณลักษณะ α_1 และ α_2 ตามลำดับ และ พารามิเตอร์ δ_{j12} คือ ค่าอิทธิพลปฏิสัมพันธ์ของทุกคุณลักษณะ α_1 และ α_2

จะเห็นว่าใน G-DINA model ผู้ตอบข้อสอบไม่จำเป็นต้องมีคุณลักษณะ α_1 และ α_2 ครบพร้อมกัน จึงจะมีโอกาสในการทำข้อที่ j ถูก เนื่องจากความน่าจะเป็นในการตอบถูกนั้นขึ้นอยู่กับ 1) ความน่าจะเป็นในการตอบข้อ j ถูก กรณีผู้ตอบข้อสอบไม่มีคุณลักษณะที่ต้องการ 2) ค่าอิทธิพลหลักของคุณลักษณะที่ต้องการ และ 3) ค่าอิทธิพลปฏิสัมพันธ์

ตัวอย่างที่ 2 สมการความน่าจะเป็นที่นักเรียนคนที่ i จะตอบข้อสอบข้อที่ j ใน G-DINA model กรณีที่ข้อนั้น ๆ ต้องการคุณลักษณะ (attributes) α_1 α_2 และ α_3 สามคุณลักษณะ คือ

$$P(X_{ij} = 1 | \alpha_1, \dots, \alpha_k) = \delta_{j0} + \delta_{j1}\alpha_1 + \delta_{j2}\alpha_2 + \delta_{j3}\alpha_3 + \delta_{j12}\alpha_1\alpha_2 + \delta_{j13}\alpha_1\alpha_3 + \delta_{j23}\alpha_2\alpha_3 + \delta_{j123}\alpha_1\alpha_2\alpha_3$$

โดยที่ พารามิเตอร์ δ_{j0} คือ ความน่าจะเป็นในการตอบข้อนั้น ๆ ถูก กรณีผู้ตอบข้อสอบไม่มีคุณลักษณะที่ต้องการในข้อนั้น ๆ

พารามิเตอร์ δ_{j1} δ_{j2} and δ_{j3} คือ ค่าอิทธิพลหลักของคุณลักษณะ α_1 α_2 และ α_3 ตามลำดับ

และ พารามิเตอร์ δ_{j123} คือ ค่าอิทธิพลปฏิสัมพันธ์ของทุกคุณลักษณะ α_1 α_2 และ α_3

เช่นเดียวกับตัวอย่างที่ 1 จะเห็นว่าใน G-DINA model ผู้ตอบข้อสอบไม่จำเป็นต้องมีคุณลักษณะ α_1 α_2 และ α_3 ครบพร้อมกัน จึงจะมีโอกาสในการทำข้อที่ j ถูก เนื่องจากความน่าจะเป็นในการตอบถูกนั้นขึ้นอยู่กับความน่าจะเป็นในการตอบข้อ j ถูก กรณีผู้ตอบข้อสอบไม่มีคุณลักษณะที่ต้องการ ค่าอิทธิพลหลักของคุณลักษณะที่ต้องการ ค่าอิทธิพลปฏิสัมพันธ์ที่ละคู่ และค่าอิทธิพลปฏิสัมพันธ์ร่วมของทุกคุณลักษณะที่จำเป็นต้องใช้ในข้อนั้น ๆ

2.6 การประมาณค่าพารามิเตอร์ในโมเดลวินิจฉัยเชิงพุทธิปัญญา CDM

โปรแกรมซอฟต์แวร์ส่วนใหญ่ที่พบในการประมาณค่าโมเดลวินิจฉัยเชิงพุทธิปัญญา CDM ตัวอย่างเช่น ดำเนินการวิเคราะห์ GDM, ใช้โปรแกรม Arpeggio Suite (Bolt et. al., 2008) ดำเนินการ Fusion model (NC-RUM) ด้วยโปรแกรม M-plus (Muthen & Muthen, 2013) ซึ่ง

เป็นซอฟต์แวร์ที่มีวัตถุประสงค์ทั่วไป Arpeggio และ M-plus เป็นซอฟต์แวร์เชิงพาณิชย์ โปรแกรม M-plus โดยขอรับ Commercial version ได้ที่ jtemplin@uga.edu นอกจากนี้ยังวิเคราะห์ได้ด้วย โปรแกรม R ด้วย CDM package ซึ่งเป็น Freeware และโปรแกรม Ox ซึ่งเป็น Freeware ขอรับได้ที่ j.delatorre@rutgers.edu (Rupp & Templin, 2008) อย่างไรก็ตามพบปัญหาอื่นการใช้ซอฟต์แวร์ที่พบบ่อยที่สุด เช่น Arpeggio และ M-plus คือการเตรียมการคำสั่งในโปรแกรมเป็นกระบวนการที่ไม่สะดวกสบายและโดยเฉพาะอย่างยิ่งในกรณีของ M-plus จะใช้เวลาวิเคราะห์เป็นเวลานาน ตัวอย่างเช่น การทดสอบที่ประกอบด้วย 4 คุณลักษณะ ต้องเขียนคำสั่งเป็นจำนวนหลายหน้าสำหรับ M-plus การประมาณค่าพารามิเตอร์ CDM เป็นตัวเลขของคุณลักษณะที่เกี่ยวข้องกับการทดสอบเพิ่มขึ้น เวลาที่ใช้ไปประมาณการแบบจำลองเพิ่มขึ้นแบบทวีคูณ Run แต่ละครั้ง สำหรับ 4 คุณลักษณะจะใช้เวลาประมาณ 28 นาทีและการวนซ้ำ 1,000 รอบด้วยคอมพิวเตอร์ที่มี RAM 2GB และ CPU Core i3 ซอฟต์แวร์อื่นที่มีลักษณะที่น่าสนใจมากที่สุดคือ โปรแกรม R โดยเรียกใช้แพ็คเกจ CDM มีจุดเด่นที่สำคัญ 4 ประการ (Robitzsch, Kiefer, Cathrice George, & Uenlue, 2014) ซึ่งมีความมุ่งหมายไว้สำหรับการวิเคราะห์โมเดลการวินิจฉัยทางพุทธิปัญญาคือ: (ก) มีประสิทธิภาพด้านเวลามาก: การประมาณค่าของพารามิเตอร์ตัวอย่างเช่น โมเดลที่มีคุณลักษณะ 5 อย่าง ใช้เวลาน้อยกว่าหนึ่งนาที (ข) มีประสิทธิภาพในการ RUN โมเดล CDM ตัวหลัก ๆ เช่น DINA, DINO, NC-RUM, GDM และ G-DINA (ค) สามารถดาวน์โหลดมาใช้โดยไม่เสียเงินและ (ง) ใครก็สามารถใช้โปรแกรม R ได้ โดยพิมพ์คำสั่ง 2-3 บรรทัดเท่านั้น (Ravand and Robitzsch, 2015, p. 5)

Ma and de la Torre (2020, pp. 18-19) กล่าวไว้ว่า การประมาณค่าพารามิเตอร์ δ จากแพ็คเกจ GDINA โดยการเปรียบเทียบประสิทธิภาพในการประมาณค่าพารามิเตอร์ระหว่างแพ็คเกจ GDINA กับ Templin และ Hoffman (2013) โดยใช้ M-plus พบว่าประมาณค่าได้เหมือนกันโดยมีความแตกต่างเพียงเล็กน้อยที่ตำแหน่งทศนิยมที่สาม จาก Templin and Hoffman (2013). แพ็คเกจ GDINA ใช้เวลาในการวิเคราะห์ 21.42 วินาที ในขณะที่ M-plus ใช้เวลา 32 นาที 4 วินาที เห็นได้ชัดเจนว่าแพ็คเกจ GDINA ใช้เวลาในการ Run น้อยกว่า M-plus มากและประมาณค่าพารามิเตอร์ได้ค่าที่เหมือนกัน

2.7 แนวคิดสำหรับการออกแบบการประเมินเพื่อวินิจฉัย

แนวคิดที่สำคัญในปัจจุบันสำหรับหลักการออกแบบการประเมินโดยทั่วไปและหลักการออกแบบการประเมินเพื่อวินิจฉัยมีอยู่ 2 แนวคิดคือ การออกแบบแบบอิงหลักฐาน (Evidence centered design: ECD) ซึ่งเสนอโดย Mislevy, Almond and Lukas และการออกแบบจากระบบการคิด (Cognitive design system: CDS) ซึ่งเสนอโดย Susan Embretson และคณะ (2003, Rupp et al., 2010, p. 15, อ้างถึงโดย สุมาลี มีสกุล, 2558, หน้า 25) ซึ่งมีรายละเอียดดังต่อไปนี้

2.7.1 การออกแบบแบบอิงหลักฐาน (Evidence centered design: ECD)

เป้าหมายของการสร้างแบบทดสอบแบบอิงหลักฐาน (ECD-based test construction) คือการรวบรวมเหตุการณ์ที่สนับสนุนการอ้างอิงเกี่ยวกับบุคคล ซึ่งแตกต่างจากการประเมินโดยทั่วไปที่จำกัดขอบเขตในการอ้างอิงเกี่ยวกับคะแนนที่แสดงถึงสมรรถนะที่ทำการประเมินเพียงประเภทเดียว (Leighton & Gierl, 2007, p. 179, อ้างถึงใน แคทรียา แสงดาวเทียน, 2562, หน้า 19-22) การออกแบบโดยใช้เหตุการณ์เป็นศูนย์กลางหรือชื่อเรียกโดยย่อ ECD ถูกพัฒนาขึ้นเพื่อสร้างระบบการคิดของนักออกแบบการประเมินเพื่อวินิจฉัยที่เป็นอิสระจากภาวะสันนิษฐาน การประเมินเพื่อวินิจฉัยนั้นต้องการวัดซึ่งในความเป็นจริงการประเมินเพื่อวินิจฉัยส่วนใหญ่ได้รับประโยชน์จากการออกแบบการประเมินภายใต้แนวคิด ECD หลักการออกแบบแบบอิงหลักฐานนี้จะใช้การบรรยายเหตุการณ์เกี่ยวกับสิ่งที่ต้องการประเมินว่าผู้ตอบต้องทำอะไรและใช้ความรู้ใดจึงจะบรรลุเป้าหมายของการแก้ปัญหาในสถานการณ์หรืองานที่กำหนดรายละเอียดของการออกแบบการประเมินเพื่อวินิจฉัยโดยใช้ ECD ประกอบไปด้วยองค์ประกอบทั้งหมด 5 แบบ ดังนี้

3.2.1 โมเดลนักเรียน (Student model) เป็นโมเดลที่แสดงการอ้างอิงของโครงสร้างความสามารถในการแก้ปัญหาของนักเรียนในสถานการณ์ต่าง ๆ

3.2.2 โมเดลปัญหา (Task model) เป็นโมเดลที่แสดงระบบการปฏิบัติในการแก้ปัญหาว่าควรถูกกระทำในลักษณะใด

3.2.3 โมเดลเหตุการณ์ (Evidence model) เป็นโมเดลการวัดทางจิตวิทยาที่เชื่อมโยงระหว่างโมเดลนักเรียนและโมเดลปัญหา

3.2.4 โมเดลรวม (Assembly model) เป็นโมเดลที่จัดระบบว่าโมเดลนักเรียนโมเดลปัญหาและโมเดลสถานการณ์ เชื่อมโยงกันอย่างไรในการประเมิน

3.2.5 โมเดลนำเสนอ (Presentation model) เป็นโมเดลที่แสดงว่าปัญหาที่ใช้ในการประเมินจะถูกนำเสนออย่างไร

2.7.2 การออกแบบจากระบบการคิด (Cognitive design system: CDS)

การออกแบบจากระบบการคิดหรือชื่อเรียกโดยย่อ CDS ถูกพัฒนาโดยการวิจัยที่มีมาอย่างต่อเนื่องในเรื่องของกระบวนการตอบสนองทางความคิดสำหรับการประเมิน ซึ่งวัดความสามารถทางการคิดขั้นพื้นฐาน เช่น มิติสัมพันธ์ (Spatial rotation) หรือการให้เหตุผลโดยทั่วไป (General reasoning) การแสดงออกถึงภาวะสันนิษฐานเหล่านี้ผ่านการทำแบบทดสอบในการประเมินเพื่อวินิจฉัยโดยมีพื้นฐานมาจากความเข้าใจในเรื่องของจิตวิทยาประยุกต์ทางด้านการคิดในบริบทของการประยุกต์ใช้ในอุดมคติลักษณะที่ออกแบบขึ้นของงานในการประเมินเพื่อวินิจฉัยความสามารถ (สุมาลี มีสกุล, 2558, หน้า 25)

กรอบกระบวนการออกแบบการวินิจฉัยเชิงพุทธิปัญญาของเอ็มเบรทสัน

เอ็มเบรทสัน (Embretson, 1983) ตั้งสมมติฐานว่าทฤษฎีและวิธีการของจิตวิทยาการรู้คิดหรือจิตวิทยาทางปัญญายินยอมให้ความตรงเชิงโครงสร้างถูกแยกออกเป็นสองลักษณะได้แก่ โครงสร้างที่เกี่ยวข้องกับกระบวนการเชิงกลยุทธ์ และโครงสร้างความรู้ (schema) ที่เกี่ยวข้องในการตอบสนองข้อสอบ ข้อมูลเชิงประจักษ์ที่เกี่ยวข้องทั้งหมดถูกรวมเข้าด้วยกันในโมเดล (ความแม่นยำและเวลาตอบสนอง) จากแง่มุมต่างๆ ที่หลากหลายอย่างเป็นระบบของหน้าที่ข้อสอบ Nomothetic span เกี่ยวข้องกับความสัมพันธ์ของการทดสอบกับการวัดแบบอื่นๆ ของความแตกต่างรายบุคคล ข้อมูลเชิงประจักษ์ที่เกี่ยวข้องรวมถึงความสัมพันธ์ของคะแนนสอบกับแบบทดสอบอื่นๆ เกณฑ์การวัดและอื่นๆ ดังนั้น Nomothetic span จึงหมายถึง ความเที่ยงตรงและความสัมพันธ์ระหว่างภาวะสันนิษฐานซึ่งถูกวัดโดยการประเมินกับทฤษฎีที่เกี่ยวข้องกับภาวะสันนิษฐานนั้น จึงมีบทบาทที่ชัดเจนในการกำหนดความตรงเชิงโครงสร้างหลักการของขั้นตอนการออกแบบการประเมินด้วย CDS ประกอบด้วย 7 ขั้นตอน (Embretson, 1994, อ้างถึงโดย สุมาลี มีสกุล, 2558, หน้า 25) ดังต่อไปนี้

- 1.1 กำหนดเป้าหมายของการประเมิน
- 1.2 ระบุคุณลักษณะที่เกี่ยวข้องในขอบข่ายของสิ่งที่ต้องการประเมิน
- 1.3 พัฒนาโมเดลการคิดที่ต้องใช้ในการหาคำตอบของปัญหาและกำหนดความสำคัญของโมเดลการคิดที่พัฒนาขึ้น
- 1.4 สร้างข้อคำถามตามโมเดลการคิด
- 1.5 ประเมินโมเดลการคิดในเชิงประจักษ์ด้วยการนำแบบทดสอบไปทดลองใช้
- 1.6 คัดเลือกข้อคำถามเก็บไว้โดยพิจารณาจากความซับซ้อนของการใช้ความคิดในการแก้ปัญหา
- 1.7 ตรวจสอบความเที่ยงตรงของโมเดลการวัดที่ได้จากการทดสอบด้วยการตรวจสอบ Nomothetic span หมายถึง ความเที่ยงตรงและความสัมพันธ์ระหว่างภาวะสันนิษฐานซึ่งถูกวัดโดยการประเมิน (เช่น ความหตหุใจ ความสามารถทางคณิตศาสตร์) กับทฤษฎีที่เกี่ยวข้องกับภาวะสันนิษฐานนั้น

จากที่กล่าวมา สรุปได้ว่ากรอบกระบวนการออกแบบจากระบบการคิด (CDS) คือ ชุดของขั้นตอนที่จำเป็นในการรวมบทบาทสำคัญของทฤษฎีทางปัญญาในการพัฒนาแบบทดสอบ ซึ่งเป็นการออกแบบทางพุทธิปัญญาที่เป็นภาวะสันนิษฐานเบื้องต้นสำหรับกรอบกระบวนการที่เป็นระบบโดยใช้ทฤษฎีทางพุทธิปัญญามาอธิบายและถูกนำเสนอด้วยโมเดลทางคณิตศาสตร์ที่เชื่อมโยงข้อสอบทดสอบ สิ่งเร้าหรือสถานการณ์ของข้อสอบ เงื่อนไขการนำเสนอ หรือประสบการณ์เดิมกับการแสดงออกถึงภาวะสันนิษฐานเหล่านี้ผ่านการทำแบบทดสอบในการประเมินเพื่อวินิจฉัย จากความสำคัญดังกล่าว ในการศึกษาครั้งนี้ ผู้วิจัยได้สร้างแบบสอบวินิจฉัยทางพุทธิปัญญา โดยอิงกรอบแนวคิดของหลักการออกแบบการทดสอบเชิงวินิจฉัยแบบซีดีเอส (Cognitive Design System: CDS)

ของเอมเบรทสัน (Embretson, Susan E. 1994, p. 110) ซึ่งเป็นวิธีการที่ให้ความตรงเชิงโครงสร้างสูง ประกอบด้วย 6 ขั้นตอน ดังนี้

1) กำหนดจุดมุ่งหมายของการวัดให้ชัดเจน โดยระบุสิ่งที่มุ่งวัดซึ่งมีลักษณะเป็นตัวแปรเชิงโครงสร้าง จึงต้องมีการนิยามให้ชัดเจน ซึ่งในงานวิจัยฉบับนี้มุ่งวัดความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ พร้อมทั้งระบุถึงโครงสร้างอื่นที่เกี่ยวข้องในเชิงทฤษฎีกับสิ่งที่มุ่งวัด (nomothetic span)

2) ระบุคุณลักษณะที่เกี่ยวข้องกับสิ่งที่ต้องการวินิจฉัย โดยระบุทั้งรูปแบบหรือเงื่อนไขต่าง ๆ และคุณลักษณะย่อย (attributes) หรือทักษะ (skills) ซึ่งในงานวิจัยฉบับนี้กำหนดสิ่งที่มุ่งวินิจฉัยในการวิจัยนี้ คือ ความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษ ประกอบด้วยคุณลักษณะย่อย 6 คุณลักษณะคือ มีความรู้คำศัพท์ (Vocabulary knowledge) ระบุรายละเอียดที่ต้องการซึ่งปรากฏชัดเจน (identifying explicit information) มีความรู้ทางไวยากรณ์ (Syntactic knowledge) ระบุใจความสำคัญของเรื่อง (Generalization Main Idea) การสรุปความ (Getting Inference) และการเชื่อมโยงและสังเคราะห์ข้อมูล (Connecting and Synthesizing information)

3) พัฒนาโมเดลพุทธิปัญญา (cognitive model) โดยเริ่มจากการศึกษาทฤษฎีที่เกี่ยวข้องกับสิ่งที่มุ่งวัด โดยพิจารณาความสัมพันธ์เชิงโครงสร้างของคุณลักษณะย่อย แล้วกำหนดเป็นโมเดลทางพุทธิปัญญาของสิ่งที่มุ่งวัด และกำหนดเมตริกซ์คิว (Q-matrix) ซึ่งเป็นเมตริกซ์แสดงความสัมพันธ์ระหว่างข้อคำถามและคุณลักษณะที่ใช้ในการตอบข้อคำถามนั้น ๆ แล้วตรวจสอบคุณภาพโมเดลโดยการเสนอผู้เชี่ยวชาญพิจารณา

4) สร้างข้อสอบตามโมเดลพุทธิปัญญาและเมตริกซ์คิวที่พัฒนาขึ้น โดยมีลักษณะเป็นกลุ่มของข้อสอบที่มีสิ่งเร้าร่วมกันเป็นสถานการณ์/บทความ (passages) หรือข้อสอบแบบเทสต์เลท (Testlet test) ซึ่งผู้วิจัยสร้างข้อสอบด้วยตนเอง โดยกำหนดขอบเขตเนื้อหาตามกรอบอ้างอิงความสามารถทางภาษา CEFR ระดับ A2

5) ประเมินโมเดลพุทธิปัญญาที่สร้างขึ้นด้วยหลักฐานเชิงประจักษ์ โดยการนำแบบทดสอบไปทดลองใช้

6) คัดเลือกข้อคำถามที่ผ่านกระบวนการตรวจสอบคุณภาพ เพื่อจัดระเบียบเรียงเป็นฉบับ และนำไปใช้จริง

2.8 การกำหนดคุณลักษณะของโมเดลการวินิจฉัยเชิงพุทธิปัญญา

คุณลักษณะเป็นตัวแปรที่นำเข้าที่สำคัญในการประเมินเพื่อวินิจฉัยทางการศึกษา เพราะให้การสรุปอ้างอิงถึงทักษะเชิงพุทธิปัญญาของผู้สอบ แต่จากการศึกษาเอกสารที่เกี่ยวข้องกับการนำการประเมินเพื่อวินิจฉัยไปใช้ในบริบททางการศึกษาส่วนใหญ่ไม่ได้มีการให้คำจำกัดความคุณลักษณะไว้อย่างชัดเจน แต่จะมีการให้คำจำกัดความที่แตกต่างกันไปตามบริบทของการนำไปใช้ นักวิจัยบางคน

กล่าวถึงคุณลักษณะในความหมายทั่ว ๆ ไป โดยพิจารณาทักษะเชิงพุทธิปัญญาที่จำเป็นต้องใช้ในการปฏิบัติงาน บางคนกล่าวถึงคุณลักษณะในความหมายเฉพาะในลักษณะการรวมองค์ประกอบย่อยเชิงพุทธิปัญญาเพื่อใช้ในการแก้ปัญหาในเนื้อหาเฉพาะคุณลักษณะจึงถูกนำไปใช้ทั้งในความหมายเฉพาะและความหมายทั่วไป

Birenbaum Kelly and Tatsuoka (1993 cited in Gierl, Leighton & Hunka, 2000) ให้คำจำกัดความคุณลักษณะไว้ว่า “เป็นการบรรยายกระบวนการ ทักษะหรือความรู้ของนักเรียนที่ต้องใช้เพื่อให้สามารถทำงานได้บรรลุเป้าหมาย แต่ไม่จำกัดเฉพาะความสามารถของนักเรียนในบางครั้งคุณลักษณะอาจรวมถึงสิ่งที่นักเรียนใช้ช่วยในการเรียนรู้หรือใช้ในการแก้ปัญหา Gierl, Leighton and Hunka (2000, p. 89) กล่าวถึงคุณลักษณะไว้ว่า คุณลักษณะเป็นการบรรยายวิธีการหรือการอธิบายความรู้ที่ต้องใช้ในการทำงานที่มีขอบเขตเฉพาะ แม้ว่าคุณลักษณะจะไม่ใช่วิธีการแก้ปัญหาแต่คุณลักษณะก็จะช่วยสร้างวิธีการแก้ปัญหามากไปกว่านั้นชุดของคุณลักษณะจะถูกจัดเข้าสู่วิธีการแก้ปัญหาเพื่อสนับสนุนบทบาทในการแก้ปัญหาที่เกิดขึ้นคุณลักษณะเป็นเอกลักษณ์เฉพาะตนที่มีลักษณะเป็นพลวัตชุดคุณลักษณะของนักเรียนที่มีความสามารถในเวลาหนึ่งอาจไม่เหมือนชุดของคุณลักษณะของนักเรียนที่มีความสามารถในอีกเวลาหนึ่ง ทั้งนี้อาจเป็นเพราะพัฒนาการของนักเรียนหรือประสบการณ์ที่ได้รับจากการเรียนการสอน

2.8.1 การตั้งชื่อคุณลักษณะ

โดยทั่วไปคำว่า “คุณลักษณะแฝง (Latent characteristic)” จะหมายถึงองค์ประกอบทางจิตซึ่งเป็นสิ่งที่สนใจในทางทฤษฎีที่ไม่สามารถสังเกตได้สำหรับในบริบทของ DCMs จะใช้คำว่า “ตัวแปรแฝง (Latent variables)” แทนคำว่าลักษณะแฝง ส่วนคำว่า Latent trait แสดงให้เห็นถึงองค์ประกอบทางจิตที่สนใจซึ่งเชื่อว่าจะมีความคงที่ไม่เปลี่ยนแปลงเมื่อเวลาเปลี่ยนไปในขณะที่ Latent states สามารถเปลี่ยนแปลงได้เมื่อเวลาเปลี่ยนไป

คำว่า Element ใช้อ้างอิงถึงองค์ประกอบทางจิตซึ่งสร้างรูปแบบแนวคิดของสิ่งที่ใหญ่กว่าซึ่งเรียกว่ากระบวนการตอบสนองทางความคิด (Cognitive response process)

คำว่าทักษะ (Skill) และคุณลักษณะ (Attribute) อาจถูกนำมาใช้อยู่เสมอในวรรณกรรมที่เกี่ยวกับการวัดซึ่งเป็นการประเมินเพื่อวินิจฉัย คำเหล่านี้ถูกใช้มาเป็นระยะเวลาอันยาวนานในการวิเคราะห์องค์ประกอบโดยมีคำที่ใช้ซึ่งมีความหมายคล้ายกัน ได้แก่ Latent trait และ Latent characteristic

สำหรับในการศึกษาเกี่ยวกับการประเมินเพื่อวินิจฉัยนี้จะใช้คำว่าคุณลักษณะ (Attribute) แทนสิ่งที่ต้องการวินิจฉัย และใช้คำว่าแบบแผนคุณลักษณะ (Attribute profile) แทนคุณลักษณะเฉพาะของผู้ตอบแต่ละคน สำหรับการกำหนดชื่อตัวแปรแฝงหรือคุณลักษณะนั้น เป็นเรื่องปกติที่จะมีการกำหนดชื่อตัวแปรแฝงที่ต้องการศึกษาขึ้นเอง เนื่องจากไม่มีชื่อที่ถูกต้องหรือไม่ถูกต้อง

อย่างแท้จริงตัวแปรแฝงที่สนใจศึกษาท้ายที่สุดสิ่งที่สำคัญที่สุดคือการแปลผลตัวแปรแฝงจะต้องอยู่บนพื้นฐานทางทฤษฎีที่เป็นพื้นฐานและการตัดสินใจเกี่ยวกับการกำหนดลักษณะของผู้ตอบเพื่อที่จะให้ความเข้าใจเกี่ยวกับประเภทของคุณลักษณะและภาวะสันนิษฐานที่เกี่ยวข้องกันซึ่งถูกกำหนดในงานวิจัยที่เกี่ยวข้องกับ DCMs

2.8.2 การกำหนดนิยามของคุณลักษณะย่อย

การกำหนดนิยามของคุณลักษณะย่อย (Definitional grain size of the attribute) ที่ใช้ในขั้นตอนการคิดอย่างละเอียดเป็นสิ่งที่มีความจำเป็นในการประเมินเพื่อวินิจฉัยเนื่องจากจะช่วยให้ทราบคุณลักษณะที่เป็นต้องใช้สำหรับการแก้ปัญหาที่กำหนด และส่งผลทำให้สามารถวินิจฉัยว่าผู้ตอบมีความบกพร่องของคุณลักษณะย่อยที่ใช้ในขั้นตอนการคิดสำหรับปัญหาที่มีความยากและซับซ้อน ในการกำหนดคุณลักษณะย่อยควรมีกำหนดคุณลักษณะย่อยที่จำเป็นต้องใช้ในการแก้ปัญหาทั้งหมดซึ่งส่งผลทำให้มีจำนวนตัวแปรแฝงที่ใช้ในโมเดลมากขึ้น ซึ่งอาจทำให้ในการคำนวณด้วย DCMs ไม่สามารถทำได้จึงมีความจำเป็นที่จะต้องกำหนดจำนวนของคุณลักษณะให้อยู่ในขอบเขตที่จะสามารถประมาณค่าในทางสถิติได้ตามความยาวของข้อคำถามและขนาดของกลุ่มตัวอย่างที่ใช้

ด้วยเหตุนี้ทำให้ในบางครั้งการประเมินเพื่อวินิจฉัยที่ต้องประเมินการคิดที่ซับซ้อนมีความจำเป็นที่จะต้องใช้นิยามคุณลักษณะที่ละเอียดน้อยกว่า การนิยามคุณลักษณะและกระบวนการคิดอย่างหยابมักถูกใช้ในตารางกำหนดรายละเอียดของข้อสอบหรือพิมพ์เขียวสำหรับการประเมินทางการศึกษา การบรรยายขั้นตอนการคิดและพฤติกรรมในระดับที่ละเอียดจะถูกใช้ในการประเมินอิงมาตรฐานที่เป็นเครื่องมือในการตรวจสอบการบรรลุผลของหลักสูตรการสอนและการประเมิน กล่าวโดยสรุปความเหมาะสมของการกำหนดนิยามของคุณลักษณะขึ้นอยู่กับวัตถุประสงค์ของการประเมินเพื่อวินิจฉัย คุณลักษณะที่ต้องการศึกษาและระดับประสบการณ์ของกลุ่มบุคคลที่ต้องการประเมิน

2.8.3 การกำหนดรายละเอียดของคุณลักษณะ

การกำหนดรายละเอียดของคุณลักษณะ(The nature of the attribute specifications) ซึ่งเป็นสิ่งที่มีประโยชน์และจำเป็นมีอยู่ 3 สิ่ง คือชื่อของคุณลักษณะ คำนิยามของคุณลักษณะ และคำแนะนำในการกำหนดรหัสของคุณลักษณะ (Gierl, Leighton & Hunka, 2000, p. 102, อ้างถึงใน แคทรียา แสงดาวเทียน, 2562, หน้า 26) ดังนี้

- 1) คุณลักษณะ คือ คำ ข้อความ หรือกลุ่มคำ ที่แสดงความหมายหลักของคุณลักษณะ
- 2) คำนิยามของคุณลักษณะ คือ การบรรยายหรือข้อความสั้น ๆ ที่ให้รายละเอียดซึ่งใช้ในการแสดงด้านที่ต่างกันของคุณลักษณะ

3) คำแนะนำในการกำหนดรหัสของคุณลักษณะ คือ การเชื่อมโยงการประเมินเพื่อวินิจฉัยซึ่งช่วยผู้เชี่ยวชาญในการกำหนดว่าเมื่อใดที่คุณลักษณะนั้นจะถูกวัดโดยข้อคำถามที่สร้างขึ้น

2.8.4 การกำหนดเมตริกซ์ถ่วงน้ำหนัก หรือเมตริกซ์คิว (Q-matrix)

การนำโมเดลการวินิจฉัยเชิงพุทธิปัญญา (DCMs) ไปใช้ในทางปฏิบัติต้องมีการกำหนดเมตริกซ์ซึ่งเป็นส่วนประกอบที่สำคัญของการวิเคราะห์ข้อมูลที่มีชื่อว่า Q-matrix ซึ่ง Q-matrix นี้เป็นเมตริกซ์ที่แสดงถึงการเชื่อมโยงระหว่างทฤษฎีที่ใช้ในการออกแบบการประเมินเพื่อวินิจฉัยนอกเหนือไปจากนี้ยังแสดงให้เห็นถึงสมมติฐานในเชิงประจักษ์เกี่ยวกับโครงสร้างของคุณลักษณะที่ทำการทดสอบความถูกต้องของ Q-matrix เป็นสิ่งที่จำเป็นสำหรับการประมาณค่าที่แม่นยำของรูปแบบทักษะความสามารถของผู้สอบเมื่อใช้โมเดลการประเมินวินิจฉัยเชิงจำแนก (DCMs)

Q-matrix เป็นเมตริกซ์ที่มีแถวเป็นลำดับที่ของข้อคำถามและมีหลักเป็นคุณลักษณะที่ต้องการวินิจฉัยโดยมีสมาชิกในเมตริกซ์เพียง 2 ค่า คือ 0 และ 1 กำหนดให้ α_{ij} แทนสมาชิกใน Q-matrix α_{ij} มีค่าเป็น 1 เมื่อข้อคำถามที่ i ต้องใช้คุณลักษณะ j ในการแก้ปัญหา และ α_{ij} มีค่าเป็น 0 เมื่อข้อคำถามที่ i ไม่จำเป็นต้องใช้คุณลักษณะ j ในการแก้ปัญหการพัฒนา Q-matrix และกระบวนการตรวจสอบความเที่ยงตรงของ Q-matrix ถูกกำหนดด้วยกระบวนการในการพัฒนา 2 วิธีการหลัก คือ วิธีการในเชิงปริมาณและวิธีการในเชิงคุณภาพ

1) วิธีการในเชิงปริมาณ วิธีการนี้เป็นวิธีที่ใช้ความคิดเห็นของผู้เชี่ยวชาญในการพิจารณาว่าข้อคำถามที่กำหนดใช้วัดคุณลักษณะใดของผู้สอบ ตามนิยามของคุณลักษณะที่กำหนด

2) วิธีการในเชิงคุณภาพ วิธีการนี้เป็นการวิเคราะห์กระบวนการในการแก้ปัญหาของผู้สอบโดยการให้ผู้สอบแสดงกระบวนการคิดแก้ปัญหาในแบบฝึกหัดที่กำหนดโดยเฉพาะ (Specific tasks) ด้วยเทคนิคการคิดออกเสียง (Think-aloud protocol) ซึ่งเป็นวิธีที่ให้ผู้สอบแสดงสิ่งที่กำลังคิดในการถูกนำมาใช้ในการให้ข้อมูลในเชิงการพัฒนาเกี่ยวกับความสัมพันธ์ระหว่างข้อคำถามและทักษะเชิงพุทธิปัญญาและวิธีที่ใช้ในการแก้ปัญหา (Q-matrix) สำหรับบริบทของการประเมินเพื่อวินิจฉัยทางด้านพุทธิปัญญา (Cognitive diagnostic assessment: CDA) วิธีการในเชิงปริมาณที่เข้าร่วมกันในการพัฒนา Q-matrix ประกอบไปด้วยการวิเคราะห์หองค์ประกอบเชิงยืนยันการวิเคราะห์ความเป็นพหุมิติ และการตรวจการประมาณค่าพารามิเตอร์ของข้อคำถามที่เป็นผลมาจากการใช้โมเดลการประเมินเพื่อวินิจฉัยเชิงพุทธิปัญญา (DCMs)

การกำหนดเมตริกซ์ถ่วงน้ำหนัก หรือเมตริกซ์คิว (weight matrix: matrix Q) หรือโดยมากเรียก Q-matrix ใช้ในการระบุคุณลักษณะหรือทักษะที่ถูกวัดในแต่ละข้อว่าในแต่ละข้อนั้น ๆ วัด คุณลักษณะหรือทักษะใดบ้าง อีกนัยหนึ่งคือในแต่ละข้อนั้น ๆ ต้องการคุณลักษณะ (attributes)

หรือทักษะ (skills) ใดบ้าง ซึ่งมีลักษณะเดียวกันกับเมตริกซ์รูปแบบปัจจัย (factor pattern matrix) ที่จัดนำหน้าองค์ประกอบในการวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) (วรางคณา ก้อมน้อย, 2563, หน้า 69-70) ซึ่งในการกำหนด Q-matrix นี้ต้องผ่านการพิจารณาโดยผู้เชี่ยวชาญในสาขาวิชานั้น ๆ เป็นการตรวจสอบคุณภาพเบื้องต้น ประกอบด้วย 2 ขั้นตอน ดังนี้

ขั้นที่ 1 ผู้เชี่ยวชาญทำการแบ่งขอบเขตทั้งหมดของการสอบออกเป็นทักษะหรือคุณลักษณะที่ต้องใช้ในการทำแบบสอบนั้น ๆ ซึ่งทักษะย่อยเหล่านี้ใช้สัญลักษณ์แทนด้วย α_k , $k = 1, 2, \dots, K$ จากนั้นผู้วิจัยทำการสร้างข้อสอบ

ตัวอย่าง การศึกษาของโตเนกาโบนี และราวัน (Tonekaboni & Ravand, 2021, p. 60) ทำการศึกษากับตัวอย่าง 2,625 คน ได้ทำการระบุคุณลักษณะของความเข้าใจในการอ่านภาษาอังกฤษซึ่งถูกกำหนดขึ้นจากการตัดสินของผู้เชี่ยวชาญ 7 คุณลักษณะ ดังนี้

VOCAB: Vocabulary knowledge (ความรู้คำศัพท์)

INF: Inferencing (การสรุปความ)

BSM: Building situation model of a text (สร้างรูปแบบสถานการณ์)

TSK: Text structure knowledge (ความรู้เกี่ยวกับโครงสร้างข้อความ)

ICM: Identifying and constructing main idea (ระบุสาระสำคัญของเรื่อง)

IRT: Identifying relevant text material to answer a question (ระบุส่วนประกอบที่เป็นประเด็นหลักซึ่งใช้ในการตอบคำถาม)

SI: Synthesizing information from text (สังเคราะห์ข้อมูลจากบทอ่าน)

ขั้นที่ 2 ผู้เชี่ยวชาญระบุว่าแบบสอบแต่ละข้อ ผู้ตอบต้องมีคุณลักษณะใดบ้าง จึงจะสามารถตอบข้อคำถามข้อนั้น ๆ ถูก โดยจากตัวอย่างข้างต้น ผลการพิจารณาของผู้เชี่ยวชาญเป็น ดังภาพ 12

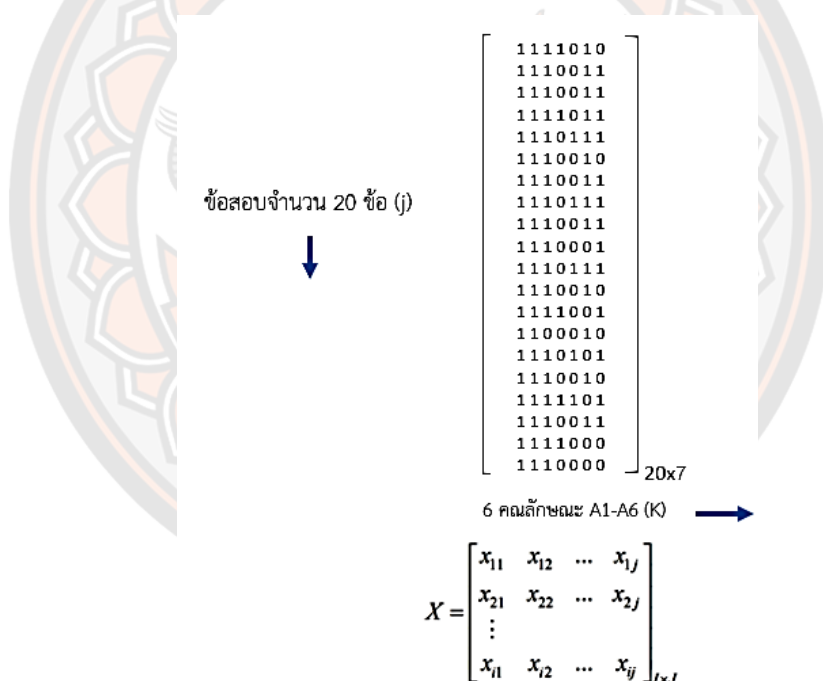
Item	VOCAB	INF	BSM	TSK	ICM	IRT	SI
1	1	1	1	1	0	1	0
2	1	1	1	0	0	1	1
3	1	1	1	0	0	1	1
4	1	1	1	1	0	1	1
5	1	1	1	0	1	1	1
6	1	1	1	0	0	1	0
7	1	1	1	0	0	1	1
8	1	1	1	0	1	1	1
9	1	1	1	0	0	1	1
10	1	1	1	0	0	0	1
11	1	1	1	0	1	1	1
12	1	1	1	0	0	1	0
13	1	1	1	1	0	0	1
14	1	1	0	0	0	1	0
15	1	1	1	0	1	0	1
16	1	1	1	0	0	1	0
17	1	1	1	1	1	0	1
18	1	1	1	0	0	1	1
19	1	1	1	1	0	0	0
20	1	1	1	0	0	0	0

ภาพ 12 แสดงผลการพิจารณาเมตริกซ์คิวของผู้เชี่ยวชาญ

จากภาพ 12 แปลงเป็นเมตริกซ์คิว (Q-matrix) ซึ่งเป็นเมตริกซ์ที่มีขนาด J แถว และมี K หลัก โดยที่ J คือ จำนวนข้อคำถามของแบบสอบ และ K คือ จำนวนทักษะ (skills) หรือคุณลักษณะ (attributes) ที่ต้องใช้ในการตอบข้อคำถามข้อนั้น ๆ ถูก หรือเขียนได้เป็น

$$Q = \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1k} \\ q_{21} & q_{22} & \dots & q_{2k} \\ \vdots & \vdots & \dots & \vdots \\ q_{j1} & q_{j2} & \dots & q_{jk} \end{bmatrix}_{J \times K}$$

โดยที่ $q_{jk} = 1$ เมื่อ ข้อสอบข้อที่ j ต้องการใช้ทักษะหรือคุณลักษณะ k และในทิศทางตรงกันข้าม $q_{jk} = 0$ เมื่อข้อสอบข้อที่ j ไม่ต้องการใช้ทักษะหรือคุณลักษณะ k ดังนั้นจัดเตรียมเป็นเมตริกซ์คิว ได้ดังภาพ 13



ภาพ 13 แสดงตัวอย่างเมตริกซ์คิวของข้อสอบ 20 ข้อ ที่ใช้วัดคุณลักษณะ 6 คุณลักษณะ

$$\begin{bmatrix} 01110110010011100101 \\ 11100111101010110101 \\ \vdots \\ 10111011000010111011 \end{bmatrix}_{2625 \times 20}$$

เมื่อ $x_{jk} = 1$ เมื่อ ผู้ตอบ ตอบแบบสอบข้อนั้น ๆ ถูก และ $x_{jk} = 0$ เมื่อผู้ตอบตอบข้อคำถามนั้น ไม่ถูกต้อง ซึ่งจากตัวอย่างของ Tonekaboni & Ravand (2021) ข้างต้น มีจำนวน

ผู้ตอบแบบสอบ 2,625 คน ($i = 2,625$) และ แบบสอบมี 20 ข้อ ($j=20$) จัดเตรียมเป็น matrix X ขนาด 2625 แถว 20 หลัก ได้ดังนี้

ขั้นที่ 3 นำข้อมูลที่ได้จากการตอบแบบสอบของผู้สอบแต่ละคน (Response data: matrix X) และเมตริกซ์คิว (matrix Q) นำเข้าไปวิเคราะห์ด้วยโมเดลจีดีไอเอ็นเอ (GDINA) ในโปรแกรม R

ต่อจากนี้ เป็นการตรวจสอบความถูกต้องของเมตริกซ์เชิงประจักษ์รอบแรกใช้ Q-matrix อันล่าสุดใช้วัดผลทั้งหมดเจ็ดคุณลักษณะ ข้อมูลการตอบสนองของผู้สอบจำนวน 2,625 คน และแพ็คเกจ GDINA เวอร์ชัน 2.7.4 (Ma & de la Torre, 2018) ในโปรแกรม R (R Core Team, 2019) นอกจากนี้ค่าของ LR, AIC และ BIC ยังถูกพิจารณาสำหรับการตรวจสอบเชิงประจักษ์ตามวัตถุประสงค์วิจัยในภาพรวมขั้นตอนนี้ส่งผลให้มีการขจัดออกและนำคุณลักษณะ IRT “identifying relevant text material to answer a question” หรือ ระบุส่วนประกอบที่เป็นประเด็นหลักซึ่งใช้ในการตอบคำถาม ถูกลำเอียงจาก Q-matrix ผลการประเมิน Q-matrix ครั้งที่สองของคณะกรรมการ ซึ่งเป็นผู้เชี่ยวชาญได้มีการจัดเตรียมบทอ่านและข้อความสำหรับใช้ในแบบสอบ และรายการคุณลักษณะพร้อมกับนิยามและการยกตัวอย่างประกอบเพื่อจัดลำดับบน Q-matrix อย่างอิสระ จากนั้นจึงได้ Q-matrix ที่เป็นเอกฉันท์จากการประเมินหลักฐานเชิงประจักษ์ของผู้เชี่ยวชาญที่ถูกสร้างขึ้นสำหรับการตรวจสอบและแก้ไขในรอบต่อไป ต่อมา มีการตรวจสอบความตรง Q-matrix ซึ่งในรอบนี้มีการพิจารณาค่า LR, AIC และ BIC พร้อมกับ mesa plots ซึ่ง mesa plots นี้ถูกนำเสนอครั้งแรกโดย de la Torre และ Ma (2016) เพื่อแสดงภาพเวกเตอร์ q ที่ดีที่สุดสำหรับข้อความรายข้อ โดยที่ q -vector ที่อยู่บนขอบเขตของ mesa เป็น q -vector ที่ดีที่สุด ดังนั้น ในการศึกษากระบวนการสร้างและการตรวจสอบความถูกต้องของ Q-matrix ซึ่งผ่านการแก้ไขหลายครั้งทำให้ได้คุณลักษณะ 6 ประการ ได้แก่ Vocabulary ความรู้ (VOCAB), การอนุมาน (INF), สร้างแบบจำลองสถานการณ์ (BSM), โครงสร้างข้อความความรู้ (TSK) การระบุและสร้างแนวคิดหลัก (ICM) และการสังเคราะห์ข้อมูลจาก Text (SIT) ดังที่แสดงไว้ในตารางที่ 12

ตาราง 12 แสดงตัวอย่างผลการตรวจสอบคุณภาพของเมตริกซ์คิวทั้งสองครั้ง

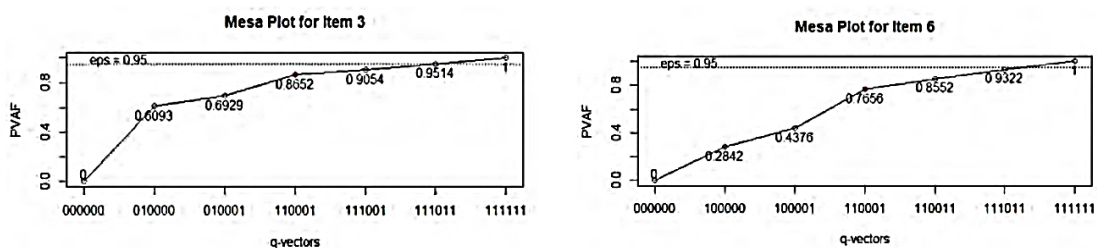
คุณลักษณะ	คำนิยาม	ข้อความ	จำนวนข้อ
VOCAB:	ความสามารถในการสร้างความหมายของคำโดย	1,2,3,4,5,6,	14
Vocabulary	ใช้พื้นความรู้เดิม หรือการเดาความหมายจาก	7,9,10,12,1	
knowledge	เงื่อนไขทางภาษา และบริบท	3,14,17,19	
(ความรู้คำศัพท์)			

คุณลักษณะ	คำนิยาม	ข้อคำถาม	จำนวนข้อ
INF: Inferencing (การสรุปความ)	ความสามารถในการสมมติหรือสร้างการเชื่อมต่อ (ระหว่างข้อความและองค์ประกอบตามบริบท) ซึ่งไม่ได้ระบุไว้อย่างชัดเจนโดยอาศัยความรู้เดิม มากกว่าข้อมูลที่ไม่ได้อยู่ในข้อความหรือไม่ใช่โดยอัตโนมัติ/เชิงกลยุทธ์โดยข้อมูลที่รวมกันอยู่ในประโยคหรือความสามารถในการสมมติหรือทำให้การเชื่อมต่อ (ระหว่างองค์ประกอบที่เป็นข้อความและบริบท)	2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 16, 17, 18, 19, 20	14
BSM: Building situation model of a text (สร้างรูปแบบสถานการณ์)	ความสามารถในการประมวลผลและบูรณาการแนวคิดจากข้อความและแนวคิดที่เกี่ยวข้องของโดเมนและโลกทั่วไปความรู้เพื่อสร้างแผนภาพซึ่งเป็นตัวแทนของเนื้อหาของข้อความ	4, 7, 8, 9, 11, 12, 15, 16, 17, 19, 20	11
TSK: Text structure knowledge (ความรู้เกี่ยวกับโครงสร้างข้อความ)	ความสามารถในการแยกแยะว่าข้อความมีโครงสร้างอย่างไร โดยใช้พื้นความรู้เดิม หรือความรู้ทางวากยสัมพันธ์ และความรู้เกี่ยวกับความสัมพันธ์ระหว่างและระหว่างประโยคและย่อหน้าเพื่ออนุมานเกี่ยวกับข้อความเพื่อจัดระเบียบเนื้อหาและสร้างการแสดงแทนเนื้อหาข้อความ	1, 13, 17, 20	4
ICM: Identifying and constructing main idea (ระบุสาระสำคัญของเรื่อง)	ความสามารถในการระบุและสร้างส่วนสำคัญของย่อหน้าแนวคิดหลักหรือชื่อเรื่องของข้อความ	5, 7, 8, 11, 13, 15, 17, 18, 20	9
SIT: Synthesizing information from text (สังเคราะห์ข้อมูล)	ความสามารถในการสร้างและทำขึ้นมาใหม่โดยเจตนาอย่างต่อเนื่องการเชื่อมโยง (การเชื่อมต่อหรือการเชื่อมโยง) อนุมานว่าเชื่อมโยงกลับไปยังประโยคและความคิดก่อนหน้าเพื่อที่จะสร้างความหมายของข้อความ	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 13, 14, 15, 18	14

คุณลักษณะ	คำนิยาม	ข้อคำถาม	จำนวนข้อ
จากบทอ่าน)			

ขั้นตอนที่ 1 ทำการตรวจสอบ Q-matrix

ในการตรวจสอบ Q-matrix เชิงประจักษ์ โมเดล G-DINA ที่มีแอตทริบิวต์อิมัลชัน การกระจายถูกปรับให้เข้ากับข้อมูลที่ใช้ข้อจำกัดแบบโมโนโทนิคและฟังก์ชัน Qval ผลลัพธ์แสดงการแก้ไข 6 รายการสำหรับรายการที่ 3, 6, 11 และ 17 เพื่อตรวจสอบเพิ่มเติมความเป็นไปได้ของการปรับเปลี่ยนที่เสนอตามรายการเฉพาะของ de la Torre และ Chiu (2016) ดัชนีอำนาจจำแนก ζ^2 จากนั้น mesa plots ที่สอดคล้องกันสำหรับข้อคำถาม 3, 6, 11, และ 17 ถูกนำมาพิจารณา mesa plots ใช้เพื่อระบุ q-vector ที่ดีที่สุดสำหรับข้อคำถามแต่ละข้อคล้ายกับ screen plot ในการวิเคราะห์ห้องประกอบและแผนภูมิเส้นโดยที่แกน x แทนเวกเตอร์ q ที่มีสัดส่วนความแปรปรวนสูงสุด (PVAFs) สำหรับจำนวนแอตทริบิวต์ที่จำเป็นที่แตกต่างกันและแกน y ที่นำเสนอ ค่า PVAF (proportion of variance accounted for) ค่าสัดส่วนความแปรปรวนคือความแปรปรวนในตัวแปรทั้งหมดที่คิดด้วยปัจจัยเดียวที่เป็นลักษณะทั่วไป กล่าวอีกนัยหนึ่งคือค่าลักษณะเฉพาะ/จำนวนตัวแปร = ค่าสัดส่วนความแปรปรวน ยิ่งผลรวมของความแปรปรวน (SS) สูง (เช่น ค่าลักษณะเฉพาะ) องค์ประกอบหรือปัจจัยที่มากกว่าจะช่วยอธิบายความแปรปรวนในตัวแปรได้มากขึ้นเท่านั้น (Ma, 2019) จุดทึบสีแดงแสดงถึง q-vector ดั้งเดิม q-vector ที่ถูกต้องสำหรับข้อคำถามที่กำหนดคืออันที่ดีที่สุดขอบเขต mesa (de la Torre & Ma, 2019). ความจริงแล้ว q-matrix ที่ถูกต้องคือค่าที่ได้จากโมเดลประหยัด (parsimonious) ที่ให้ค่าดัชนีอำนาจจำแนกเฉพาะของข้อสอบ (item-specific discrimination index) สูงสุด ζ^2 โดยประมาณแทนที่จะเป็น q-vector ที่สร้าง ζ^2 สูงสุด (de la Torre & Akbay, 2019) parsimonious q-vector คือค่าที่ประมาณ ζ^2 สูงสุดด้วยคุณลักษณะ (Attributes) ที่กำหนดไว้ในจำนวนที่น้อยที่สุดผลการศึกษาของ Tonekaboni & Ravand (2021, p.58-87) พบว่าเฉพาะ q-vector ดั้งเดิมสำหรับข้อคำถามข้อที่ 3 และ 6 เท่านั้นที่มี PVAF น้อยกว่า .95 แสดงว่าอาจต้องตรวจสอบเพิ่มเติม แสดงตัวอย่างค่า PVAF ในภาพที่ 14



ภาพ 14 แสดงค่า PVAF ของข้อคำถามข้อที่ 3 และ 6 ก่อนปรับ

นอกจากนี้สถิติการทดสอบ Wald หรือ Wald Test (de la Torre & Lee, 2013) สำหรับคุณลักษณะเฉพาะทั้งหมดสำหรับข้อคำถามแต่ละข้อถูกคำนวณ ในกรณีที่มีสมมติฐานศูนย์ (คุณลักษณะเฉพาะที่ถูกวัดด้วยข้อสอบหนึ่งข้อ) ถูกปฏิเสธที่ $p < .05$ โมเดลลดรูปจะถูกปฏิเสธ ในกรณีของการเก็บรักษาโมเดลลดรูปมากกว่าหนึ่งแบบไว้ในระบบประมวลผลของโปรแกรมและการมีอยู่ของโมเดล DINA, DINO, DINA หรือ DINO ด้วยค่านัยสำคัญทางสถิติ (p value) ที่ใหญ่ที่สุดยังคงอยู่ มิฉะนั้น โมเดลลดรูปอื่นๆ ที่คงไว้ซึ่งการมีนัยสำคัญทางสถิติ (p value) ที่ใหญ่ที่สุดค่าถูกเลือก เหตุผลสำหรับการเลือกตั้งกล่าวต้องทำด้วยวิธีการทางสถิติลดความซับซ้อนให้เหลือน้อยที่สุด โมเดลวินิจฉัยเชิงพหุปัญญา (DCMs) ที่ถูกลดรูปแล้วจะเป็นที่ต้องการมากกว่าโมเดลวินิจฉัยเชิงพหุปัญญา (DCMs) เฉพาะอื่นๆ (Rupp & Templin, 2008) ดังนั้น ในขั้นตอนสุดท้ายการพิจารณาค่าทางสถิติเหล่านี้รวมทั้ง mesa plots และผลการประเมิน Q-matrix จากคณะผู้เชี่ยวชาญได้ผลสรุปออกมาดังภาพที่ 15

Item	VOCAB	INF	BSM	TSK	ICM	SI
1	1	0	0	1	0	1
2	0	1	0	0	0	1
3	1	1	0	0	0	0
4	1	1	1	0	0	1
5	1	1	0	0	1	1
6	1	1	0	0	0	1
7	1	0	1	0	1	1
8	0	0	1	0	1	1
9	1	1	1	0	0	0
10	1	1	0	0	0	1
11	0	1	1	0	1	1
12	1	1	1	0	0	0
13	1	1	0	1	1	1
14	1	0	0	0	0	1
15	0	0	1	0	1	1
16	0	1	1	0	0	0
17	1	1	1	1	1	0
18	0	1	0	0	1	1
19	1	1	1	0	0	0
20	0	1	1	1	1	0

ภาพ 15 แสดง Final Q-matrix

แบบสอบวินิจฉัยการอ่านของโตนิกาโบนี และราวัน (Tonekaboni & Ravand, 2021, p. 76) ถูกกำหนดจากเทคนิคการคิดออกเสียง (Think aloud: TA) ของผู้ให้ข้อมูล (participants) 10 คน และการตัดสินของผู้เชี่ยวชาญวิเคราะห์ข้อมูลทางสถิติด้วยโปรแกรมอาร์ ด้วย G-DINA package ได้ 6 คุณลักษณะดังนี้

A1 หมายถึง ความรู้คำศัพท์ (Vocabulary knowledge)

A2 หมายถึง การสรุปความ (Inferencing)

A3 หมายถึง สร้างรูปแบบสถานการณ์ (Building situation model)

A4 หมายถึง ความรู้เกี่ยวกับโครงสร้างข้อความ (Text structure knowledge)

A5 หมายถึง ระบุส่วนที่เป็นประเด็นหลัก (Identifying and constructing the main idea)

A6 หมายถึง สังเคราะห์ข้อมูลจากบทอ่าน (Synthesizing information from text)

โปรแกรม R มีฟังก์ชันทั้งสำหรับโมเดลการวินิจฉัยทางพุทธิปัญญา (cognitive diagnosis modeling) และโมเดลการตอบสนองต่อข้อสอบแบบพหุมิติ (multidimensional item response modeling) สำหรับข้อมูลแบ่งสองค่า (dichotomous data) และข้อมูลแบ่งหลายค่า (polytomous data) ในแพ็คเกจ CDM สามารถใช้ประมาณค่าโมเดลได้หลายโมเดล ทั้ง DINA model, DINO model, G-DINA model แบบตอบหลายค่า (polytomous), DINA model แบบหลายตัวเลือก (multiple choice), GDM, โมเดลการตอบสนองต่อข้อสอบแบบชดเชยได้เชิงเส้นพหุมิติ (multidimensional linear compensatory item response model) (Alexander Robitzsch, Thomas Kiefer, Ann Cathrice George, and Ali Uenlue, 2016, อ้างถึงใน วารสารคณา ก้อนน้อย, 2563, หน้า 74-77) และ ในโปรแกรม R มีคำสั่งในการขอ output หลายคำสั่ง ผู้วิจัยขอยกตัวอย่างบางคำสั่งที่เกี่ยวข้อง โดยตรง ดังตาราง 13

ตาราง 13 แสดงตัวอย่างคำสั่งขอ output ในโปรแกรม R

คำสั่ง	คำอธิบาย	โมเดล		
		DINA	GDINA	GDM
Summary (ชื่อโมเดล)	ภาพรวมของโมเดล	✓	✓	✓
Plot (ชื่อโมเดล)	แสดงกราฟ	✓	✓	✓
Coef (ชื่อโมเดล) หรือ ชื่อโมเดล\$coef	พารามิเตอร์ของโมเดล	✓	✓	✓
cdm.est.class.accuracy (ชื่อ โมเดล)	ความตรงของการจำแนกและความเที่ยง ของการจำแนก	✓	✓	
ชื่อโมเดล \$pattern	รูปแบบทักษะส่วนบุคคล	✓	✓	
ชื่อโมเดล \$skill.patt	ความน่าจะเป็นในการรอบรู้แต่ละทักษะ	✓	✓	
ชื่อโมเดล \$attribute.patt	การแจกแจงการจำแนกของทักษะ	✓	✓	

ที่มา: George, Robitzsch, Kiefer, GroB and Ünlü (2016, อ้างถึงใน วรางคณา ก้อมน้อย (2563, หน้า 74)

วรางคณา ก้อมน้อย (2563, หน้า 74) แนะนำว่า ในการวิเคราะห์ข้อมูลด้วยโปรแกรม R จะแบ่งเป็น 3 ส่วน คือ ส่วนที่ 1 การเขียนคำสั่งขอสารสนเทศรายบุคคล ส่วนที่ 2 การเขียนคำสั่งขอสารสนเทศในภาพรวม และส่วนที่ 3 การเขียนคำสั่งขอสารสนเทศของแบบสอบ มีตัวอย่างคำสั่งและการแปลความหมาย ดังนี้

ส่วนที่ 1 การเขียนคำสั่งขอสารสนเทศรายบุคคล `>fsgmodel$pattern` แล้วกด enter จะได้ output ที่ให้ผลการวิเคราะห์

- pattern หมายถึง รูปแบบความรอบรู้
 - mle.est หมายถึง ผลจากการประมาณค่าพารามิเตอร์ความสามารถผู้สอบด้วยวิธี Maximum likelihood estimation (MLE) ว่าผู้สอบมีความรอบรู้ในคุณลักษณะใดบ้าง
 - mle.post หมายถึง ผลจากการประมาณค่าพารามิเตอร์ความสามารถผู้สอบด้วยวิธี Maximum likelihood estimation (MLE) ว่ามีโอกาสที่ผู้สอบจะมีความรอบรู้ร้อยละเท่าใด
 - map.est หมายถึง ผลจากการประมาณค่าพารามิเตอร์ความสามารถผู้สอบด้วยวิธี Maximum a posteriori (MAP) ว่าผู้สอบมีความรอบรู้ในคุณลักษณะใดบ้าง
 - map.post หมายถึง ผลจากการประมาณค่าพารามิเตอร์ความสามารถผู้สอบด้วยวิธี Maximum a posteriori (MAP) ว่ามีโอกาสที่ผู้สอบจะมีความรอบรู้ร้อยละเท่าใด
 - post.attr1 หมายถึง ผู้สอบมีโอกาสที่จะมีความรอบรู้ในคุณลักษณะที่ 1 ร้อยละเท่าใด
- สารสนเทศข้างต้นจะทำให้นักเรียนทราบว่าตนเองมีจุดอ่อน (weakness) หรือยังไม่รอบรู้ (not mastered) และจุดแข็ง (strange) หรือ มีความรอบรู้ (mastered) ในทักษะหรือ คุณลักษณะย่อยใดเพื่อประโยชน์ในการพัฒนาตนเองได้ตรงจุดอ่อนของตน

ส่วนที่ 2 การเขียนคำสั่งขอสารสนเทศในภาพรวม `>fsgmodel$skill.patt` แล้วกด enter จะได้ output

นักเรียนทั้งหมดมีโอกาสรอบรู้ (mastered) ในแต่ละทักษะร้อยละเท่าไร สารสนเทศข้างต้นจะทำให้ครูทราบว่า ในภาพรวมนักเรียนของตนมีจุดอ่อน (weakness) หรือ ยังไม่รอบรู้ (not mastered) และจุดแข็ง (strange) หรือ มีความรอบรู้ (mastered) ในทักษะหรือ คุณลักษณะย่อยใด เพื่อประโยชน์ในการจัดการส่งเสริมพัฒนานักเรียน ได้ตรงกับทักษะย่อยที่นักเรียนยังบกพร่อง

ส่วนที่ 3 การเขียนคำสั่งขอสารสนเทศของแบบสอบ `>cdm.est.class.accuracy(fsgmodel)` แล้วกด enter จะได้ output ที่แสดงความตรงและความเที่ยง ตัวอย่าง output แสดงดังภาพ 16

```
> cdm.est.class.accuracy(fsgmodel)
```

		P_a	P_c	
การประมาณค่าการจำแนกด้วยวิธี Maximum likelihood estimation	MLE	0.263	0.643	ความเที่ยงในการจำแนก
	MAP	0.568	0.767	
	MAP_Skill11	0.661	0.699	ความตรงในการจำแนก
	MAP_Skill12	0.920	0.854	
MAP_Skill13	0.388	0.876		
MAP_Skill14	0.949	0.912		
การประมาณค่าการจำแนกด้วยวิธี Maximum a posteriori	MAP_Skill15	0.863	0.780	
	MAP_Skill16	0.528	0.507	
	MAP_Skill17	0.938	0.893	
	MAP_Skill18	0.839	0.731	

ภาพ 16 แสดงตัวอย่างผลการวิเคราะห์สารสนเทศของแบบสอบ
(วารางคณา ก้อนน้อย, 2563, หน้า 77)

จากภาพ 16 แปลผลได้ว่า แบบสอบฉบับนี้ มีความตรงในการจำแนกเท่ากับ .263 ด้วยวิธีการ ประมาณค่าแบบ MLE และเท่ากับ .563 ด้วยวิธีการประมาณค่าแบบ MAP และแบบสอบฉบับนี้ มีความเที่ยงในการจำแนกเท่ากับ .643 ด้วยวิธีการประมาณค่าแบบ MLE และเท่ากับ .767 ด้วย วิธีการประมาณค่าแบบ MAP ซึ่งเมื่อพิจารณาแต่ละทักษะ พบว่า ทักษะที่ 1 มีความตรงในการจำแนกเท่ากับ .661 และมีความเที่ยงในการจำแนกเท่ากับ .699 เป็นต้น (วารางคณา ก้อนน้อย, 2563, หน้า 77)

การขอค่าความน่าจะเป็นในการตอบข้อสอบถูก > `fsgmodel$probitem` แล้วกด enter

โมเดลที่ใช้

ข้อสอบใช้คุณลักษณะ

คุณลักษณะที่คาดหวั่ง

ความน่าจะเป็นในการตอบถูก

```
> fsgmodel$probitem
```

	item	rule	partype.attr	itemno	skillcomb	prob
1	Item1	GDINA	alpha4-alpha6-alpha7	1	A000	1.110223e-16
2	Item1	GDINA	alpha4-alpha6-alpha7	1	A100	1.000000e+00
3	Item1	GDINA	alpha4-alpha6-alpha7	1	A010	2.451475e-02
4	Item1	GDINA	alpha4-alpha6-alpha7	1	A001	7.771561e-16
5	Item1	GDINA	alpha4-alpha6-alpha7	1	A110	2.944379e-01
6	Item1	GDINA	alpha4-alpha6-alpha7	1	A101	7.946157e-01
7	Item1	GDINA	alpha4-alpha6-alpha7	1	A011	2.816541e-01
8	Item1	GDINA	alpha4-alpha6-alpha7	1	A111	1.000000e+00
9	Item2	GDINA	alpha4-alpha7	2	A00	1.731304e-02
10	Item2	GDINA	alpha4-alpha7	2	A10	5.580272e-01
11	Item2	GDINA	alpha4-alpha7	2	A01	1.110223e-16
12	Item2	GDINA	alpha4-alpha7	2	A11	9.836602e-01
13	Item3	GDINA	alpha4-alpha7	3	A00	0.000000e+00
14	Item3	GDINA	alpha4-alpha7	3	A10	3.014730e-01

ภาพ 17 แสดงตัวอย่างค่าความน่าจะเป็นในการตอบข้อสอบถูก
(วารางคณา ก้อนน้อย, 2563, หน้า 78)

จากภาพ 17 ข้อสอบข้อที่ 1 ต้องใช้ทักษะที่ 4 6 และ 7 ในการตอบถูก แต่จากผล การทำข้อสอบพบว่า เมื่อผู้ตอบมีหรือไม่มีทักษะในรูปแบบต่าง ๆ จะมีความน่าจะเป็นในการตอบถูก ดังนี้

A000 หมายถึง ผู้ตอบที่ไม่มีคุณลักษณะใด ๆ มีโอกาสตอบถูกร้อยละ $1.11 \times 10^{-16} \times 100 \approx 0$

A100 หมายถึง ผู้ตอบมีคุณลักษณะที่ 4 อย่างเดียว มีโอกาสตอบถูกร้อยละ $1.00 \times 10^0 \times 100 = 100$

นั่นแสดงว่าข้อสอบข้อนี้ ผู้ตอบมีเพียงทักษะที่ 4 ทักษะเดียวก็เพียงพอในการ ตอบข้อสอบข้อนี้ถูก

A010 ผู้ตอบมีทักษะที่ 6 ทักษะเดียวมีโอกาสตอบถูกร้อยละ $2.45 \times 10^{-2} \times 100 \approx 2.45$

A111 ผู้ตอบมีครบทั้ง 3 ทักษะ มีโอกาสตอบถูกร้อยละ $1.00 \times 10^0 \times 100 = 100$

ซึ่งหากพิจารณาเฉพาะกรณีผู้ตอบไม่มีทักษะหรือคุณลักษณะใด แต่มีโอกาสดอบ ข้อสอบถูก นั่นคือค่าพารามิเตอร์การเดา (guessing) นั่นเอง หากเป็นใน DINA model จะมีคำสั่ง สำหรับขอ output การเดาและความสะเพร่า โดยเฉพาะ (วารางคณา ก้อมน้อย, 2563, หน้า 78)

การใช้กฎวิเคราะห์ข้อมูล ทั้ง CDM และ G-DINA MODEL โดยเรียก package CDM และ package GDINA จะให้ผลการวิเคราะห์ที่ละเอียด ในส่วนที่เป็นคำสั่งเพื่อการวิเคราะห์ด้วย โมเดล G-DINA ผลการวิเคราะห์ที่ได้ จะสรุปออกมาสั้น ๆ เข้าใจง่าย (เกศรินทร์ สมราช, 2563, หน้า 90)

3. แบบสอบวินิจัยทางพุทธิปัญญา

การทดสอบทางการศึกษาออกแบบเพื่อวินิจัยเชิงพุทธิปัญญาทั้งตรรกะรายละเอียด เนื้อหาตามวัตถุประสงค์การเรียนรู้ การให้เหตุผล การแก้ปัญหา Nichols (1994, อ้างถึงใน แคทรียา แสงดาวเทียน, 2562, หน้า 32) ได้ให้แนวทางในการพัฒนาแบบสอบไว้ 5 ขั้นตอน ดังนี้

1. การสร้างทฤษฎีเชิงเนื้อหา (Substantive theory construction) ขั้นตอนนี้แรกนี้ ต้องการพัฒนาโมเดลหรือทฤษฎีตามลักษณะโครงสร้างความรู้ตามสมมติฐาน และกระบวนการ ต้องการตอบสนองต่อการประเมิน และกำหนดข้อสอบที่เกี่ยวข้องกับทักษะและโครงสร้างความรู้

2. การออกแบบ (Design selection) ขั้นตอนนี้เป็นการวางแนวทางโดยโมเดลหรือ ทฤษฎีที่พัฒนาขึ้นจากขั้นตอนที่ 1 ซึ่งต้องการให้ผู้พัฒนาแบบสอบเลือกคุณลักษณะและการออกแบบ

3. การบริหารการสอบ (Test administration) ขั้นตอนนี้เกี่ยวข้องกับรายละเอียดที่ สำคัญของสภาพและบริบทของผู้สอบ เช่น รูปแบบข้อสอบ ตัวแทนของการเสนอข้อสอบ และการวาง แบบสอบ

4. การให้คะแนน (Response scoring) ขั้นตอนนี้เกี่ยวข้องกับการให้คะแนนผู้สอบ ซึ่งเป็นสารสนเทศของคุณลักษณะที่วัดโดยแบบสอบ

5. การปรับปรุงการออกแบบ (Design revision) ขั้นตอนนี้เกี่ยวข้องกับการตรวจสอบการออกแบบโมเดลหรือทฤษฎีที่ใช้ ผลของการประเมินถูกใช้เพื่อปรับปรุงต่อไป

Jang (2008) ได้กล่าวเกี่ยวกับการวินิจฉัยสำหรับผู้เรียนภาษาอังกฤษเป็นภาษาที่ 2 ว่า นักวิจัยทางการทดสอบทางภาษาหันมาสนใจในการสร้างแบบทดสอบที่สามารถให้ข้อมูลย้อนกลับ (Feedback) เพื่อนำมาเป็นแนวทางในการออกแบบการ เรียนการสอนและเป็นแนวทางในการเรียนรู้แก่นักเรียน อย่างไรก็ตามงานวิจัยที่เกี่ยวข้องกับการใช้ผล คะแนนเชิงวินิจฉัยในบริบทของการเรียนและการสอนยังมีน้อยเป็นเพราะว่าการให้ข้อมูลย้อนกลับในเชิงวินิจฉัยส่งผลแก่ผู้เรียนที่แตกต่างกันขึ้นอยู่กับระดับความสามารถ รูปแบบการเรียนรู้ และบริบท ของการเรียนรู้ของนักเรียน ดังนั้นการให้ข้อมูลย้อนกลับในเชิงวินิจฉัยควรถูกอธิบายอย่างละเอียด (Descriptive) และสามารถเข้าใจได้ง่าย เพื่อช่วยให้ผู้เรียนเกิดการพัฒนาและปรับปรุงตนเอง ทั้งยัง เป็นการปิดช่องว่างระหว่างความสามารถในปัจจุบันและเป้าหมายการเรียนรู้ที่ต้องการ ซึ่ง Jang ได้ นำเสนอกรอบแนวคิดเกี่ยวกับขั้นตอนการวินิจฉัยด้านพุทธิปัญญาไว้ดังนี้

ขั้นที่ 1 การระบุเป้าหมายของหลักสูตรและระบุวัตถุประสงค์ของการเรียนรู้ ครูผู้สอนต้อง ระบุทักษะทางภาษาที่คาดหวัง กิจกรรมการเรียนที่จะช่วยพัฒนาทักษะทางภาษา ระดับทางภาษาที่ คาดหวัง

ขั้นที่ 2 การเขียนข้อสอบและแบ่งข้อสอบ ในขั้นนี้ผู้เชี่ยวชาญด้านการประเมินทำงานร่วมกับ ครูผู้สอนที่พัฒนาข้อสอบและทักษะที่เจาะจง เมื่อข้อสอบที่ทดสอบทักษะเบื้องต้นได้ถูกพัฒนาขึ้น ข้อสอบพวกนี้จะถูกนำไปลองใช้เบื้องต้น ข้อมูลเกี่ยวกับความสามารถในการวินิจฉัยของข้อสอบแต่ละ ข้อจะถูกเก็บไว้ในระบบคอมพิวเตอร์ และข้อสอบจะถูกนำมาใช้

ขั้นที่ 3 หลังจากที่ครูผู้สอนใช้แบบทดสอบวินิจฉัยกับนักเรียนก่อนการสอนทำให้ทราบว่า ข้อบกพร่องของนักเรียนคืออะไร และเมื่อทำการสอนต่อไปอีกสักพัก ครูมีความอยากรู้ว่า ทักษะที่เป็น จุดอ่อนของนักเรียนได้รับการพัฒนาขึ้นหรือไม่ ครูจึงเข้าไปเลือกข้อสอบในคอมพิวเตอร์ที่จะทดสอบ ข้อบกพร่องของนักเรียน เมื่อนักเรียนทำข้อสอบเสร็จข้อมูลการทำข้อสอบของนักเรียนจะถูกส่งไปให้ คะแนน คอมพิวเตอร์จะรายงานผลคะแนนเชิงวินิจฉัย ซึ่งผลคะแนนที่ถูกอ่านโดยครูและนักเรียนที่ทำ ข้อสอบ ครูเข้าไปคุยกับนักเรียนแต่ละคนเพื่ออธิบายถึงหนทางในการพัฒนาจุดอ่อนของนักเรียน

ขั้นที่ 4 การนำผลวินิจฉัยไปใช้ ครูสามารถนำผลวินิจฉัยมาสะท้อนถึงรูปแบบการสอน และวางแผนจัดกิจกรรมที่ช่วยนักเรียนแต่ละคน

ลีและซาวากิ (Lee & Sawaki, 2009, อ้างถึงใน กุลพร พูนสวัสดิ์, 2556, หน้า 42-43) ได้ชี้แจงว่า นักวิจัยใช้วิธีการที่หลากหลายในการระบุระดับความรู้ด้านภาษาศาสตร์และระดับความชำนาญ (Mastery level) ของผู้สอบโดยดูจากคะแนนสอบในการทำข้อสอบทางภาษาอังกฤษ ซึ่งการวินิจฉัยเป็นวิธีการหนึ่งที่สำคัญที่สามารถระบุจุดแข็งจุดอ่อนของผู้เรียน และการวินิจฉัยทางภาษาควรให้ข้อมูลย้อนกลับเกี่ยวกับสถานะความรู้และทักษะในปัจจุบันของผู้เรียน เพื่อให้ผู้เรียนและครูสามารถแก้ไขปัญหาในด้านจุดอ่อนในการเรียนรู้ภาษาที่สองของผู้เรียน ลีและซาวากิ ได้อธิบายว่า วิธีวินิจฉัยทางพุทธิปัญญาประกอบด้วย 2 ส่วนหลัก ดังนี้

ส่วนที่ 1 การวิเคราะห์เนื้อหาของข้อสอบเพื่อระบุความสัมพันธ์ระหว่างข้อสอบและคุณลักษณะทางพุทธิปัญญาที่สนใจ ซึ่งในการวิเคราะห์เนื้อหานี้มักทำโดยผู้เชี่ยวชาญในการกำหนดกรอบของเนื้อหาและสร้างเมทริกซ์คิว (Q-matrix) ซึ่ง Q-matrix ถูกพัฒนาขึ้นโดย Tatsuoka ในปี 1983 และ 1990 เพื่อใช้ในการระบุความสัมพันธ์ระหว่างคุณลักษณะและข้อสอบ คุณลักษณะ (Attributes) คือ ชุดของทักษะ ความรู้ ความสามารถ กระบวนการทางสมอง และกลยุทธ์ที่ผู้สอบควรมีความชำนาญและใช้ในการตอบข้อสอบได้อย่างถูกต้อง

ส่วนที่ 2 การเลือกใช้โมเดลการวัดทางจิตวิทยา (Psychometric modeling) ที่เหมาะสมกับข้อสอบและคุณลักษณะที่สนใจ ซึ่งโมเดลการวัดทางจิตวิทยาเป็นโมเดลทางสถิติเพื่อใช้ในการค้นหาระดับความรู้หรือระดับความชำนาญของผู้เรียนในทักษะทางพุทธิปัญญาที่ระบุใน Q-matrix ซึ่งข้อมูลของระดับความรู้หรือระดับความชำนาญของผู้เรียนจะดูได้จากรูปแบบการตอบข้อสอบ (item response pattern) ของผู้เรียน Lee and Sawaki (2009, อ้างถึงใน กุลพร พูนสวัสดิ์, 2556, หน้า 42-43) ได้นำเสนอ 4 ขั้นตอนในการวินิจฉัยทางพุทธิปัญญา ดังนี้

- 1) ระบุคุณลักษณะ (Attribute) คือ ความรู้หรือทักษะที่ต้องใช้ในการตอบข้อสอบให้ถูกต้อง
- 2) สร้าง Q-Matrix เพื่อเป็นการระบุว่าคุณลักษณะอะไรที่ต้องการใช้ในการตอบข้อสอบให้ ถูกต้อง
- 3) วิเคราะห์ข้อมูลโดยใช้ Cognitive diagnostic psychometric models (CDPMs) เช่น Fusion model ในการประเมินคุณลักษณะของผู้สอบแต่ละคนว่ามี Skill profile แบบใด
- 4) รายงานผลการสอบและให้ข้อมูลวินิจฉัยย้อนกลับ โดยรายงานจาก Skill profile ที่ได้จากการวิเคราะห์แบบแผนการตอบ

อย่างไรก็ตามการศึกษาของ Lee & Sawaki (2009) เป็นการนำเอาข้อมูลการตอบข้อสอบที่มีอยู่แล้วมาวินิจฉัย และแบบทดสอบที่ใช้เก็บข้อมูลมิได้สร้างขึ้นโดยมีจุดมุ่งหมายเพื่อวินิจฉัย

จากประเด็นเกี่ยวกับขั้นตอนการวินิจฉัยเชิงพุทธิปัญญาตามแนวคิดของนักวิชาการทางภาษาศาสตร์ สรุปได้ว่า กระบวนการประเมินวินิจฉัยเชิงพุทธิปัญญาเป็นวิธีที่อยู่บนฐานของทฤษฎีการเรียนรู้ปัญญานิยมที่มุ่งอธิบายกระบวนการคิดในสมองของผู้เรียนซึ่งเป็นกระบวนการทางพุทธิปัญญา ดังนั้นการศึกษาครั้งนี้ผู้วิจัยจึงพัฒนาแบบสอบวินิจฉัยตามแนวคิดการออกแบบจากระบบการคิด (Cognitive design system: CDS) ของ เอ็มเบรทสัน (Embretson, Susan E., 1994, p. 110)

3.1 การตรวจสอบคุณภาพของแบบสอบ

การตรวจสอบคุณภาพของแบบสอบตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) คุณภาพของแบบทดสอบนั้นแบ่งออกเป็นคุณภาพของข้อสอบรายข้อ และ คุณภาพของเครื่องมือทั้งฉบับ ดังนั้น คุณภาพของข้อสอบรายข้อ การตรวจสอบคุณภาพข้อสอบรายข้อตามแนวคิดทฤษฎีการทดสอบแบบ ดั้งเดิมโดยทั่วไปจะพิจารณาเกี่ยวกับค่าความยากและค่าอำนาจจำแนกเป็นตัวบ่งชี้ที่สำคัญใน การพิจารณาตัดสินคุณภาพข้อสอบแต่ละข้อตามหลักการประเมินแบบอิงกลุ่ม รายละเอียดของ สูตรในการคำนวณและเกณฑ์การแปลความหมายคุณภาพข้อสอบแต่ละด้าน ดังรายละเอียด ต่อไปนี้ (ศิริชัย กาญจนวาสี, 2552; ฌัญฐภรณ์ หลาวทอง, 2548; เอมอร จังศิริพรภรณ์, 2549; รณิดา เขยชุม, 2551; เยาวดี ราชชัยกุล วิบูลย์ศรี, 2552, อ้างถึงใน เกศรินทร์ สมราช, 2563, หน้า 76)

1. คุณภาพของข้อสอบรายข้อ

1.1 ความตรง (Validity) จากการศึกษาของนักการศึกษาไทยเกี่ยวกับ Validity นิยามของ Bingham (1937; Guilford, 1946; Cureton, 1950, อ้างถึงใน เกศรินทร์ สมราช, 2563, หน้า 79) กล่าวว่า ความตรงคือ ความสัมพันธ์อย่าง ง่ายของคะแนนสอบที่สร้างขึ้นกับแบบทดสอบอื่น ๆ มีสหสัมพันธ์ของคะแนนที่สังเกตได้ของ แบบทดสอบกับเกณฑ์ของคะแนนจริง มีจุดมุ่งหมายเฉพาะโดยค่าความเที่ยงตรงที่สูงที่สุดของ แบบทดสอบควรจะมีสหสัมพันธ์กับคะแนนของแบบทดสอบที่วัดได้กับคะแนนจริงของ แบบทดสอบ ซึ่งมีวิธีที่สามารถตรวจสอบความสัมพันธ์นั้น โดยใช้ชื่อที่หลากหลายแตกต่างกันออกไป ความตรงที่เหมาะสมที่สุดสำหรับแบบสอบวินิจฉัย คือ ความตรงเชิงเนื้อหา ความสามารถในการวัดพฤติกรรมการเรียนรู้ได้ครอบคลุมเป็นตัวแทนของเนื้อหาทั้งหมดที่ต้องการจะวัด โดยให้ผู้เชี่ยวชาญตรวจสอบนิยามขอบเขตครอบคลุมเนื้อหาของสิ่งที่ต้องการวัด ด้วยการใช้ดัชนีตรวจสอบความสอดคล้องของเนื้อหา (IOC: Item-Objective Congruence) ให้ผู้เชี่ยวชาญเปรียบเทียบสัดส่วนข้อคำถามว่าเหมาะสมกับน้ำหนัก ความสำคัญของแต่ละเนื้อหาที่มุ่งวัดผ่านตารางวิเคราะห์หลักสูตรหรือผังการสร้างข้อสอบ (ฉัตรศิริ ปิยะพิมลสิทธิ์, 2541; ศิริชัย กาญจนวาสี, 2556, อ้างถึงใน เกศรินทร์ สมราช, 2563, หน้า 79)

1.2 ความยาก (P: Power of Difficulty or Difficulty Index) (เกศรินทร์ สมราช, 2563, หน้า 76-77) หมายถึง สัดส่วน ของจำนวนคนที่ตอบข้อสอบข้อนั้นได้ถูกต้องต่อ

จำนวนผู้เข้าสอบทั้งหมด สำหรับกรณีที่มีการให้ คะแนนแบบ 0 – 1 ดังนั้น หากข้อสอบที่มีผู้สอบตอบ ข้อสอบได้เป็นจำนวนมาก P จะมีค่าสูง คือ มีค่าเข้าใกล้ 1 แสดงว่าข้อสอบข้อนั้นง่าย ในทางตรงกันข้ามถ้าข้อสอบข้อใดมีผู้สอบได้จำนวนน้อย P จะมีค่าต่ำ คือ มีค่าเข้าใกล้ 0 แสดงว่าข้อสอบข้อนั้นยาก ความยากนั้นถือเป็นลักษณะที่สำคัญอย่างหนึ่งของเครื่องมือวัดทางด้านสติปัญญา ได้แก่ แบบสอบ วัดผลสัมฤทธิ์ แบบสอบวัดความถนัด โดยเครื่องมือเหล่านี้ต้องมีความยากพอเหมาะข้อสอบที่ง่ายเกินไปนั้นไม่มีประโยชน์เพราะทั้งคนเก่งคนไม่เก่งสามารถตอบถูก เหมือนกันหมดข้อสอบที่ยากมากเกินไปถือว่าไม่มีประโยชน์เช่นกัน เพราะมีผู้ตอบถูกเพียงไม่กี่คน แสดงว่าข้อสอบไม่สามารถรื้อให้ ผู้ตอบแสดงคุณลักษณะที่ต้องการวัดออกมาได้ ดังนี้

เกณฑ์การแปลความหมายผลการวิเคราะห์คำตอบถูก

0.80-1.00 ข้อสอบง่าย ควรตัดทิ้งหรือปรับปรุง

0.60-0.79 ข้อสอบค่อนข้างง่าย สามารถเก็บไว้ใช้ได้ต่อไป

0.40-0.59 ข้อสอบยากพอเหมาะ สามารถเก็บไว้ใช้ได้ต่อไป

0.20-0.39 ข้อสอบค่อนข้างยาก สามารถเก็บไว้ใช้ได้ต่อไป

0.00-0.19 ข้อสอบยาก ควรตัดทิ้งหรือปรับปรุง

1.3 อำนาจจำแนก (D: Discrimination Index หรือค่า r) หมายถึง ความสามารถของข้อสอบในการจำแนก/แยกความแตกต่างระหว่างผู้สอบที่มีผลสัมฤทธิ์ทางการเรียน ต่างกัน ให้ออกจากกันได้ (ศิริชัย, 2556, อ้างถึงใน เกศรินทร์ สมราช, 2563, หน้า 77-78) โดย คำนวนจาก ผลต่างระหว่างสัดส่วนจำนวนคนตอบถูกในกลุ่มเก่ง กับสัดส่วนจำนวนคนตอบถูกในกลุ่ม อ่อน อำนาจจำแนกของข้อสอบมีค่าตั้งแต่ -1 ถึง +1 ค่าอำนาจจำแนกที่ใช้ได้จะต้องมีค่าสูงกว่า 0.20 ขึ้นไป อำนาจจำแนกของเครื่องมือถือเป็นลักษณะที่สำคัญอีกอย่างหนึ่ง เป็นการ พิจารณาว่าเครื่องมือ ที่สร้างขึ้นนั้นสามารถวัดได้ตรงตามความเป็นจริงหรือไม่ ซึ่งแสดงถึง ความสัมพันธ์ของระดับ คุณลักษณะของสิ่งที่ต้องการวัดกับระดับคำตอบที่เลือกตอบ โดยคนที่มีความสามารถสูงควรเลือกตอบ ถูก คนที่มีความสามารถต่ำควรเลือกตอบผิด ดังนั้นเครื่องมือวัด ควรีค่าอำนาจจำแนกที่สามารถแยก หรือจำแนกกลุ่มผู้ตอบออกเป็นระดับต่าง ๆ ได้ถูกต้องว่า ผู้ตอบคนใดมีความสามารถมากคนใดมี ความสามารถน้อยออกจากกันได้

เกณฑ์การพิจารณาค่าอำนาจจำแนก (Ebel, 1965, p. 364, อ้างถึงใน เกศรินทร์ สม ราช, 2563, หน้า 77-78) ดังนี้

อำนาจจำแนกของข้อสอบ	ความหมาย
0.60-1.00	จำแนกได้ดีมาก
0.40-0.59	จำแนกได้ดี
0.20-0.39	จำแนกได้พอใช้แต่ต้องปรับปรุง

0.10-0.19	จำแนกได้ค่อนข้างต่ำ ต้องปรับปรุง หรือตัดทิ้ง
0.00-0.09	จำแนกไม่ได้ ควรตัดทิ้ง

ตาราง 14 ตัวอย่างการพิจารณาค่าอำนาจจำแนกของข้อสอบ

ข้อ	ค่าอำนาจจำแนก (r)	ความหมายค่า r	ประเมินคุณภาพจากค่าอำนาจจำแนกของข้อสอบ
1	.36	จำแนกได้	เป็นข้อสอบที่ดี
2	.12	จำแนกไม่ได้	ควรปรับปรุงให้คนเก่งตอบถูกมากกว่าคนอ่อน
3	-.45	จำแนกกลับ	ควรปรับปรุงให้คนเก่งตอบถูกมากกว่าคนอ่อน
4	.68	จำแนกได้	เป็นข้อสอบที่ดี
5	.92	จำแนกได้	เป็นข้อสอบที่ดี
6	-.04	จำแนกไม่ได้	ควรปรับปรุงให้คนเก่งตอบถูกมากกว่าคนอ่อน
7	.33	จำแนกได้	เป็นข้อสอบที่ดี
8	.46	จำแนกได้	เป็นข้อสอบที่ดี
9	.82	จำแนกได้	เป็นข้อสอบที่ดี
10	-.21	จำแนกกลับ	ควรปรับปรุงให้คนเก่งตอบถูกมากกว่าคนอ่อน

ดังนั้น ในแบบสอบหนึ่ง ๆ ควรประกอบด้วยข้อสอบที่มีอำนาจจำแนกตั้งแต่ .40 ถึง 1.00 หากข้อสอบที่มีค่าอำนาจจำแนกเป็นลบหรือเป็นศูนย์ ควรตัดทิ้งเพราะไม่สามารถแยกคนเก่งและคนอ่อนออกจากกันได้

2. คุณภาพของเครื่องมือทั้งฉบับ

การตรวจสอบคุณภาพของเครื่องมือเป็นคุณสมบัติเบื้องต้นของทุกแบบวัด เพื่อให้ได้ข้อมูลที่มีคุณภาพ ถูกต้อง สามารถนำไปใช้ประโยชน์ได้ สิ่งที่ต้องบอกคุณภาพของ เครื่องมือเป็นคุณสมบัติที่ต้องผ่านการตรวจสอบและรายงานคุณภาพเครื่องมือทั้งฉบับเสมอคือค่าความเชื่อมั่น (Reliability) หรือความเที่ยง (เกศรินทร์ สมราช, 2563, หน้า 80-82) หมายถึง ความคงที่หรือความคงเส้นคงวาของคะแนนในช่วงเวลาต่างกัน หรือความสอดคล้องของคะแนนที่วัดในเนื้อหาเดียวกัน เป็นการทำให้ได้ผลการทดสอบที่ใกล้เคียงของเดิมเมื่อผู้สอบทำการทดสอบซ้ำภายใต้สภาพการทดสอบที่เหมือนเดิม ความเที่ยงแบ่งเป็น 4 ประเภท ได้แก่ ความเที่ยงแบบความคงที่ ความเที่ยงแบบสมมูลหรือแบบ คู่ขนาน ความเที่ยงแบบความสอดคล้องภายใน และความเที่ยงแบบอิงเกณฑ์ ซึ่งความเที่ยงแต่ละประเภทมีความหมาย วิธีประมาณค่า ข้อดีและข้อจำกัดในการนำไปใช้แตกต่างกัน จากการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง พบว่า ความเชื่อมั่นแบบความสอดคล้องภายใน (Measure

of Internal Consistency) ความสอดคล้องกันระหว่างคะแนนรายข้อหรือความเป็นเอกพันธ์ของเนื้อหาข้อสอบ เป็นตัวแทนของ คุณลักษณะที่วัดอย่างเดียวกัน โดยใช้วิธีคำนวณต่าง ๆ มีวิธีประมาณค่า โดยคำนวณค่าสัมประสิทธิ์ความเป็นเอกพันธ์ระหว่างคะแนนของกลุ่ม ข้อสอบ 2 กลุ่ม จากการวัดด้วยแบบสอบเดียวกันค่าความเที่ยงที่ได้จากแต่ละสูตรจะมีความแตกต่างกันไม่เหมาะสมสำหรับใช้หาความเที่ยงแบบสอบจำกัดเวลา (Speed Test) เหมาะสมหรับแบบสอบที่มีความเป็นเอกพันธ์ (Homogeneity) เหมาะสมกับแบบสอบที่ให้คะแนนแบบ 0 – 1 ด้วยวิธี KR-20 ของคูเดอร์-ริชาร์ดสัน ซึ่งมีข้อดีคือเหมาะสำหรับแบบสอบที่ให้คะแนนแบบ 0 – 1 มีข้อดกลงเบื้องต้น คือ ข้อสอบทุกข้อมีความยากง่ายเท่ากัน ถ้าข้อสอบมีความยากง่ายเท่ากัน ค่าความเที่ยงของ KR-20 = KR-21 หากข้อสอบมีความยากง่ายต่างกัน $KR-20 > KR-21$ โดยค่าความเชื่อมั่นของแบบทดสอบใดๆ มีค่าอยู่ระหว่าง .00 ถึง 1.00 เกณฑ์การแปลผล ค่าความเชื่อมั่นของเครื่องมืออยู่ระหว่าง 0.00 – 1.00 ยิ่งใกล้ 1.00 ยิ่งมีความเชื่อมั่นสูง แบบทดสอบที่ดีควรมีค่าความเชื่อมั่นมากกว่า 0.70 (ปกรณ ประจันบาน, 2552) เกณฑ์การแปลผลความเชื่อมั่นมีดังนี้

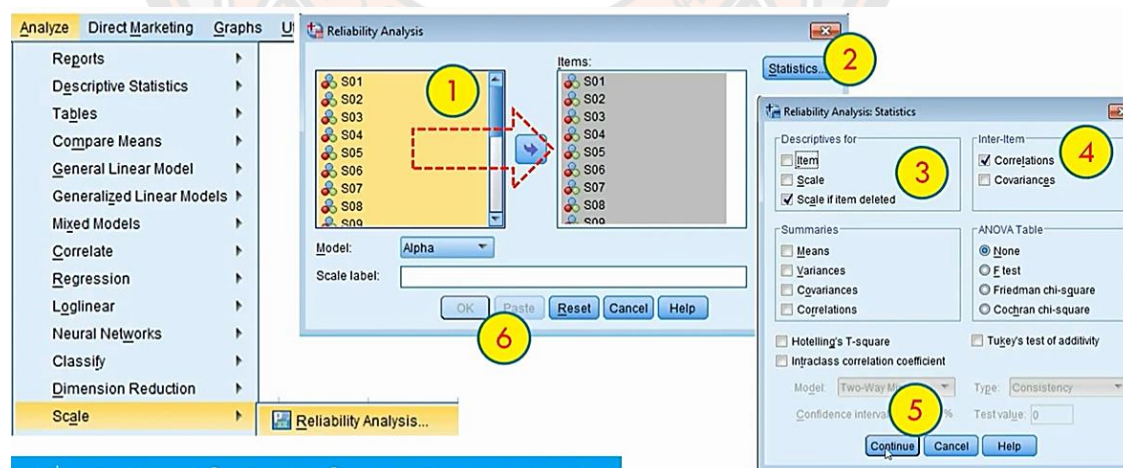
0.00 – 0.20 ความเชื่อมั่นต่ำมาก/ไม่มีเลย

0.21 – 0.40 ความเชื่อมั่นต่ำ

0.41- 0.70 ความเชื่อมั่นปานกลาง

0.71 – 1.00 ความเชื่อมั่นสูง

ค่าความเชื่อมั่นแบบความสอดคล้องภายในด้วยวิธี KR-20 ของคูเดอร์-ริชาร์ดสัน สามารถคำนวณได้ด้วยโปรแกรมคำนวณทางสถิติมีวิธีการ ดังภาพ 18



คำสั่ง: Analyze → Scale → Reliability Analysis

ภาพ 18 แสดงการใช้โปรแกรมคำนวณค่าความเชื่อมั่นด้วยวิธี KR-20

(ดร.ธณัฐ วงศ์สายเชื้อ. การวิเคราะห์ความเชื่อมั่น)

3.2 ความตรงของเมตริกซ์คิว (Q-matrix validation)

ความตรงของเมตริกซ์คิว เป็นการตรวจสอบความสอดคล้องระหว่างคุณลักษณะ/ ทักษะซึ่งในการศึกษาครั้งนี้เป็นคุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษที่ ถูกวัดโดยข้อคำถามข้อนั้น ๆ พิจารณาโดยผู้เชี่ยวชาญ และการตรวจสอบกับข้อมูลเชิงประจักษ์ โดย พิจารณาจากดัชนีจำแนก (discrimination index) หรือ ζ_j^2 ที่ดัชนีอำนาจจำแนก GDI เสนอโดย de la Torre และ Chiu (2016, pp. 40-41) สำหรับการตรวจสอบเชิงประจักษ์ Q-matrix ที่นำไป เชื่อมโยงกับโมเดล G-DINA สำหรับการให้คะแนนแบบสองค่า เช่น ตอบถูกได้ 1 ตอบผิดได้ 0 (dichotomous responses) สามารถขยายโมเดล G-DINA ที่ต่อเนื่องกัน และถูกกำหนดลักษณะที่ ชัดเจนสำหรับศูนย์ไม่แท้ของการที่ให้คะแนนแบบหลายค่า (polytomous responses) ได้อีกด้วย คำนวณได้จาก

$$\begin{aligned}\zeta^2 &= \zeta_{K':K''}^2 = \sum_{\alpha_{K'}=0}^1 \cdots \sum_{\alpha_{K''}=0}^1 w(\alpha_{K':K''}) [p(\alpha_{K':K''}) - \bar{p}(\alpha_{K':K''})]^2 \\ &= \sum_{\alpha_{K'}=0}^1 \cdots \sum_{\alpha_{K''}=0}^1 w(\alpha_{K':K''}) p^2(\alpha_{K':K''}) - \bar{p}^2(\alpha_{K':K''}),\end{aligned}$$

$$\text{เมื่อ } \bar{p}(\alpha_{K':K''}) = \sum_{\alpha_{K'}=0}^1 \cdots \sum_{\alpha_{K''}=0}^1 w(\alpha_{K':K''}) p(\alpha_{K':K''})$$

คือ ความน่าจะเป็นแบบถ่วงน้ำหนัก ดัชนีอำนาจจำแนก หรือ ζ_j^2 จะวัดความแปรปรวน ถ่วงน้ำหนักของความน่าจะเป็นในการตอบข้อสอบข้อที่ j อย่างถูกต้อง ให้คุณลักษณะ K_j แรกเข้าไป ในสมการดังตัวอย่างของ De la Torre และ Chiu (2016, pp. 40-41)

```
R> Qv <- Qval(GDINA.obj = GDINA_est, method = "wald")
R> Qv
```

Q-matrix validation based on Stepwise Wald test

Suggested Q-matrix:

	A1	A2	A3	A4
1	0	0	1	0
2	1	0	0	0
3	0	0	0	1
4	0	1	0	0
5	1	0	1	0
6	1	0	1	0
7	0	0	1	1
8	0	0	1	1
9	0	0*	1	0
10	0*	1	0	0
11	0*	1	0*	0
12	0	1	0*	0*

Note: * denotes a modified element.

ภาพ 19 เมตริกซ์คิวที่แนะนำ (Suggest Q-matrix)

จากภาพ 19 มีคุณลักษณะ 4 คุณลักษณะ ในข้อสอบข้อที่ 9 โปรแกรมให้ข้อเสนอแนะในการปรับปรุงเมตริกซ์คิวในคุณลักษณะที่ 2 ข้อสอบข้อที่ 10 โปรแกรมให้ข้อเสนอแนะในการปรับปรุงเมตริกซ์คิว ในคุณลักษณะที่ 1 ข้อสอบข้อที่ 11 โปรแกรมให้ข้อเสนอแนะในการปรับปรุงเมตริกซ์คิว ในคุณลักษณะที่ 1 และ 3 ข้อสอบข้อที่ 12 โปรแกรมให้ข้อเสนอแนะในการปรับปรุงเมตริกซ์คิว ในคุณลักษณะที่ 3 และ 4 เป็นต้น

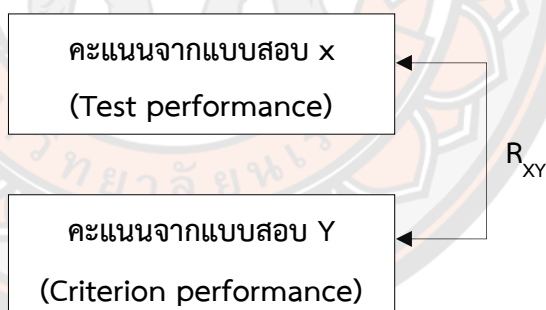
การทดสอบวาร์ด (Wald test) (Morrison, 1967) เป็นการทดสอบทางสถิติที่ได้รับความนิยมมานานหลายทศวรรษ ในเปรียบเทียบการทดสอบ Wald กับ Lagrange multiplier (LM), and likelihood ratio (LR) tests บุษ (Buse, 1982) เน้นย้ำว่าการทดสอบ Wald มีข้อได้เปรียบที่ว่าต้องการการประมาณค่าโมเดลที่มีขนาดใหญ่ต้องใช้ตัวอย่างจำนวนมากเพียงเท่านั้น (เช่น ไม่จำกัด) และไม่ต้องการอนุพันธ์ การทดสอบวาร์ดถูกนำมาใช้เพิ่มเติมโดย de la Torre และ Lee (2013) เพื่อตรวจสอบ CDM ที่เหมาะสมที่สุดที่ระดับรายการโดยเปรียบเทียบความพอดีของโมเดล (Model fit) ที่อิมพิวเตอร์กับความพอดีของโมเดลที่ลดลง พวกเขาตรวจสอบประสิทธิภาพของการทดสอบ Wald ด้วยการวิเคราะห์ข้อมูลทั้งแบบจำลองและข้อมูลจริง จากการศึกษาโมเดลพบว่า การทดสอบวาร์ด (Wald test) มีพลังที่ยอดเยี่ยมในการระบุโมเดลพื้นฐานที่แท้จริงแม้กับกลุ่มตัวอย่างขนาดเล็ก ขณะเดียวกันควบคุมข้อผิดพลาด Type-I สำหรับขนาดตัวอย่างขนาดใหญ่ด้วยจำนวนคุณลักษณะที่น้อย ซึ่งข้อผิดพลาด Type-I คือความน่าจะเป็นในการปฏิเสธสมมติฐานศูนย์เมื่อสมมติฐานศูนย์เป็นจริง

การทดสอบ Wald (Wald, 1943) เป็นการทดสอบสมมติฐานที่ใช้กันอย่างแพร่หลายในด้านสถิติ ในบริบทของ CDM ใช้สำหรับเปรียบเทียบโมเดล G-DINA กับโมเดล CDM ที่ถูกลดขนาดลง (Ma & de la Torre, 2019; Ma, Iaconangelo, & de la Torre, 2016; Sorrel, Abad, Olea, de la Torre, & Barrada, 2017; de la Torre, 2011) และการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (Hou, de la Torre, & Nandakumar, 2014; Ma, Terzi, Lee, & de la Torre, 2017) วิธีการใช้การทดสอบ Wald เพื่อประเมินแอตทริบิวต์ ในทางสถิติมีเวกเตอร์ q เกี่ยวข้อง 2 หรือมากกว่าหนึ่งขึ้นไป โดยเฉพาะอย่างยิ่ง หากการเปลี่ยนปัจจัยจาก 1 เป็น 0 ในเวกเตอร์ q นำไปสู่ความพอดีของโมเดลอย่างมีนัยสำคัญ แอตทริบิวต์ดังกล่าวจะถูกเรียกว่าไม่มีความจำเป็นหรือไม่จำเป็นต้องมีคุณลักษณะข้อนี้ สิ่งนี้ทำให้เราสามารถดำเนินการตรวจสอบความตรงของ Q-matrix จากมุมมองการเปรียบเทียบโมเดล

ในการศึกษาครั้งนี้ผู้วิจัยจึงตรวจสอบคุณภาพด้านความตรงของเมตริกซ์คิวด้วยดัชนีจำแนก (discrimination index) หรือ G_j^2 โดยใช้โปรแกรม R ด้วยการทดสอบวาร์ดใช้วิธี Stepwise Wald Test

3.3 ความตรงตามเกณฑ์สัมพันธ์ (Criterion – related Validity)

ความตรงตามเกณฑ์สัมพันธ์ เป็นคุณสมบัติของเครื่องมือด้านความสอดคล้องสัมพันธ์กันระหว่างคะแนนจากแบบสอบกับเกณฑ์ภายนอกที่สามารถใช้วัดคุณลักษณะที่ต้องการนั้นได้ เกณฑ์ภายนอกอาจเป็นสมรรถนะ ผลการปฏิบัติ พฤติกรรม หรือผลการดำเนินงานที่ปรากฏของคุณลักษณะนั้น เช่น ในสถานการณ์ของการสร้างแบบสอบคัดเลือกนักเรียนเข้าศึกษาในสถานศึกษา ก่อนที่จะนำคะแนนจากแบบสอบไปตัดสินใจเกี่ยวกับการคัดเลือกได้อย่างถูกต้องเหมาะสม ก็จะต้องรู้ถึงคุณภาพของคะแนนสอบที่ได้ว่าสามารถนำไปสรุปอ้างอิงถึงความสามารถทางการเรียนวิชาการในสถานศึกษาได้เป็นอย่างดีก่อน หลักฐานที่แสดงระดับความสัมพันธ์ระหว่างคะแนนสอบกับเกณฑ์ที่เป็นสมรรถนะ พฤติกรรม หรือความสามารถที่ปรากฏนี้เรียกว่า ความตรงตามเกณฑ์สัมพันธ์ จำแนกออกเป็นความตรงเชิงสภาพการณ์ หรือ ความตรงตามสภาพ หรือความตรงร่วมสมัย (Concurrent Validity) หมายถึง ความสามารถในการวัดคุณลักษณะที่สนใจได้ตรงตามสมรรถนะของสิ่งนั้นในสภาพปัจจุบัน เป็นคุณสมบัติของแบบสอบที่จะบ่งชี้ได้ว่าผู้เข้าสอบมีความสามารถตามเกณฑ์ที่กำหนดให้ในสภาพปัจจุบันเพียงใด การประมาณค่าความตรงตามสภาพของแบบสอบ ให้ความสนใจประมาณ สถานภาพปัจจุบันของสมรรถนะที่มุ่งวัด โดยคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง คะแนนจากแบบสอบ กับ คะแนนเกณฑ์ จากเครื่องมืออื่นที่สามารถใช้บ่งบอกสถานภาพปัจจุบันของลักษณะที่มุ่งวัดได้ดังภาพ 20



ภาพ 20 แสดงความสัมพันธ์ระหว่างคะแนนสอบกับเกณฑ์

(ศิริชัย กาญจนวาสี, 2556)

เนื่องจากเครื่องมือทั้งสองทำการวัดในช่วงเวลาเดียวกัน สัมประสิทธิ์สหสัมพันธ์ในทางบวกที่สูง จะแสดงว่า คะแนนจากแบบสอบสามารถใช้เป็นตัวบ่งชี้ที่ดีของสถานภาพของลักษณะที่มุ่งวัดนั้น

ตัวอย่าง สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบสอบข้อเขียนความรู้ความสามารถเกี่ยวกับงานช่าง (X) กับคะแนนเกณฑ์ที่ประเมินได้จากการสังเกตการปฏิบัติงานช่าง (Y) ในเวลาเดียวกันหรือเวลาใกล้เคียงกัน มีค่าเท่ากับ 0.90 แสดงว่า แบบสอบข้อเขียนนี้มีความตรงตามสภาพสูง สามารถวัดได้ตามสภาพที่เป็นอยู่ในปัจจุบัน คะแนนสอบจึงอาจนำไปใช้บ่งบอก

ความสามารถเกี่ยวกับงานช่างของผู้สอบแทนการติดตามสังเกตและประเมินความสามารถเป็นรายบุคคล ซึ่งต้องใช้เวลาและค่าใช้จ่ายสูง เป็นต้น (อภาพรรณ ประทุมไทย, 2017)

วิธีการตรวจสอบความตรงตามเกณฑ์สัมพัทธ์ โดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์เพื่อตรวจสอบความสอดคล้องระหว่างคะแนนที่ได้จากเครื่องมือที่พัฒนาขึ้น และคะแนนจากเครื่องมือที่เป็นเกณฑ์ภายนอกซึ่งมีคุณภาพสูง หรือมีมาตรฐานเป็นที่ยอมรับ สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson product-moment correlation coefficient) เพื่อหาความสัมพันธ์ระหว่างแบบวัดที่สร้างขึ้นกับแบบวัดมาตรฐานฉบับอื่นที่ใช้เป็นเกณฑ์ โดยข้อมูลที่น่ามาหาความสัมพันธ์ในสูตรนี้ควรมีลักษณะเป็นข้อมูลที่อยู่ในระดับการวัดมาตราอันดับขึ้นไป มาตราอันดับ (interval scale) และมาตราอัตราส่วน (ratio scale) และข้อมูลทั้ง 2 ชุดนั้นเป็นอิสระต่อกันคำนวณได้จากสูตร

$$r_{xy} = \frac{n \sum XY - \sum X \sum Y}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

เมื่อ r_{xy} แทน สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

X แทน คะแนนจากแบบวัดที่สร้างขึ้น

Y แทน คะแนนจากแบบวัดอื่นที่มีคุณภาพ, แบบวัดมาตรฐาน, คะแนนเกณฑ์

n แทน จำนวนผู้เข้าสอบ

การแปลความหมาย: ค่า r จะมีค่าอยู่ระหว่าง -1 ถึง 1

- ถ้าค่า r เป็นบวก และเข้าใกล้ 1 ➤ แบบวัดมีความตรงตามเกณฑ์สัมพัทธ์สูง
- ถ้าค่า r เป็นศูนย์ ➤ แบบวัดไม่มีความตรงตามเกณฑ์สัมพัทธ์สูง
- ถ้าค่า r เป็นลบ และเข้าใกล้ -1 ➤ แบบวัดมีความสัมพันธ์ในทิศทางตรงกันข้าม

การคำนวณโดยใช้โปรแกรม SPSS แสดงดังภาพ 21

คำสั่ง: Analyze ➤ Correlate ➤ Bivariate ➤ Pearson

		X	Y
X	Pearson Correlation	1	.984**
	Sig. (2-tailed)		.002
	N	5	5
Y	Pearson Correlation	.984**	1
	Sig. (2-tailed)	.002	
	N	5	5

** Correlation is significant at the 0.01 level (2-tailed).

ภาพ 21 แสดงตัวอย่าง out put การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์โดยใช้โปรแกรม SPSS

จากภาพ 21 ได้ผลลัพธ์ ค่า $r_{xy} = .984$ ค่าสัมประสิทธิ์สหสัมพันธ์เป็นบวก และเข้าใกล้ 1 หมายความว่า แบบวัดมีความตรงตามเกณฑ์สัมพัทธ์สูง (มีนัยสำคัญทางสถิติที่ระดับ .01)

3.4 การตรวจสอบคุณภาพแบบสอบตามแนวคิดทฤษฎีตอบสนองข้อสอบ (IRT)

หลักการพื้นฐานของทฤษฎีการตอบสนองข้อสอบเป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างความสามารถ ที่มีอยู่ภายในบุคคล (Latent trait or ability) กับผลการตอบข้อสอบหรือข้อคำถาม โดยใช้โค้งลักษณะ ข้อสอบ (Item Characteristic Curve; ICC) ซึ่งมีการกำหนดลักษณะของข้อสอบด้วยพารามิเตอร์ความยาก (b) อำนาจจำแนก (a) และโอกาสการตอบข้อสอบถูก (c) IRT จึงอยู่บนฐานความคิดที่สำคัญ 2 ประการคือ 1) ผลการตอบข้อสอบหรือข้อคำถามของผู้ตอบ สามารถอธิบายได้ด้วยความสามารถที่มีอยู่ภายในของผู้ตอบ และ 2) ความสัมพันธ์ระหว่างผลการตอบข้อสอบกับความสามารถที่มีอยู่ภายใน สามารถอธิบายได้ด้วย ฟังก์ชันลักษณะข้อสอบ หรือโค้งลักษณะข้อสอบ (ICC) อันมีลักษณะเป็นฟังก์ชันทางคณิตศาสตร์ เรียกว่าฟังก์ชัน โลจิส (Logistic function) หรือ ใกล้เคียงกับฟังก์ชันปกติสะสม (Normal ogive function) ทฤษฎีการตอบสนองข้อสอบ (IRT) พยายามอธิบายความสัมพันธ์ระหว่างคุณลักษณะภายใน หรือความสามารถที่มีอยู่ภายในตัวบุคคล กับพฤติกรรมการตอบสนองข้อสอบของบุคคลนั้นว่ามี โอกาสตอบข้อสอบถูกมากน้อยเพียงไร ทฤษฎีนี้มีพื้นฐานความเชื่อว่าพฤติกรรมการตอบสนองต่อ ข้อสอบของ ผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายในหรือความสามารถที่ อยู่ในตัวบุคคล ซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง ทฤษฎีนี้ได้อธิบายความสัมพันธ์ดังกล่าวในรูปของฟังก์ชันคณิตศาสตร์ หรือโมเดลที่แสดงความสัมพันธ์ระหว่างระดับความสามารถ คุณลักษณะของข้อสอบ และโอกาสของการตอบข้อสอบได้ถูก ที่เรียกว่า ฟังก์ชันการตอบสนองข้อสอบ ซึ่งมีลักษณะความสัมพันธ์เป็นแบบฟังก์ชันโลจิสหรือฟังก์ชันปกติสะสม (ศิริชัย กาญจนวาสี, 2550, หน้า 53)

3.5 โมเดลของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบมีการพัฒนารูปแบบต่าง ๆ ทำให้เป็นโมเดลเฉพาะขึ้น โดยที่แต่ละโมเดลจะแตกต่างกันขึ้นอยู่กับฟังก์ชันคณิตศาสตร์ และจำนวนพารามิเตอร์ของข้อสอบ ที่นำมาใช้อธิบายโค้งลักษณะข้อคำถาม ประกอบด้วย (ศิริชัย กาญจนวาสี, 2550, อ้างถึงใน แคนทรียา แสงดาวเทียน, 2560, หน้า 78) ทฤษฎีการตอบสนองข้อสอบมีแนวคิดที่สำคัญสองประการ คือ 1) พฤติกรรมการตอบข้อคำถามของผู้ทดสอบ (Examinee) สามารถทำนาย หรืออธิบายได้ด้วยชุดขององค์ประกอบ ที่เรียกว่า คุณลักษณะ (Traits) คุณลักษณะแฝง (Latent Traits) หรือความสามารถ (Abilities) และ 2) ความสัมพันธ์ระหว่างพฤติกรรมการตอบข้อคำถามของผู้ทดสอบ กับชุดของคุณลักษณะแฝงซึ่งมีอิทธิพลอยู่เบื้องหลังพฤติกรรมการตอบข้อคำถามนั้น สามารถอธิบายได้ด้วยฟังก์ชันคณิตศาสตร์ ที่เรียกว่า ฟังก์ชันลักษณะข้อสอบ (Item Characteristic Function) หรือ โค้ง

ลักษณะข้อสอบ (Item Characteristic Curve: ICC) ซึ่งเมื่อระดับของคุณลักษณะเพิ่มขึ้น ความน่าจะเป็นในการตอบข้อคำถามถูก (Correct Response) จะเพิ่มขึ้นด้วย (Hambleton et al., 1991, p. 7) โค้งลักษณะข้อสอบมีลักษณะเป็นฟังก์ชันโลจิส (Logistic Function) หรือฟังก์ชันปกติสะสม (Normal Ogive Function) บางครั้งอาจเรียกว่า โมเดลโลจิส หรือโมเดลปกติสะสม (ศิริชัย กาญจนวาสี, 2555, หน้า 53)

3.5.1 โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า

โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า ประกอบด้วย โมเดลการตอบสนองข้อสอบ แบบ 1, 2 และ 3 พารามิเตอร์ ศิริชัย กาญจนวาสี (2555, หน้า 77, อ้างถึงใน สารนิ

ติ บัญจรสพ, 2560, หน้า 65) กล่าวว่าโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ เหมาะสำหรับใช้กับข้อสอบเลือกตอบแบบทั่วไปข้อสอบแบบหลายตัวเลือก เนื่องจากผู้สอบสามารถเดาคำตอบได้

1. โมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์

โมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ (One-parameter model) โมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ หรือ Rasch model มีโค้งลักษณะข้อสอบที่เขียนด้วยฟังก์ชันโลจิส (แคทรียา แสงดาวเทียน, 2560, หน้า 78-79) ดังสมการ

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

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้ตอบซึ่งมีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้

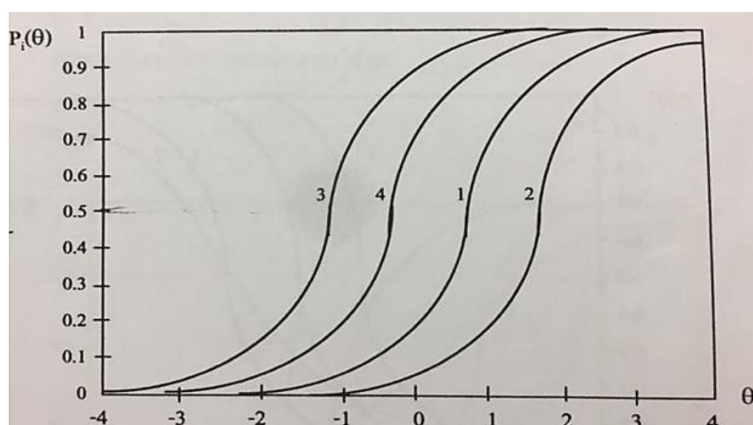
ถูกต้อง

b_i คือ ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่ง

ของ ICC

ณ จุด θ ที่มีโอกาสตอบข้อสอบถูก 0.50

$$e = 2.718$$



ภาพ 22 โค้งลักษณะข้อสอบแบบ 1 พารามิเตอร์ของตัวอย่างข้อสอบ 4 ข้อ

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ มี
ลักษณะที่
สำคัญดังนี้

b_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

a_i = มีค่าคงที่

$c_i = 0$

2. โมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์

ข้อสอบแบบ 2 พารามิเตอร์ มีโค้งลักษณะข้อสอบที่เขียนด้วย ฟังก์ชันโลจิส
(แคทรียา แสงดาวเทียน, 2560, หน้า79-80) ดังสมการ

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

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้ตอบซึ่งมีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้
ถูกต้อง

b_i คือ ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่ง
ของ ICC

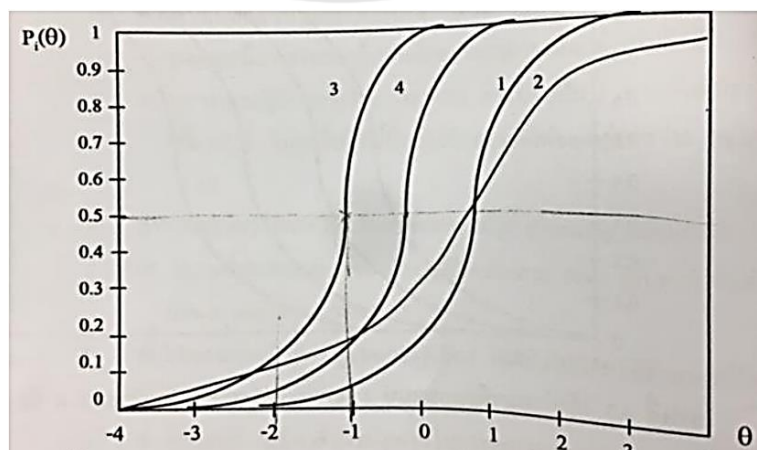
ณ จุด θ ที่มีโอกาสตอบข้อสอบถูก 0.50

a_i คือ ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นค่าที่ความชัน
ของ ICC

ณ ตำแหน่ง b_i

$e = 2.718$

$D = 1.70$



ภาพ 23 โค้งลักษณะข้อสอบแบบ 2 พารามิเตอร์ของตัวอย่างข้อสอบ 4 ข้อ

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ มีลักษณะสำคัญดังนี้

b_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

a_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

$c_i = 0$

3. โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (Three-parameter model: 3PL)

เป็นโมเดลที่พัฒนามาจาก Two-parameter logistic model เพื่อให้เหมาะกับแบบทดสอบที่มีอิทธิพลจากการเดาเข้ามารวมอยู่ด้วยและเป็นโค้งลักษณะข้อสอบที่แสดงถึงลักษณะข้อสอบที่มีค่าพารามิเตอร์ของข้อสอบ 3 ตัวข้อสอบแต่ละข้อมีความแตกต่างกัน ได้ทั้งพารามิเตอร์ a , b และ c โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ มีโค้งลักษณะข้อสอบที่เขียนด้วย ฟังก์ชันโลจิส (แคทรียา แสงดาวเทียน, 2560, หน้า 80-81) ดังสมการ

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

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้ตอบซึ่งมีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ ถูกต้อง

b_i คือ ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC

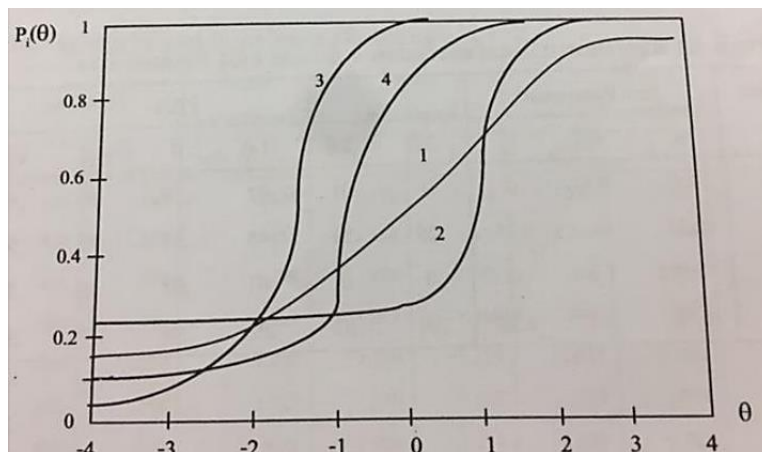
ณ จุด θ ที่มีโอกาสตอบข้อสอบถูก $\frac{1+c_i}{2}$
 a_i คือ ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นค่าที่ความชันของ ICC

ณ ตำแหน่ง b_i

c_i คือ ค่าพารามิเตอร์โอกาสเดาข้อสอบได้ถูก

$e = 2.718$

$D = 1.70$



ภาพ 24 โค้งลักษณะข้อสอบแบบ 3 พารามิเตอร์ของตัวอย่างข้อสอบ 4 ข้อ

จากภาพ 24 โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ มีลักษณะที่สำคัญดังนี้

b_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

a_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

c_i = มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ

ในการศึกษาครั้งนี้ผู้วิจัยใช้โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (3PL) ในขั้นตอนการหาคุณภาพของแบบสอบ (ฉบับร่าง) ครั้งที่ 2

3.5.2 ค่าพารามิเตอร์ของผู้สอบ

ค่าพารามิเตอร์ของผู้สอบ ได้แก่ 1) ความสามารถของผู้สอบ (θ) หมายถึง ความสามารถทักษะหรือศักยภาพของผู้ตอบข้อสอบแต่ละข้อว่ามีโอกาสทำข้อสอบว่ามีโอกาสทำข้อสอบข้อนั้น ถูกต้องมากน้อยเท่าใด หรือค่าประมาณได้จากโมเดลตามทฤษฎีการตอบสนองข้อสอบ ค่า θ มีพิสัยอยู่ระหว่าง $-\infty$ ถึง ∞ แต่ผลการวิเคราะห์ส่วนใหญ่ให้ค่าอยู่ในช่วง -3 ถึง $+3$ และ 2) ค่า $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้ตอบซึ่งมีความสามารถ θ จะตอบข้อสอบข้อ i ได้ถูกต้อง

3.5.3 ค่าพารามิเตอร์ของข้อสอบประกอบด้วย

1) พารามิเตอร์อำนาจจำแนก (α -parameter) หรือค่าบนเส้นโค้งลักษณะข้อสอบ (ICC) ณ จุดโค้งที่มีความชันมากที่สุดซึ่งอยู่ตรงกันข้ามกับค่าความยาก จุดนี้เป็นจุดแสดงความสามารถในการจำแนกผู้สอบที่มีความสามารถต่ำและสูงออกจากกันได้เด่นชัด ค่าอำนาจจำแนกยิ่งมีค่ามาก เส้นโค้งจะยิ่งมีความชันมาก ในทางทฤษฎี ค่าอำนาจจำแนก α มีค่าระหว่าง $-\infty$ ถึง ∞ ซึ่ง Baker (2001) ได้ให้แนวทางในการแปลความหมายค่าพารามิเตอร์อำนาจจำแนก (Baker, 2001, pp. 34-35, อ้างถึงใน สารนิติ บุญประสพ, 2560, หน้า 66) ดังนี้

- ข้อสอบที่มีค่า $\alpha + \text{infinity}$ หมายถึง อำนาจจำแนก สมบูรณ์
- ข้อสอบที่มีค่า α มากกว่า 1.70 หมายถึง มีอำนาจจำแนกสูงมาก
- ข้อสอบที่มีค่า α ระหว่าง 1.35 ถึง 1.69 หมายถึง มีอำนาจจำแนกสูง
- ข้อสอบที่มีค่า α ระหว่าง 0.65 ถึง 1.34 หมายถึง มีอำนาจจำแนกปานกลาง
- ข้อสอบที่มีค่า α ระหว่าง 0.35 ถึง 0.64 หมายถึง มีอำนาจจำแนก ต่ำ
- ข้อสอบที่มีค่า α ระหว่าง 0.01 ถึง 0.34 หมายถึง มีอำนาจจำแนก ต่ำมาก
- ข้อสอบที่มีค่า α เท่ากับ 0 หมายถึง ไม่มีอำนาจจำแนก

การวิจัยในครั้งนี้ ผู้วิจัยเลือกใช้เกณฑ์การพิจารณาค่าพารามิเตอร์อำนาจจำแนกตามแนวทางการแปลความหมายของ Baker (2001)

2) พารามิเตอร์ความยาก (b -parameter) คือ ระดับความสามารถของข้อสอบที่อยู่ตรงกันข้ามกับโค้งที่ชันที่สุดของ ICC และเป็นจุดที่ผู้ตอบข้อสอบมีความสามารถ $[P_i(\theta)]$ เท่ากับความยากของข้อสอบนั้นมีโอกาสที่จะตอบข้อสอบข้อนั้นได้อย่างถูกต้อง เท่ากับ $(1+c)/2$ สำหรับ 1PL และ 2PL ค่า $P_i(\theta)$ มีค่าเท่ากับ 0.5 ส่วนโมเดล 3PL ค่า $P_i(\theta)$ มีค่าเท่ากับ $(1+C_i)/2$ ในทางทฤษฎีค่า b มีค่าอยู่ระหว่าง $\pm \infty$ ข้อสอบที่ง่ายสอดคล้องกับผู้สอบที่มีความสามารถต่ำ ส่วนข้อสอบที่ยากสอดคล้องกับผู้สอบที่มีความสามารถสูง ในทางปฏิบัติ นิยมใช้ข้อสอบหรือข้อคำถามที่มีค่า b อยู่ระหว่าง -3 ถึง +3 (เขาวดี รวงชัยกุล วิบูลย์ศรี, 2552, หน้า 172; Baker, 2001, p. 22, อ้างถึงใน สารานิติ บุญประสพ, 2560, หน้า 66) ข้อสอบที่มีค่า b เท่ากับ -3 เป็นข้อสอบง่าย ค่า b เท่ากับ +3 เป็นข้อสอบยากค่า b เท่ากับ 0 ข้อสอบมีความยากปานกลาง (สุวิมล ตีรกานันท์, 2551, หน้า 159, อ้างถึงใน สารานิติ บุญประสพ, 2560, หน้า 66)

3) พารามิเตอร์โอกาสในการเดา (c-parameter) คือ ค่าโอกาสที่ผู้ตอบข้อสอบที่มีความสามารถต่ำจะสามารถเดาข้อสอบข้อนั้นได้อย่างถูกต้อง มากน้อยเพียงใด โดยมากค่า c จะมีค่าระหว่าง 0.00-0.40 ข้อสอบที่มีค่า $c > 0.30$ ขึ้นไป จะเป็นข้อสอบที่ไม่ดี ข้อสอบที่ดีควรมีค่า c น้อยกว่า 0.20 ค่า c ยิ่งน้อยข้อสอบยิ่งดี และถ้าค่า $c = .00$ จะดีมาก ปกติแล้ว ค่า c จะมีค่าประมาณ $(1/k) - 0.05$ เช่น ข้อสอบที่มีตัวเลือก 4 ตัว ค่า c ควรมีค่าประมาณ $(1/4) - 0.05$ หรือประมาณ 0.20

4) ค่าคงที่ e มีค่าเท่ากับ 2.718 และค่าคงที่ D มีค่าเท่ากับ 1.702 ซึ่งค่า D จะใช้ เป็นค่าองค์ประกอบของการปรับสเกลเพื่อทำให้ Logistic function กับ Normal Ogive function ใกล้เคียงกัน หรือมีค่า θ ต่างกัน ไม่เกิน .01

3.5.4 การประมาณค่าความสามารถและพารามิเตอร์ของข้อสอบ

ในการประมาณค่าความสามารถและพารามิเตอร์ของข้อสอบ (Ability and Item Parameter Estimation) ขั้นตอนสำคัญมากที่สุดของการนำ IRT ไปประยุกต์ใช้คือ การประมาณค่าพารามิเตอร์ตามโมเดล การตอบสนองข้อสอบที่เลือก ตามทฤษฎี $P_i(\theta)$ ขึ้นอยู่กับความสามารถของผู้สอบ (θ) และค่าพารามิเตอร์ (β) ที่แสดงลักษณะของข้อสอบ (a, b, c) ทั้ง θ และ β เป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรงและไม่ทราบค่า เราทราบแต่เพียงคำตอบของผู้สอบแต่ละคนที่ตอบข้อสอบแต่ละข้อ ปัญหาของการวิเคราะห์จึงอยู่ที่จะทำ อย่างไร จึงจะสามารถประมาณค่า θ ของผู้สอบแต่ละคนและพารามิเตอร์ของข้อสอบแต่ละข้อ จากข้อมูลการตอบข้อสอบของผู้สอบที่มีอยู่

ปัญหาการประมาณค่าพารามิเตอร์ IRT คล้ายการประมาณค่าพารามิเตอร์ในโมเดลการถดถอย (Regression model) ซึ่งนำค่าที่สังเกตได้ของตัวแปรเพื่อประมาณค่าพารามิเตอร์ของโมเดล แต่มีความแตกต่าง กันที่สำคัญ 2 ประการ คือ 1) โมเดลการถดถอยเป็นโมเดลเชิงเส้นตรง ในขณะที่โมเดลการตอบสนอง ข้อสอบเป็นโมเดลที่ไม่ใช่เส้นตรง มีลักษณะเป็นฟังก์ชันโลจิส และ 2) โมเดลการถดถอย มีตัวแปรทำนาย และตัวแปรเกณฑ์เป็นตัวแปรที่สามารถสังเกตได้ จึงประมาณค่าสัมประสิทธิ์การถดถอย เพื่อให้สมการ ทำนายสอดคล้องกับข้อมูลมากที่สุด โดยใช้เทคนิคความคลาดเคลื่อนกำลังสองน้อยที่สุด (Ordinary Least Squares : OLS) แต่โมเดลการตอบสนองข้อสอบทั้งตัวแปร θ และ β ต่างเป็นตัวแปรที่ไม่สามารถสังเกต โดยตรง จึงต้องอาศัยเทคนิคการประมาณค่าที่เป็นไปได้สูงสุด (Maximum Likelihood Estimate : MLE) การแจกแจงการสุ่มตัวอย่างของการประมาณค่าดังกล่าว จะประมาณค่าพารามิเตอร์ได้ถูกต้อง จำเป็นต้องใช้กลุ่มตัวอย่างขนาดใหญ่

3.5.5 ข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบ

การนำโมเดลการตอบสนองข้อสอบไปใช้ทางการศึกษาหรือทางจิตวิทยาจะต้องมีการตรวจสอบข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบที่สำคัญ 2 ประการ ได้แก่ (ศิริชัย กาญจนาสี, 2555, หน้า 75-76; Cohen, Swerdlik & Sturman, 2013, pp. 170-171; Embretson & Reise, 2000, pp. 45-48; Hambleton, Swaminathan, & Rogers, 1991, pp. 9-12, อ้างถึงใน สุชาติ สกลกิจรุ่งโรจน์, 2558, หน้า 27)

1. ความเป็นเอกมิติ (Unidimensionality) หมายถึง การที่ชุดของข้อคำถาม (หรือข้อสอบ) ในมาตรวัด (หรือแบบทดสอบ) มุ่งวัดคุณลักษณะแฝงเพียงคุณลักษณะเดียว (หรือความสามารถของบุคคลเพียงความสามารถเดียว) ในความเป็นจริงข้อตกลงเบื้องต้นดังกล่าวไม่สามารถพบได้อย่างสมบูรณ์ เนื่องจากปัจจัยทางด้านการรู้คิด (Cognitive) บุคลิกภาพ (Personality) และปัจจัยของการทดสอบ เช่น แรงจูงใจ ความกังวลในการทดสอบความรวดเร็วในการทำแบบทดสอบ การเดาคำตอบ และทักษะทางการรู้คิด อาจส่งผลร่วมกับพฤติกรรมที่แสดงออกในการทดสอบอยู่เสมอ ผลการทดสอบที่เกิดขึ้นจึงมิได้เกิดจากคุณลักษณะที่ต้องการวัดเพียงคุณลักษณะ

เดียวอย่างแท้จริงอย่างไรก็ดี ในทางปฏิบัติการตรวจสอบความเป็นเอกมิติ ต้องการเพียงหลักฐานที่แสดงให้เห็นว่า ชุดของข้อคำถามมุ่งวัดองค์ประกอบหลัก (Dominant Factor) เพียงองค์ประกอบเดียวแบบทดสอบใดที่มีคุณลักษณะดังกล่าว จะถือว่าเป็นแบบทดสอบที่มีคุณลักษณะเป็นโมเดลเอกมิติ (Unidimensional Model) ขณะที่แบบทดสอบใดมีการมุ่งวัดองค์ประกอบหลักมากกว่าหนึ่งองค์ประกอบ แบบทดสอบนั้นแสดงให้เห็นถึงโมเดลพหุมิติ (Multidimensional Model) ซึ่งเป็นโมเดลที่มีความซับซ้อนมากกว่า การตรวจสอบความเป็นเอกมิติของเครื่องมือวัด สามารถกระทำได้โดยการใช้เทคนิควิเคราะห์ทางสถิติ ได้แก่ การวิเคราะห์องค์ประกอบ (Factor Analysis) เพื่อคำนวณค่าไอเกน (Eigen Value) สำหรับศึกษาอัตราส่วนระหว่างค่าไอเกนขององค์ประกอบแรกกับองค์ประกอบถัดไป ถ้ามีอัตราส่วนสูง แสดงว่าเครื่องมือวัดคุณลักษณะหลักเพียงคุณลักษณะเดียว (Single Dominant Factor) นอกจากนี้ สามารถวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) เพื่อเพิ่มความมั่นใจ และช่วยยืนยันได้ว่าเครื่องมือวัดเพียงคุณลักษณะเดียว

2. ความเป็นอิสระ (Local Independence) หมายถึง เมื่อควบคุมคุณลักษณะหรือความสามารถของบุคคลให้คงที่แล้วผลการตอบข้อคำถามในข้อคำถามแต่ละคู่ จะต้องมีความเป็นอิสระจากกันทางสถิติ หรือกล่าวอีกนัยหนึ่งได้ว่า คุณลักษณะแฝงภายในบุคคล ต้องเป็นปัจจัยเดียวที่ทำให้การตอบข้อคำถามของผู้ทดสอบมีความแตกต่างกัน ความแตกต่างกันของการตอบข้อคำถามต้องไม่เกิดขึ้นจากความสัมพันธ์รูปแบบอื่น หากพิจารณาอย่างผิวเผิน แนวคิดเรื่องความเป็นอิสระของข้อคำถามอาจดูขัดแย้งกับแนวคิดเรื่องความสอดคล้องภายในของข้อคำถามตามทฤษฎีการทดสอบแบบดั้งเดิม ที่ระบุว่าข้อคำถามที่มีความสัมพันธ์กันมาก แสดงให้เห็นว่ามาตรวัดนั้นมีความเที่ยงในระดับสูง อย่างไรก็ตามเมื่อพิจารณาอย่างละเอียดจะพบว่า แนวคิดตามทฤษฎีทั้งสองนั้น มิได้ขัดแย้งกัน เพราะความเป็นอิสระของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบนั้นเป็นความอิสระของข้อคำถามที่เกิดขึ้นภายหลังการควบคุมพารามิเตอร์บุคคล และพารามิเตอร์ข้อสอบที่มีอิทธิพลโดยตรงต่อการตอบข้อคำถามในโมเดลแล้ว (สุชาติ สกลกิจรุ่งโรจน์, 2558, หน้า 27-28)

3.5.6 ฟังก์ชันสารสนเทศของแบบทดสอบ (Test information)

การวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบจะใช้แบบแผนการตอบสนองแบบสอบ เป็นรายข้อในการประมาณค่าความสามารถของผู้สอบ ดังนั้น การประเมินคุณภาพของแบบสอบจึงสามารถพิจารณาจากความถูกต้องแม่นยำ ในการประมาณความสามารถของผู้ตอบ โดยใช้ดัชนี ตัวหนึ่ง เรียกว่า สารสนเทศของแบบสอบ (Test information: $I(\theta)$) ซึ่งเป็นค่าฟังก์ชันสารสนเทศ ของแบบสอบอันเกิดจากผลรวมเชิงพีชคณิตของค่าฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อ รวมเข้าด้วยกันทั้งฉบับ ณ ตำแหน่ง θ เดียวกัน ดังสูตร

$$I(\theta) = \sum_{i=1}^k I_i(\theta) \quad , i = 1, 2, \dots, k$$

เมื่อ $I(\theta)$ แทน ค่าฟังก์ชันสารสนเทศหรือค่าสารสนเทศที่ได้จากแบบสอบสำหรับ ผู้ตอบที่มีความสามารถ θ

โปรแกรมสำหรับวิเคราะห์ทางสถิติ

โปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นมาเพื่อใช้ในการประมาณค่าพารามิเตอร์ของโมเดลการตอบสนองข้อสอบแบบพหุมิติมีอยู่หลากหลายโปรแกรม แต่ละโปรแกรมมีคุณสมบัติที่ต่างกันเช่น วิธีการประมาณค่าลักษณะโมเดลการตอบสนองข้อสอบแบบพหุมิติที่สามารถประมาณค่าได้จำนวนโมเดลการตอบสนองข้อสอบแบบพหุมิติที่สามารถประมาณค่าได้ ความยากง่ายในการใช้งาน คู่มือการใช้งานโปรแกรม และค่าใช้จ่ายที่เกิดขึ้นจากการใช้โปรแกรม ซึ่งผู้วิจัยจะนำเสนอโปรแกรมคอมพิวเตอร์ที่สามารถนำมาใช้งานบางโปรแกรมที่ผู้ใช้งานสามารถหามาใช้ได้ง่าย ดังนี้ (Reckase, 2009, pp. 148-175; Yao & Schwarz, 2005; Wu et al., 2007)

1. TESTFACT เป็นโปรแกรมคอมพิวเตอร์ที่ใช้ประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบสำหรับโมเดลการตอบสนองข้อสอบแบบพหุมิติด้วยวิธีการ Exploratory Factor Analysis(EFA) ที่ใช้ประมาณค่า Multidimensional Extension of the Two-parameter Normal Ogive Model และใช้ประมาณค่าพารามิเตอร์ Multidimensional Extension of the Three-parameter Normal Ogive Model โดยผู้ใช้ต้องคำนวณค่าการเดาจากโปรแกรมอื่นๆ แล้วนำมาป้อนข้อมูลให้กับโปรแกรม TESTFACT การประมาณค่าพารามิเตอร์ข้อสอบใช้วิธีการความน่าจะเป็นสูงสุด(Maximum Likelihood Method) ที่เรียกว่า Marginal Maximum Likelihood ส่วนการประมาณค่าความสามารถของผู้สอบใช้วิธีการแบบเบย์ (Bayesian Estimation Method) ผลของการประมาณค่าพารามิเตอร์จะได้ค่าอำนาจจำแนกที่มีหลายค่าต่อข้อสอบหนึ่งข้อ (Parameters- a) และ ได้ ค่าพารามิเตอร์ความยากหนึ่งค่าต่อข้อสอบหนึ่งข้อ (Parameters-d) รวมทั้งประมาณค่าองค์ประกอบของเวกเตอร์ความสามารถ (θ -vector) ให้ด้วย TESTFACT

2. NOHARM เป็นโปรแกรมคอมพิวเตอร์ที่ใช้ประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบสำหรับโมเดลการตอบสนองข้อสอบแบบพหุมิติได้ด้วยวิธีการทั้งสองแบบคือ Exploratory Factor Analysis (EFA) และ Confirmatory Factor Analysis (CFA) โดยชื่อโปรแกรมมาจากคำว่า Normal Ogive by Harmonic Analysis Robust Method การประมาณค่าพารามิเตอร์ข้อสอบใช้วิธีที่เรียกว่า Unweighted Least Squares เป็นโปรแกรมที่ใช้ประมาณค่าพารามิเตอร์ Multidimensional Extension of the Two-Parameter Normal Ogive Model และใช้ประมาณค่าพารามิเตอร์ Multidimensional Extension of the Three-Parameter Normal Ogive

Model โดยผู้ใช้ต้องคำนวณค่าการเดาจากโปรแกรมอื่นๆ แล้วนำมาป้อนข้อมูลให้กับโปรแกรม NOHARM การประมาณค่า พารามิเตอร์ข้อสอบของโปรแกรม NOHARM ใช้วิธีการที่แตกต่างจากโปรแกรม TESTFACT และผลของการประมาณค่าพารามิเตอร์จะได้ค่าอำนาจจำแนกที่มีหลายค่าต่อข้อสอบหนึ่งข้อ และได้ค่าพารามิเตอร์ความยากหนึ่งค่าต่อข้อสอบหนึ่งข้อ และเป็นโปรแกรมคอมพิวเตอร์ที่สามารถใช้งานได้ทั้งบนระบบดอส (Dos System) และบนระบบวินโดวส์ (Windows System) เป็นโปรแกรมที่ทำงานได้รวดเร็วมาก อีกทั้งไม่มีข้อจำกัดเกี่ยวกับจำนวนพารามิเตอร์ที่จะประมาณค่าแต่ขึ้นอยู่กับขีดความสามารถของเครื่องคอมพิวเตอร์ที่นำมาใช้งาน แต่โปรแกรม NOHARM จะไม่ประมาณค่าองค์ประกอบของเวกเตอร์ความสามารถ (θ -vector) ซึ่งหากผู้ใช้งานต้องการประมาณค่าส่วนประกอบของเวกเตอร์จะต้องใช้โปรแกรมอื่นๆ ช่วยในการประมาณค่า เช่น TESTFACT

3. ConQuest เป็นโปรแกรมคอมพิวเตอร์ที่ใช้ประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบสำหรับ MIRT ได้ด้วยวิธีการทั้งสองแบบคือ Exploratory Factor Analysis (EFA) และ Confirmatory Factor Analysis (CFA) เป็นโปรแกรมที่ใช้ประมาณค่าได้หลากหลายโมเดล เช่น Multifaceted Item Response Models, Multidimensional Item Response Models, Latent Regression Models โดยการประมาณค่าพารามิเตอร์โมเดลการตอบสนองข้อสอบแบบพหุมิติในตระกูลของ Rasch Models และการประมาณค่าพารามิเตอร์ข้อสอบใช้วิธีการความน่าจะเป็นสูงสุด (Maximum Likelihood Method) ที่แตกต่างกันสองแบบ คือ Quadrature Approach ใช้เมื่อจำนวน θ มีหนึ่งหรือสองค่า และ Monte Carlo Approach ใช้เมื่อจำนวน θ มีสามค่าหรือสามมิติขึ้นไป และสามารถประมาณค่า θ ได้ถึง 15 ค่าหรือ 15 มิติ

4. BMIRT (Bayesian Multivariate Item Response Theory) โปรแกรมดำเนินการโดยใช้วิธี Markov Chain Monte Carlo (MCMC) ด้วย Metropolis-Hasting Sampling Algorithm เพื่อประมาณค่าพารามิเตอร์ข้อสอบและผู้สอบ และการแจกแจงประชากรสำหรับชุดของโมเดลการตอบสนองข้อสอบแบบพหุมิติสำหรับข้อสอบทั้งที่ให้คะแนนแบบสองค่าและให้คะแนนแบบหลายค่า โปรแกรมนี้ใช้โมเดล M-2PPC และ M-3PL model สำหรับรูปแบบผสม

5. POLYFACT เป็นโปรแกรมสำหรับวิเคราะห์ข้อมูลที่มีการให้คะแนนแบบหลายค่า และใช้กับโมเดล Generalized Partial Credit Model ซึ่งใช้วิธี Marginal Maximum Likelihood ในการประมาณค่า และกำหนดโครงสร้างของพารามิเตอร์เชิงยืนยัน

6. IRTPRO (Item Response Theory for Patient-Reported Outcomes) เป็นลิขสิทธิ์ของ Scientific Software International (SSI) เป็นโปรแกรมใหม่ล่าสุดสำหรับการเปรียบเทียบคะแนนจากข้อสอบและคะแนนจากการทดสอบโดยใช้แนวคิดพื้นฐานจาก IRT ซึ่งอยู่บน

พื้นฐานของการวิเคราะห์ได้ทั้งแบบเอกมิติ (Unidimensional) และแบบพหุมิติ (Multidimensional) ทั้งแบบการวิเคราะห์องค์ประกอบเชิงสำรวจและการวิเคราะห์องค์ประกอบเชิงยืนยัน โดยใช้วิธีการประมาณค่าแบบ Maximum Likelihood (ML) และ Maximum a posteriori (MAP) ซึ่งสามารถใช้ได้กับฟังก์ชันการตอบสนองข้อสอบแบบต่าง ๆ เช่น Two-parameter Logistic, Three-parameter Logistic และ Generalized Partial Credit เป็นต้น

7. โปรแกรม R เป็นโปรแกรมคอมพิวเตอร์ ที่สามารถใช้ในการวิเคราะห์ข้อมูลเชิงสถิติเหมาะกับการใช้เพื่อการเรียนรู้ทางสถิติ และการใช้วิธีเชิงสถิติในการวิเคราะห์ข้อมูล ปัจจุบันเป็นที่รู้จักสำหรับนักวิจัยในวงกว้าง ถูกนำมาใช้ในการวิเคราะห์สถิติเพื่องานวิจัยอย่างแพร่หลายสาเหตุที่ R ได้รับความนิยมมากในปัจจุบันก็เนื่องจาก R เป็น Software ประเภท Open Source ที่ทุกคนสามารถนำมาใช้ได้โดยไม่มีค่าใช้จ่ายใด ๆ ผู้ใช้ไม่ต้องกังวลกับเรื่องการละเมิดลิขสิทธิ์เหมือนกับโปรแกรมสำเร็จรูปทางสถิติอื่น ๆ โปรแกรม R เป็นโปรแกรม ที่อยู่ภายใต้การดูแลของมูลนิธิที่ไม่แสวงหาผลกำไร ชื่อ R Foundation มี Robert Gentleman และ Ross Ihaka จากภาควิชาสถิติ มหาวิทยาลัย Auckland เป็นผู้ริเริ่มก่อตั้งและพัฒนา โปรแกรม R ขึ้นมาใช้ประโยชน์ทางวิชาการมีสมาชิกจำนวนหนึ่ง ช่วยดูแลจัดการเกี่ยวกับโปรแกรม R ให้กับผู้ใช้ตั้งแต่ปี 1997 ถึงปัจจุบัน (ดูรายละเอียดเพิ่มเติมได้จาก <http://www.r-project.org>) โปรแกรม R เป็นส่วนหนึ่งของโครงการของ GNU การใช้งานโปรแกรม R สามารถใช้ได้ทั้งบนระบบปฏิบัติการ Unix Macintosh และ Windows (เกศรินทร์ สมราช, 2563, หน้า 85) ทำการติดตั้งโปรแกรมลิงก์ R Package Version 4.1 จากเว็บไซต์ <http://CRAN.R-project.org/package=CDM> ในการศึกษาครั้งนี้ ผู้วิจัยคำนวณค่าอิทธิพลทดสอบด้วยโปรแกรม R ร่วมกับโปรแกรม WinBUGS โดยเรียกใช้งาน Package R2WinBUGS เพื่อความชัดเจนเกี่ยวกับอิทธิพลทดสอบของข้อสอบแบบทดสอบที่มีความไม่เป็นอิสระกัน เมื่อถูกนำไปสร้างแบบสอบวินิจฉัยทางพุทธิปัญญา

โปรแกรม R เป็นโปรแกรมคอมพิวเตอร์ ที่สามารถใช้ในการวิเคราะห์ข้อมูลเชิงสถิติเหมาะกับการใช้เพื่อการเรียนรู้ทางสถิติ และการใช้วิธีเชิงสถิติในการวิเคราะห์ข้อมูล ปัจจุบันเป็นที่รู้จักสำหรับนักวิจัยในวงกว้าง ถูกนำมาใช้ในการวิเคราะห์สถิติเพื่องานวิจัยอย่างแพร่หลายสาเหตุที่ R ได้รับความนิยมมากในปัจจุบันก็เนื่องจาก R เป็น Software ประเภท Open Source ที่ทุกคนสามารถนำมาใช้ได้โดยไม่มีค่าใช้จ่ายใด ๆ ผู้ใช้ไม่ต้องกังวลกับเรื่องการละเมิดลิขสิทธิ์เหมือนกับโปรแกรมสำเร็จรูปทางสถิติอื่น ๆ โปรแกรม R เป็นโปรแกรม ที่อยู่ภายใต้การดูแลของมูลนิธิที่ไม่แสวงหาผลกำไร ชื่อ R Foundation มี Robert Gentleman และ Ross Ihaka จากภาควิชาสถิติ มหาวิทยาลัย Auckland เป็นผู้ริเริ่มก่อตั้งและพัฒนา โปรแกรม R ขึ้นมาใช้ประโยชน์ทางวิชาการมีสมาชิกจำนวนหนึ่ง ช่วยดูแลจัดการเกี่ยวกับโปรแกรม R ให้กับผู้ใช้ตั้งแต่ปี 1997 ถึงปัจจุบัน (ดูรายละเอียดเพิ่มเติมได้จาก <http://www.r-project.org>) โปรแกรม R เป็นส่วนหนึ่งของโครงการของ

GNU การใช้งานโปรแกรม R สามารถใช้ได้ทั้งบนระบบปฏิบัติการ Unix Macintosh และ Windows (เกศรินทร์ สมราช, 2563, หน้า 85)

การใช้ Package G-DINA การวิเคราะห์ในโปรแกรม R ทำการติดตั้งโปรแกรมลิงก์ R Package Version 4.1. จากเว็บไซต์ <http://CRAN.R-project.org/package=CDM> หลังจากการติดตั้งโปรแกรม R แล้ว จะทำการโหลด Package สำหรับ วิเคราะห์ข้อมูล โดยทำการติดตั้ง 2 Package ได้แก่ Package CDM และ Package G-DINA เพื่อให้ได้ผลการวิเคราะห์ที่ละเอียด เนื่องจากความสามารถและข้อจำกัดของ Package CDM ที่อธิบายค่าสถิติพื้นฐานได้ละเอียดกว่า เมื่อทำการวิเคราะห์ข้อมูลด้วย Package G-DINA จึงจำเป็นที่จะต้องทำการวิเคราะห์ 2 Package เพื่อให้ได้ค่าสถิติที่ถูกต้องแม่นยำต่อวัตถุประสงค์หรือจุดมุ่งหมายการวิจัย ได้ครบถ้วน มีการอธิบายการติดตั้งและวิธีการวิเคราะห์ โมเดล G-DINA (Ravand and Robitzsch, 2015], p.12)

8. โปรแกรมสำเร็จรูปทางการวัดผลและวิจัยการศึกษา : IRT-BAY โดยส่วนมากแล้วจะเป็นโปรแกรมที่มาจากต่างประเทศ สิ่งที่น่ายินดีอย่างยิ่งคือมีโปรแกรมทางการวัดผลและวิจัยการศึกษาที่เขียนโดยคนไทย 2 ท่าน ซึ่งก็คือศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี และรองศาสตราจารย์ ดร.คณิต ไช่มุก มีชื่อว่าโปรแกรม IRT-BAY (วัดผลดอทคอม, 2561) คุณลักษณะของโปรแกรม IRT-BAY เป็นโปรแกรมที่ใช้กับระบบปฏิบัติการ DOS ด้วยเนื้อที่เพียง 270 MB ใช้คำนวณหาคุณภาพของแบบทดสอบในกรณีที่ตรวจให้คะแนนแบบผิดให้ 0 คะแนน ถูกให้ 1 คะแนน วิเคราะห์ข้อสอบรายข้อตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) สามารถประมาณค่าความสามารถของผู้สอบ (θ) ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) ใช้วิธีประมาณค่าพารามิเตอร์แบบเบย์เซียน (Bayesian) สามารถวิเคราะห์ข้อสอบรายข้อตามทฤษฎีมาตรฐานเดิม (Classical Test Theory) โดยจะ แสดงค่าความยาก (p) ค่าอำนาจจำแนก (Rbis, Rpbis) และความเชื่อมั่น (KR-20) แสดงโค้งสารสนเทศของข้อสอบ (Item Information Function) คุณลักษณะของข้อมูลนำเข้า ข้อมูลที่นำเข้าวิเคราะห์จะประกอบด้วย 2 ส่วนคือ ส่วนแรกจะอยู่ในบรรทัดแรก จะเป็นตัวเลขของข้อสอบ ส่วนที่สองตั้งแต่บรรทัดที่สองเป็นต้นไป จะเป็นคำตอบของผู้สอบ

ตอนที่ 3 มโนทัศน์เกี่ยวกับข้อสอบแบบทดสอบ

1. ลักษณะสำคัญของข้อสอบแบบทดสอบ

ในปี 1987 มีนักวิจัยสองท่านคือ Wainer และ Kiely ได้เสนอ package หรือ bundle of items ซึ่งเป็นชุดของข้อสอบที่ถูกมัดรวมกัน หรือกลุ่มของข้อสอบ (set) แต่ชื่อที่รู้จักทั่วไปถูกเรียกว่า Testlet test ข้อสอบแบบทดสอบ เป็นลักษณะของข้อสอบเชิงสถานการณ์ที่พบได้บ่อย

ในทางภาษาและทางสังคมศาสตร์ ในการประมาณค่าความสามารถของบุคคลในความเข้าใจสิ่งสิ่งเร้าเดียวกัน เช่น บทความ กราฟ แผนภูมิ หรือตารางตัวเลข เป็นต้น โดยทั่วไปการทำข้อสอบที่ใช้สถานการณ์การอ่านภาษาอังกฤษเหล่านี้ทำให้ผู้สอบใช้เวลาค่อนข้าง

มากถ้าจะตอบคำถามเพียงข้อเดียว ดังนั้นในแต่ละบทความจึงมีข้อคำถามที่เกี่ยวข้องกับบทความเป็นเซตเดียวกัน อาจมีจำนวนข้อคำถาม 3-5 ข้อ หรือบางครั้ง 6-12 ข้อ โดยเรียงลำดับข้อคำถามจะเรียงจากง่ายไปหายาก ซึ่งบทความจะมีความยาวมากกว่า 300 คำ (Wainer, Bradlow, & Wang, 2007: p. 60, อ้างถึงใน พิสิฐ พิณิจสกุล, 2563: หน้า 77; Wheadon, 2010: p. 106; อริสสา เทหลิม, 2559: หน้า 2; จงกล บัวแก้ว, 2560: หน้า 4)

ในการทำข้อสอบที่ใช้สถานการณ์ในการอ่านเหล่านี้ ใช้เวลาค่อนข้างมากถ้าจะตอบคำถามเพียงข้อเดียวและอาจเกิดสภาวะภาระทางปัญญา (Cognitive Load) ซึ่งจะเกิดขึ้นมากขึ้นขึ้นอยู่กับระดับความยากง่ายของข้อมูลรวมไปถึงปริมาณข้อมูลและสื่อในการนำเสนอข้อมูล ซึ่งถ้าขณะเรียนรู้นั้นเกิดภาระทางปัญญามากเกินไป ก็จะทำให้ไม่เกิดการรับรู้ใดๆ (กอบเกียรติ สระอุบล และพัลลภ พิริยะสุรวงศ์, 2557) ด้วยลักษณะเฉพาะของเทสต์เลท (Testlets test) ที่สามารถวัดคุณลักษณะหลายด้านในข้อเดียวกัน จึงสามารถลดจำนวนข้อสอบที่ใช้ในการทดสอบได้ถึงร้อยละ 30-50 และใช้เวลาในการทำข้อสอบน้อยลงอีกด้วย (Senarat et al., 2013)

การที่กลุ่มของข้อสอบตอบสนองสิ่งเร้าหรือบทความเดียวกัน ทำให้ข้อสอบในรูปแบบของชุดเทสต์เลทนั้นอาจมีการพึ่งพาคำตอบกันในแต่ละข้อ ซึ่งถือว่าเป็นการขึ้นนำคำตอบซึ่งกันและกัน อาจมีผลต่อค่าสารสนเทศของข้อสอบรายข้อ จากประเด็นดังกล่าว พิสิฐ พิณิจสกุล (2563, หน้า 81) กล่าวไว้ว่า หากลดจำนวนข้อสอบในแต่ละเทสต์เลทลงมา ก็จะสามารถลดการขึ้นนำคำตอบซึ่งกันและกันของข้อสอบให้น้อยลงด้วย อีกทั้งข้อสอบหนึ่งข้อต้องใช้ความสามารถหลายอย่างในการตอบทำให้ข้อสอบมีความยากสูงและซับซ้อน ซึ่งจะลดการขึ้นนำคำตอบระหว่างข้อได้อีกระดับหนึ่ง การใช้ชุดข้อสอบที่มีความสัมพันธ์กันภายในแบบสอบทำให้แบบสอบที่สร้างขึ้น มีความยืดหยุ่นมากขึ้นในการประเมินทางพุทธิพิสัยข้อสอบแบบเทสต์เลทวัดคุณลักษณะหลายด้านในข้อเดียวกัน ซึ่งข้อคำถามที่วัดหลายองค์ประกอบอาจเกิดโครงสร้างเป็นพหุมิติ (Multidimensional Structure)

เวนเนอร์ (Wainer, 2000) อธิบายว่า ความสามารถทางภาษาศาสตร์นั้นถือว่าเป็นสิ่งที่ผสมผสาน มีความเกี่ยวข้องสัมพันธ์กันของทักษะ/คุณลักษณะของความสามารถหลายด้านอย่างซับซ้อน (Wainer et al., 2000; DeMars, 2006, cited in Chen, 2014, p.11-12, อ้างถึงใน จงกล บัวแก้ว, 2560: หน้า 17; Wainer et al., 2007, p. 110, อ้างถึงใน พิสิฐ พิณิจสกุล, 2563: หน้า 77) สมมติฐานเกี่ยวกับลักษณะความเป็นพหุมิติของเทสต์เลท สอดคล้องกับการศึกษาของ Chen & Chen (2015) ทำการวิเคราะห์โครงสร้างการอ่านของ PISA (PISA Reading Structure) เพื่อสำรวจความสัมพันธ์ของคุณลักษณะหรือทักษะย่อยของการอ่านเพื่อความเข้าใจภาษาอังกฤษ ด้วยชุดของ

ข้อสอบที่มีสิ่งเร้าร่วมกัน (testlet) ในประเภทบทความ (passage) จำนวน 6 บทความ 20 ข้อคำถาม โดยการวิเคราะห์ข้อมูลการทดสอบความสามารถการอ่านเพื่อความเข้าใจของนักเรียนชาวอังกฤษ ระดับชั้นมัธยมศึกษาตอนต้นจำนวน 1,029 คน โดยใช้โมเดล G-DINA ซึ่งจัดอยู่ในโมเดลเชิงวินิจฉัยทางพุทธิปัญญาแบบทั่วไป (G มาจากคำว่า General) อย่างไรก็ตามการศึกษาของ Chen & Chen (2015, pp. 1-20) มุ่งเน้นการประมาณค่าพารามิเตอร์ข้อสอบโดยเป็นการนำคะแนนที่มีอยู่แล้วมาวิเคราะห์ภายหลัง เหนือเสนอความสัมพันธ์ของคุณลักษณะย่อยความสามารถการอ่านเพื่อความเข้าใจ 5 คุณลักษณะที่เขาสังเคราะห์จากแนวคิดของนักวิชาการด้านภาษาศาสตร์ โดยเชื่อมโยงความเกี่ยวข้องกับ 3 คุณลักษณะการอ่านเพื่อความเข้าใจของ PISA

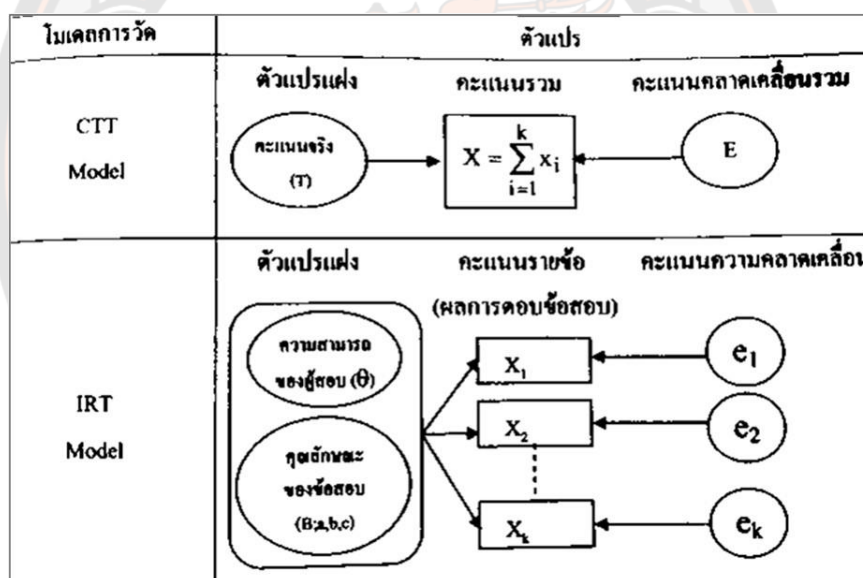
2. แนวคิดทฤษฎีที่เกี่ยวข้องกับข้อสอบแบบเทสต์เลท

ปัญหาที่พบในเมื่อใช้โมเดลการตอบสนองข้อสอบเอกมิติในการวัดทางจิตวิทยาเนื่องจากธรรมชาติของโครงสร้างคุณลักษณะแฝงทางจิตวิทยามีคุณลักษณะแฝงพหุมิติ เกือบทั้งหมด และคุณลักษณะแฝงพหุมิตินี้มีความซับซ้อนหลากหลายมิติ เช่น คุณลักษณะแฝงการปรับตัวข้อคำถามวัดการปรับตัว หลายรูปแบบที่แตกต่างกันก็จะวัดความสามารถในการปรับตัวทั้งหมดด้วย เมื่อประยุกต์ใช้โมเดลการตอบสนองข้อสอบเอกมิติ หลายโมเดลเพื่อให้มีความสอดคล้องเหมาะสมเพื่อให้ได้ข้อมูลที่ต้องการอาจจะมีปัญหา ดังนั้น เพื่อแก้ปัญหาดังกล่าว จึงมีการนำวิธีการวิเคราะห์ข้อมูลที่เรียกว่า Bi-factor model เพื่อให้ได้ข้อมูลทั้งคุณลักษณะแฝงทั่วไป (General Latent Trait) และคุณลักษณะแฝงเฉพาะ (Specific Latent Trait) ซึ่งปรากฏในการวิเคราะห์องค์ประกอบภายในชุดข้อสอบและการตอบข้อสอบ การวิเคราะห์แบบ Bi-factor จึงถูกนำมาใช้อย่างแพร่หลายในการทดสอบสมรรถภาพทางสมอง (Cognitive ability testing) เพื่ออธิบายโครงสร้างเครื่องมือวัดคุณลักษณะแฝง โดยเชื่อว่า คุณลักษณะแฝงทั่วไปหรือมิติทั่วไป อธิบายความแปรปรวนของข้อคำถามและองค์ประกอบของกลุ่ม ดังนั้นโมเดล Bi-factor จึงเป็นโมเดลที่มีความเป็นพหุมิติ และอาจจะเป็นโมเดลที่ประหยัดกว่า โมเดลการตอบสนองข้อสอบพหุมิติและเอกมิติ (Gibbons et al., 2009, อ้างถึงใน สารานิติ บุญประสพ, 2560, หน้า 88)

2.1 ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)

ทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างความสามารถ ที่มีอยู่ภายในบุคคล (Latent trait or ability) กับผลการตอบข้อสอบหรือข้อคำถาม โดยใช้โค้งลักษณะ ข้อสอบ (Item Characteristic Curve; ICC) ซึ่งมีการกำหนดลักษณะของข้อสอบด้วยพารามิเตอร์ความยาก (b) อำนาจจำแนก (a) และโอกาสการเดาข้อสอบถูก (c) IRT จึงอยู่บนฐานความคิดที่สำคัญ 2 ประการคือ 1) ผลการตอบข้อสอบหรือข้อคำถามของผู้ตอบ สามารถอธิบายได้ด้วยความสามารถที่มีอยู่ภายในของผู้ตอบ และ 2) ความสัมพันธ์ระหว่างผลการตอบข้อสอบ

กับความสามารถที่มีอยู่ภายใน สามารถอธิบายได้ด้วย ฟังก์ชันลักษณะข้อสอบ หรือโค้งลักษณะข้อสอบ (ICC) อันมีลักษณะเป็นฟังก์ชันทางคณิตศาสตร์ เรียกว่าฟังก์ชัน โลจิส (Logistic function) หรือ โกล์เคียงกับฟังก์ชันปกติสะสม (Normal ogive function) ทฤษฎีการตอบสนองข้อสอบ (IRT) พยายามอธิบายความสัมพันธ์ระหว่างคุณลักษณะภายใน หรือความสามารถที่มีอยู่ภายในตัวบุคคล กับพฤติกรรมตอบสนองข้อสอบของบุคคลนั้นว่ามี โอกาสตอบข้อสอบถูกมากน้อยเพียงไร ทฤษฎีนี้มีพื้นฐานความเชื่อว่าพฤติกรรมตอบสนองต่อ ข้อสอบของ ผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรง ว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายในหรือความสามารถที่ อยู่ภายในตัวบุคคล ซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง ทฤษฎีนี้ได้อธิบายความสัมพันธ์ดังกล่าวใน รูปของฟังก์ชันคณิตศาสตร์ หรือโมเดลที่แสดงความสัมพันธ์ระหว่างระดับความสามารถ คุณลักษณะ ของข้อสอบ และโอกาสของการตอบข้อสอบได้ถูก ที่เรียกว่า ฟังก์ชันการตอบสนองข้อสอบ ซึ่งมี ลักษณะความสัมพันธ์เป็นแบบ ฟังก์ชันโลจิสหรือฟังก์ชันปกติสะสม (ศิริชัย กาญจนวาสี, 2550, หน้า 53)



ภาพ 25 โมเดลการวัดแบบดั้งเดิมและโมเดลการตอบสนองข้อสอบ

(ศิริชัย กาญจนวาสี, 2550, หน้า 53)

โมเดลการวัด (Measurement Model) เป็นโมเดลเชิงคณิตศาสตร์ (Mathematical Model) ซึ่งแสดงความสัมพันธ์ระหว่างกลุ่มของตัวแปรต้น (Independent Variables) เพื่อทำนายตัวแปรตาม (Dependent Variables) โมเดลดังกล่าวอาจเป็นโมเดลอย่างง่ายที่ใช้ผลรวมของตัวแปรต้นเพื่อทำนายตัวแปรตามในลักษณะของโมเดลถดถอยเชิงเส้นตรง (Linear Regression Model) หรืออาจเป็นโมเดลความสัมพันธ์ที่มีความซับซ้อนมากขึ้น เช่น การรวมตัวกันของตัวแปรต้นโดยที่ตัวแปรอาจมีปฏิสัมพันธ์กัน หรือเกิดขึ้นภายใต้ฟังก์ชันการแจกแจงรูปแบบต่าง ๆ

เช่น การแจกแจงปกติสะสม (Cumulative Normal Distribution) การแจกแจงแบบโลจิสติก (Logistic Distribution)

ทฤษฎีการตอบสนองข้อสอบมีแนวคิดที่สำคัญสองประการ คือ

1) พฤติกรรมการตอบข้อคำถามของผู้ทดสอบ (Examinee) สามารถทำนายหรืออธิบายได้ด้วยชุดขององค์ประกอบ ที่เรียกว่า คุณลักษณะ (Traits) คุณลักษณะแฝง (Latent Traits) หรือความสามารถ (Abilities)

2) ความสัมพันธ์ระหว่างพฤติกรรมการตอบข้อคำถามของผู้ทดสอบ กับชุดของคุณลักษณะแฝงซึ่งมีอิทธิพลอยู่เบื้องหลังพฤติกรรมการตอบข้อคำถามนั้น สามารถอธิบายได้ด้วยฟังก์ชันคณิตศาสตร์ ที่เรียกว่า ฟังก์ชันลักษณะข้อสอบ (Item Characteristic Function) หรือ โค้งลักษณะข้อสอบ (Item Characteristic Curve: ICC) ซึ่งเมื่อระดับของคุณลักษณะเพิ่มขึ้น ความน่าจะเป็นในการตอบข้อคำถามถูก (Correct Response) จะเพิ่มขึ้นด้วย (Hambleton, Swaminathan, & Rogers, 1991, p. 7)

ทฤษฎีการตอบสนองข้อสอบ (IRT) อธิบายความสัมพันธ์ระหว่างคุณลักษณะภายใน หรือความสามารถที่มีอยู่ภายในตัวบุคคล กับพฤติกรรมการตอบสนองข้อสอบของบุคคลนั้นว่ามีโอกาสตอบข้อสอบถูกมากน้อยเพียงไร ทฤษฎีนี้มีพื้นฐานความเชื่อว่าพฤติกรรมการตอบสนองต่อข้อสอบของ ผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายในหรือความสามารถที่ อยู่ภายในตัวบุคคล ซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง ทฤษฎีนี้ได้อธิบายความสัมพันธ์ดังกล่าวใน รูปของฟังก์ชันคณิตศาสตร์ หรือโมเดลที่แสดงความสัมพันธ์ระหว่างระดับความสามารถ คุณลักษณะ ของข้อสอบ และโอกาสของการตอบข้อสอบได้ถูก ที่เรียกว่า ฟังก์ชันการตอบสนองข้อสอบ ซึ่งมีลักษณะความสัมพันธ์เป็นแบบฟังก์ชันโลจิสหรือฟังก์ชันปกติสะสม (ศิริชัย กาญจนวาสิ, 2550, หน้า 53) โดยมีข้อตกลงเบื้องต้นที่เกี่ยวข้องคือแบบทดสอบจะต้องมีมิติเดียว ซึ่งเป็นโมเดลการตอบสนองข้อสอบแบบเอกมิติ ใช้กระบวนการวัดและการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบเอกมิติ โดยมีความเชื่อว่าข้อสอบที่สร้างขึ้นนั้น วัดมิติเชิงเดียว (Single Dimension) แบบทดสอบใดมีการมุ่งวัดองค์ประกอบหลักมากกว่าหนึ่งองค์ประกอบ แบบทดสอบนั้นแสดงให้เห็นถึงโมเดลพหุมิติ (Multidimensional Model) ซึ่งเป็นโมเดลที่มีความซับซ้อนมากกว่า ทำให้ฝ่าฝืนข้อตกลงเบื้องต้นของทฤษฎีได้ง่ายและก่อให้เกิดประเด็นปัญหา คือ ข้อสอบวัดความสามารถที่ต้องการจริงหรือไม่ และค่าสถิติที่ได้จากการวัดไม่สามารถบ่งบอกความแตกต่างของมิติความสามารถที่ต้องการวัดได้ (Miller & Hirsch, 1992, อ้างถึงใน ชญารัตน์ บุญพุกฉกร, 2561, หน้า 37) ดังนั้นนักวัดผลจึงได้พยายามคิดทฤษฎีการทดสอบขึ้นใหม่เพื่อแก้ปัญหาที่เกิดขึ้น เรียกว่า “ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ” (Multidimensional Item Response Theory: MIRT)

2.2 ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ

ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory: MIRT) มีข้อตกลงเบื้องต้นที่สอดคล้องกับโมเดลตอบสนองข้อสอบแบบพหุมิติที่เกี่ยวข้องกับคุณลักษณะของผู้สอบ และคุณลักษณะของข้อสอบ มีรายละเอียด (Wilson & Hoskens, 2005; Wang et al., 2011; Wu et al., 2007; Baghaei, 2012, อ้างถึงใน ชญารัตน์ บุญพุฒิกิจ, 2561, หน้า 38) ดังนี้

ข้อตกลงเบื้องต้น

1. การเพิ่มขึ้นทางเดียว (Monotonicity Assumption) การเพิ่มขึ้นทางเดียวเป็นข้อตกลงเบื้องต้นที่กำหนดว่า เมื่อโครงสร้างตามสมมติฐานที่ต้องการวัดเพิ่มขึ้น ความน่าจะเป็นในการตอบข้อสอบถูกต้องก็น่าจะมีเพิ่มขึ้นด้วย กล่าวคือ ความน่าจะเป็นของการเลือกตอบข้อสอบที่ถูกต้องจะให้คะแนนเป็นถูกหรือไม่ถูกเพิ่มขึ้นหากสมาชิกของ θ -vector เพิ่มขึ้น

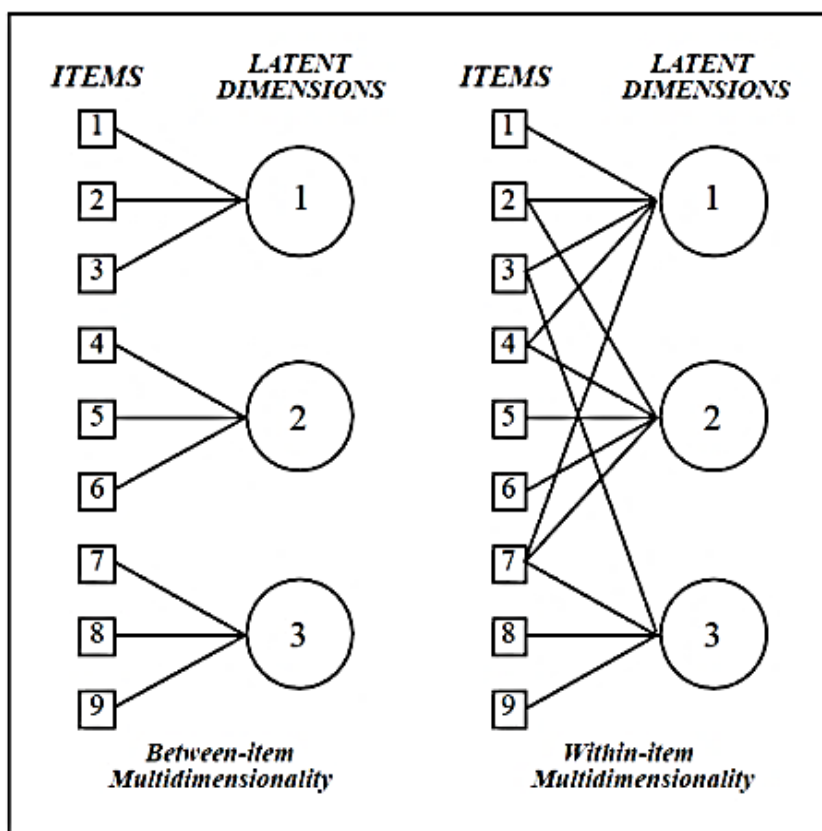
2. การนิยามฟังก์ชันความน่าจะเป็นในการตอบข้อสอบ ความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้สอบ โดยใช้โครงสร้างตามสมมติฐานสอดคล้องกับฟังก์ชันที่ได้นิยามขึ้น

3. ความเป็นอิสระในการตอบข้อสอบ (Local Independence Assumption) เป็นข้อตกลงเบื้องต้นที่กำหนดว่าความน่าจะเป็นในการตอบข้อสอบถูกหลายข้อได้ถูกต้อง พิจารณาจากผลคูณของความน่าจะเป็นในการตอบข้อสอบถูกหลายข้อ เมื่อความน่าจะเป็นคำนวณอย่างมีเงื่อนไขบนจุดในสเปซที่นิยามโดยสมมติฐาน คือการตอบข้อสอบของบุคคลหนึ่งไม่มีผลมาจากอีกบุคคลหนึ่ง

4. คุณลักษณะแฝงที่ต้องการวัดมากกว่า 1 องค์ประกอบ และแต่ละคุณลักษณะแฝงมีความสัมพันธ์กันคุณลักษณะแฝงที่ต้องการวัดมากกว่า 1 องค์ประกอบ และแต่ละคุณลักษณะแฝงมีความสัมพันธ์กันชนิดของแบบสอบพหุมิติที่มีความแตกต่างกัน 2 ชนิด ได้แก่

4.1 ความเป็นพหุมิติ ระหว่างข้อคำถาม (Between-items Multidimensionality) คือการมีลักษณะของข้อคำถามหนึ่งข้อจะวัดคุณลักษณะแฝงเดียว ซึ่งแต่ละคุณลักษณะแฝงมีความสัมพันธ์กัน

4.2 ความเป็นพหุมิติภายในข้อคำถาม (Within-items Multidimensionality) คือการมีลักษณะของข้อคำถามหนึ่งข้อจะวัดได้หลายคุณลักษณะแฝง (มากกว่าหนึ่ง) ซึ่งแต่ละคุณลักษณะแฝงมีความสัมพันธ์กัน แสดงลักษณะที่ต่างกันของโมเดลพหุมิติทั้ง 2 ชนิด ดังภาพ 25



ภาพ 26 พหุมิติระหว่างข้อคำถามและพหุมิติภายในข้อคำถาม
(Wu et al., 2007)

จากภาพ 26 แสดงลักษณะที่ต่างกันของโมเดลพหุมิติทั้ง 2 ชนิด โดยที่ความเป็นพหุมิติระหว่างข้อคำถาม (Between-items Multidimensionality) คือการมีลักษณะของข้อคำถามหนึ่งข้อจะวัดคุณลักษณะแฝงเดียว ซึ่งแต่ละคุณลักษณะแฝงมีความสัมพันธ์กัน ส่วนความเป็นพหุมิติภายในข้อคำถาม (Within-items Multidimensionality) คือการมีลักษณะของข้อคำถามหนึ่งข้อจะวัดได้หลายคุณลักษณะแฝง (มากกว่าหนึ่ง) ซึ่งแต่ละคุณลักษณะแฝงมีความสัมพันธ์กัน

2.2.1 โมเดลตอบสนองข้อสอบแบบพหุมิติ (MIRT model)

โมเดลตอบสนองข้อสอบแบบพหุมิติ (MIRT model) เป็นการขยายขีดความสามารถของในการประมาณค่าความสามารถ ของผู้สอบจากโมเดลการตอบสนองข้อสอบแบบมิติเดียว (Unidimensional IRT) (Rackase, 2009, p. 54) โดยโมเดลตอบสนองข้อสอบแบบพหุมิติ (MIRT) ยอมให้มีการวัดคุณลักษณะหลายคุณลักษณะพร้อมกันในข้อสอบข้อเดียวและคุณลักษณะ มีความสัมพันธ์กันได้ ทำให้ข้อสอบแต่ละข้อให้สารสนเทศสูงกว่าการทดสอบด้วยโมเดลตอบสนอง

ข้อสอบแบบมิติเดียว ทำให้ผลการวัดมีความแม่นยำสูง สามารถลดจำนวนข้อสอบที่ใช้ในการทดสอบได้ประมาณ 30-50 % (Senarat et al., 2013, p. 1010, อ้างถึงใน พิสิฐ พิณจสกุล, 2563, หน้า 79)

2.2.2 โมเดลตอบสนองข้อสอบพหุมิติแบบสององค์ประกอบ (Bi-factor MIRT model)

ผู้ริเริ่มพัฒนาโมเดลสององค์ประกอบ คือ Holzinger and Swineford (1937) โดยขยายแนวความคิดการทดสอบเชาว์ปัญญาของ Spearman (1904) ที่เรียกว่ากลุ่มองค์ประกอบ (Group factor) ที่มีองค์ประกอบร่วมกันแต่ไม่สัมพันธ์กันซึ่งประกอบด้วยองค์ประกอบทั่วไป (General factor : G factor) และองค์ประกอบเฉพาะ (Specific factor: S-factor) Holzinger and Swineford พัฒนาเครื่องมือวัดเชาว์ปัญญาที่เป็นคุณลักษณะรวมกันของ Spearman-Holzinger ลักษณะเฉพาะของโมเดลนี้คือองค์ประกอบทั่วไปถูกวัดภายใต้ข้อสอบข้อสอบทุกข้อในแบบทดสอบองค์ประกอบทั่วไป ทำมุมตั้งฉากกับองค์ประกอบเฉพาะ และข้อสอบทุกข้อวัดทั้งองค์ประกอบเฉพาะและองค์ประกอบทั่วไป (DeMars, 2013)

ต่อมาในปี 2017 Patarapichayatham and Kamata (2017) ทำการศึกษาความสัมพันธ์ระหว่างองค์ประกอบเฉพาะ (Testlet fator) และ องค์ประกอบทั่วไป (General factors) ยังไม่ได้รับความสนใจมากนัก อาจเป็นเพราะโมเดลการทดสอบแบบ bi-factor testlet model ถูกสันนิษฐานว่าองค์ประกอบเฉพาะ (testlet fator) เป็นปัจจัยก่อความและไม่สัมพันธ์กับองค์ประกอบทั่วไป อย่างไรก็ตาม ถ้าหาก องค์ประกอบเฉพาะ (testlet fator) มีความสัมพันธ์กับองค์ประกอบทั่วไป นั่นทำให้เกิดคำถามเกี่ยวกับความตรงของโมเดลที่ขึ้นอยู่กับทิศทางและขนาดของสหสัมพันธ์

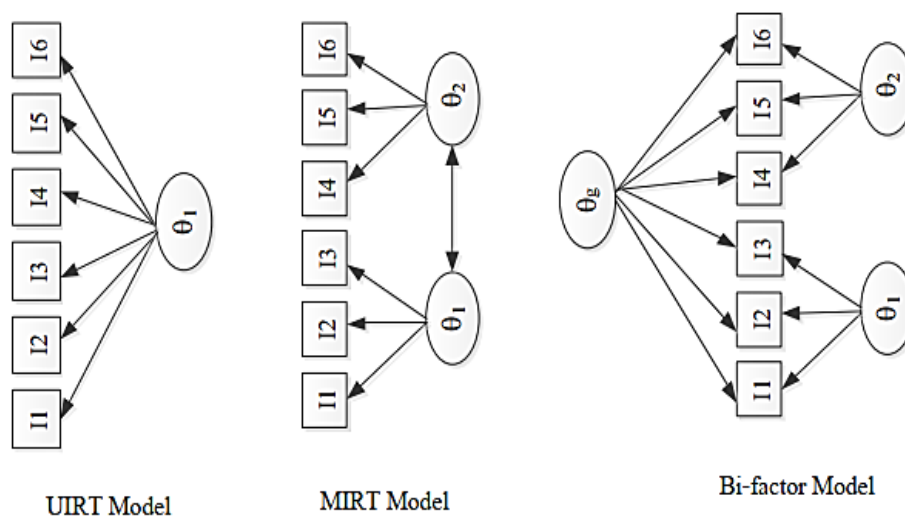
การศึกษาของ Patarapichayatham and Kamata (2017) เป็นเชิงปริมาณทำการตรวจสอบคุณภาพและลักษณะของการประมาณค่าความสามารถโดยใช้โมเดล bi-factor testlet IRT โดยมุ่งเน้นที่ความสัมพันธ์ระหว่าง องค์ประกอบเฉพาะ (testlet fator) กับองค์ประกอบทั่วไป (general factors) โดยเฉพาะอย่างยิ่ง ความสัมพันธ์ระหว่างการประมาณค่าความสามารถสำหรับองค์ประกอบเฉพาะ (testlet fator) กับองค์ประกอบทั่วไป (general factors) ถูกตรวจสอบ นอกจากนี้เรายังตรวจสอบผลรวมของคะแนนองค์ประกอบเฉพาะ (testlet fator) กับองค์ประกอบทั่วไป (General factors) ภายใต้โมเดล แบบดั้งเดิม และโมเดลแบบประยุกต์ อีกทั้งผู้วิจัยทำการประเมินความสัมพันธ์ระหว่างคะแนนรวมของการทดสอบที่สังเกตได้และคะแนนรวมที่ได้จากการจำลองข้อมูล ผลการศึกษาพบว่า ค่าความสัมพันธ์ระหว่างองค์ประกอบทั่วไป (general factors) และองค์ประกอบเฉพาะ (testlet fator) มีค่าเป็นบวกหมายความว่ามีความสัมพันธ์กันในเชิงบวกมีทิศทางเดียวกัน ในขณะที่ค่าความสัมพันธ์ระหว่างองค์ประกอบเฉพาะ (testlet fator) แต่ละองค์ประกอบเป็นลบทั้งหมด ค่าความสัมพันธ์ระหว่าง องค์ประกอบเฉพาะ (testlet fator) และการ

ประมาณค่าความสามารถมีค่าเป็นบวกหมายความว่ามีความสัมพันธ์กันในเชิงบวกมีทิศทางเดียวกัน นอกจากนี้ยังพบว่า ผลการวิเคราะห์ข้อมูลจริงมีความสอดคล้องกับผลการวิเคราะห์ด้วยข้อมูลจำลอง

โมเดลสององค์ประกอบ (Bi-factor model) เป็นโมเดลการตอบสนองข้อสอบพหุมิติที่สามารถวิเคราะห์ได้ทั้งองค์ประกอบเชิงสำรวจ (EFA) และองค์ประกอบเชิงยืนยัน (CFA) หรือเรียกว่าเป็น โมเดลแฝงภายใน (Nested factor model) ประกอบด้วยตัวแปรคุณลักษณะแฝงหลายคุณลักษณะและมีโครงสร้างเฉพาะที่ซับซ้อนจึงมีลักษณะเป็นพหุมิติ รูปแบบของโมเดลนี้ลักษณะของข้อคำถามทุกข้อวัด คุณลักษณะแฝงทั่วไป หรือองค์ประกอบทั่วไป และข้อคำถามชุดย่อยวัดคุณลักษณะแฝงเฉพาะหรือองค์ประกอบเฉพาะ

เมื่อพิจารณาเปรียบเทียบกรอบโมเดลการตอบสนองข้อสอบพหุมิติ พบว่า มีความแตกต่างจากโมเดลความเป็นพหุมิติ ระหว่างข้อคำถาม (Between-item multidimensionality) ซึ่งเป็นโมเดลอย่างง่ายที่ข้อคำถามแต่ละข้อวัดคุณลักษณะแฝงเพียงหนึ่งคุณลักษณะแฝง ในหลายๆคุณลักษณะแฝง ส่วนโมเดลพหุมิติภายในข้อคำถาม (Within-item multidimensionality) หมายถึงคุณลักษณะแฝงภายในข้อคำถามบางข้อ และคำถามบางข้อวัดมากกว่าหนึ่งคุณลักษณะแฝง ซึ่งมีโครงสร้างการวัดที่ซับซ้อน (Adams, Wilson, & Wang, 1997, อ้างถึงใน สารนิตติ บุญประสพ, 2560)

ดังนั้นสรุปได้ว่า โมเดล bi-factor เป็นโมเดลการตอบสนองข้อสอบพหุมิติที่ปรากฏลักษณะของความ เป็นพหุมิติเป็นแบบพหุมิติภายในข้อคำถาม (Within-item multidimensionality) เนื่องจากข้อคำถามที่วัด คุณลักษณะแฝงทั่วไป และคุณลักษณะแฝงเฉพาะมากกว่าหนึ่งกลุ่ม ซึ่งแตกต่างจากโมเดลการตอบสนองข้อสอบเอกมิติ (UIRT) ที่ข้อคำถามจะวัดคุณลักษณะแฝงเดียว แสดงลักษณะของโมเดลการตอบสนองข้อสอบเอกมิติ โมเดลการตอบสนองข้อสอบพหุมิติ และ โมเดลการตอบสนองข้อสอบพหุมิติแบบสององค์ประกอบ (สารนิตติ บุญประสพ, 2560) ดังภาพ 26

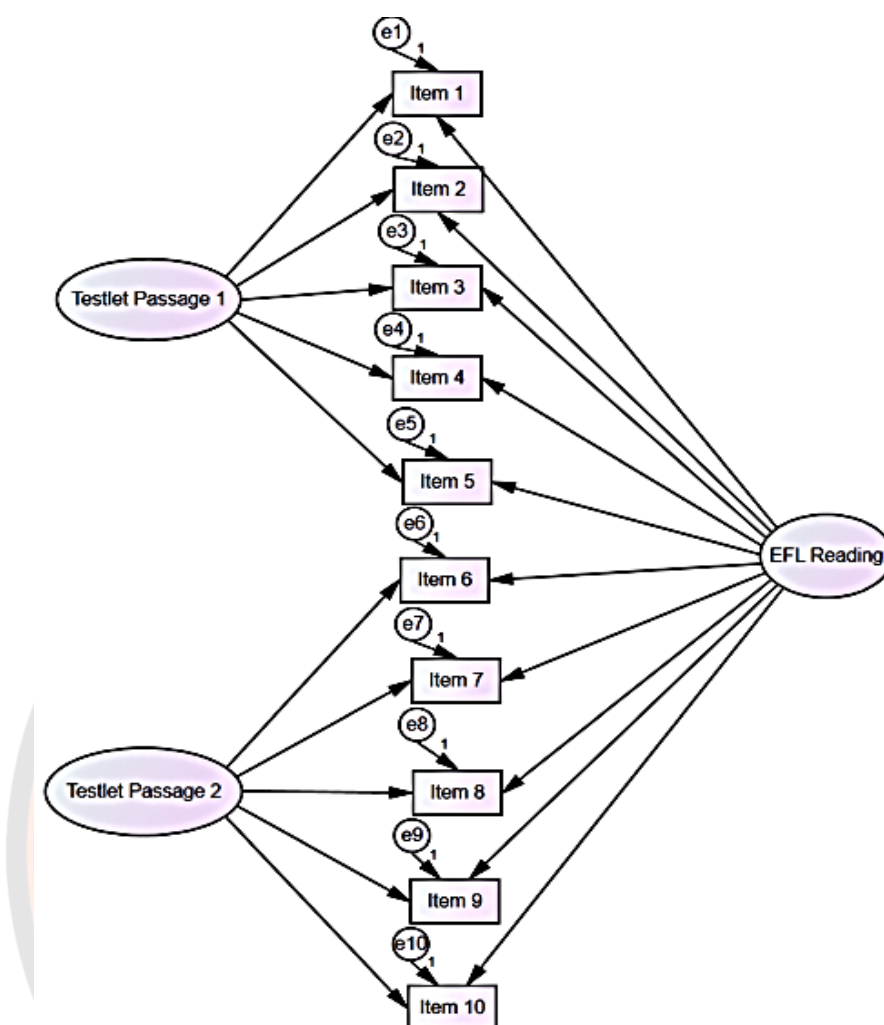


ภาพ 27 ลักษณะโครงสร้างของโมเดล UIRT, MIRT และ Bi-factor
 (สารนิตี บุญประสพ, 2560)

หมายเหตุ I (item) หมายถึง ข้อคำถาม, θ หมายถึง คุณลักษณะ /ความสามารถ

โครงสร้างของโมเดล 2 องค์ประกอบ (Bi-factor model) เป็นโครงสร้างที่ถูกนำไปประยุกต์ใช้อย่างหลากหลายของการทดสอบทางจิตวิทยาและการทดสอบทางการศึกษาที่อธิบายองค์ประกอบทั่วไปหนึ่งกลุ่มและองค์ประกอบเฉพาะหนึ่งกลุ่ม โมเดลสององค์ประกอบสามารถแทน Factorial structure ลักษณะโครงสร้างทางคณิตศาสตร์โมเดลการตอบสนองข้อสอบเปรียบเทียบ 3 โมเดล UIRT model, MIRT model และ Bi-factor model มีความแตกต่างกัน

ตัวอย่าง การศึกษาของ Geramipour M. (2021, p. 4) มีวัตถุประสงค์เพื่อใช้โมเดลการวัดกับการตอบสนองข้อสอบจากการทดสอบจริงของการทดสอบความเข้าใจในการอ่านของผู้เรียนที่ใช้ภาษาอังกฤษเป็นภาษาที่ 1 หรือที่เรียกว่า EFL ของประเทศอิหร่าน และเปรียบเทียบความถูกต้องของโมเดลสององค์ประกอบแบบพหุมิติและพารามิเตอร์ข้อสอบด้วยราชโมเดลแบบเอกมิติและราชโมเดลแบบพหุมิติ เพิ่มอิทธิพลทดสอบเลขเข้าไปในโมเดลการวัดเป็นโมเดล 3PL TRT แผนภาพเส้นทางการทดสอบสมมุติฐานโมเดลแสดงไว้ในภาพ 27



ภาพ 28 โมเดลทดสอบทดสอบมาตรฐานสำหรับวัดความเข้าใจในการอ่าน EFL Test
ที่มา: Wainer และคณะ (2000, cited in Geramipour M.,2021, p.4)

จากภาพ แบบสอบ EFL Reading ฉบับนี้มีบทอ่าน 2 เรื่อง ได้แก่ Testlet Passage 1 ที่ประกอบด้วยข้อสอบแบบทดสอบเลข (Testlet items) ข้อที่ 1- 5 จำนวน 5 ข้อ และ Testlet Passage 2 ที่ประกอบด้วยข้อสอบแบบทดสอบเลข (Testlet items) ข้อที่ 6-10 จำนวน 5 ข้อ รวมทั้งสิ้น 10 ข้อ

2.3 อิทธิพลของทดสอบเลข (Testlet effect)

อิทธิพลของทดสอบเลข (Testlet effect) หมายถึงผลกระทบของอิทธิพลแบบสุ่มที่เป็นปฏิสัมพันธ์ของการตอบกับทดสอบเลข คำนวณได้จากการใช้โมเดล Bi - factor MIRT ด้วยวิธีของเบส การวิเคราะห์ทางสถิติสามารถใช้โปรแกรม R ร่วมกับโปรแกรม WinBUGS โดยเรียกใช้งาน Package R2WinBUGS ในการคำนวณอิทธิพลของทดสอบเลข (อริสฟา เตห์ลิ้ม, 2559)

การแปลความหมายของขนาดอิทธิพล

มีนักวิชาการกำหนดเกณฑ์การแปลความหมายของขนาดอิทธิพล (Effect size) ไว้หลากหลาย ดังนี้ Cohen J. (1988, อ้างถึงใน อรุณ จิรวัดน์กุล, 2556, หน้า 935) ได้นำเสนอเป็นเกณฑ์กว้างๆ ในการแปลผลขนาดอิทธิพลดังนี้

0.2 = น้อย

0.5 = ปานกลาง

0.8 = มาก

Patarapichayatham & Kamata (2017, p.164) กำหนดเกณฑ์ขนาดอิทธิพล เทสต์เลท 3 ขนาดในระดับความสามารถของปัจจัยทั่วไปที่เป็นมาตรฐาน ได้แก่

0.0001 หมายถึง มีขนาดเล็กมาก

0.1 หมายถึง มีขนาดเล็ก

0.4 หมายถึง มีขนาดใหญ่

การกำหนดค่าอิทธิพลที่เป็นมาตรฐานซึ่งส่วนใหญ่จะนิยมค่า effect size เท่ากับ 0.3 ที่ระดับนัยสำคัญทางสถิติเท่ากับ 0.05

ค่า effect size เท่ากับ 0.1 หมายถึง มีขนาดอิทธิพลในระดับเล็ก (small)

ค่า effect size เท่ากับ 0.3 หมายถึง มีขนาดอิทธิพลในระดับปานกลาง (medium)

ค่า effect size เท่ากับ 0.5 หมายถึง มีขนาดอิทธิพลในระดับใหญ่ (large)

งานวิจัยเกี่ยวกับการพัฒนาแบบทดสอบการอ่านเพื่อความเข้าใจภาษาอังกฤษที่มีการใช้ข้อสอบแบบเทสต์เลทร่วมกับข้อสอบเดี่ยว การศึกษาเกี่ยวกับการวินิจฉัยเชิงพุทธิปัญญาการอ่านเพื่อความเข้าใจของจาง (Jang, 2005) ด้วยโมเดลวินิจฉัยเชิงพุทธิปัญญา (CDM) Fusion model การวิจัยของจางแบ่งออกเป็น 3 ระยะ จางพัฒนาแบบทดสอบการอ่าน 2 ชุด ประกอบด้วยข้อสอบเทสต์เลท 34 ข้อ และข้อสอบเดี่ยว เขาทำการประมาณค่าอิทธิพลของเทสต์เลท (Testlet Effect) เพื่อศึกษาขนาดของอิทธิพลเทสต์เลท (Jang, 2005, p. 83) ดังตาราง 15

ตาราง 15 แสดงขนาดของอิทธิพลเทสต์เลท Confirmatory DIMTEST Analyses

Passage Title	Form	T	p
The Expression of emotion	1	1.62	.05
Nineteenth-century Politics in the United States		4.66	.00
Geology and Landscape		3.44	.00
Desert Formation	2	0.31	.38

Passage Title	Form	T	p
Early Cinema		1.69	.05
The Origins Cetaceans		0.91	.18

จากตาราง 15 การศึกษาของจางพบว่า มีค่าอิทธิพลทดสอบเลขในแบบสอบชุดที่ 1 อยู่ระหว่าง 1.62 ถึง 4.66 และแบบสอบชุดที่ 2 เท่ากับ 0.31 ถึง 1.69 จากเกณฑ์ในข้างต้นสรุปได้ว่า อิทธิพลมีขนาดใหญ่ อย่างไรก็ตามจางได้กล่าวว่าการใช้โมเดลเชิงวินิจฉัยทางพุทธิปัญญาที่ประกอบด้วยทักษะย่อย ๆ ที่มีความสัมพันธ์กัน และข้อสอบหนึ่งข้อสามารถวัดทักษะที่มากกว่าหนึ่งทักษะ เมื่อวิเคราะห์ด้วยโมเดลทางสถิติไม่พบว่าเกิดปัญหา

จากที่กล่าวมาสรุปได้ว่า อิทธิพลทดสอบเลขเป็นผลกระทบของอิทธิพลแบบสุ่มที่เป็นปฏิสัมพันธ์ของการตอบข้อสอบกับทดสอบเลข ในการศึกษาครั้งนี้ ผู้วิจัยคำนวณโดยใช้โปรแกรม R ร่วมกับโปรแกรม WinBUGS โดยเรียกใช้งาน Package R2WinBUGS เพื่อได้ความชัดเจนเกี่ยวกับอิทธิพลทดสอบเลขของข้อสอบแบบทดสอบเลขที่มีความไม่เป็นอิสระกัน เมื่อถูกนำไปสร้างแบบสอบวินิจฉัยทางพุทธิปัญญา

2.4 การกำหนดขนาดของทดสอบเลขและความยาวของบทอ่าน

การศึกษาเกี่ยวกับการประมาณค่าโมเดลตอบสนองข้อสอบทดสอบเลขนั้น การกำหนดขนาดของทดสอบเลขที่เหมาะสมจะส่งผลต่อความน่าเชื่อถือของผลการวิจัย (Chen, 2010) จากการศึกษาขนาดตัวอย่างและขนาดของแบบสอบทดสอบเลขของงานวิจัยที่ผ่านมา มีรายละเอียดในตาราง 16 ดังนี้

ตาราง 16 แสดงขนาดของทดสอบเลขในการศึกษาเกี่ยวกับความสามารถการอ่านเพื่อความเข้าใจ

นักการศึกษา	ขนาดตัวอย่าง	ขนาดแบบสอบ
Jang (2005)	2,700	รวม 34 ข้อ (11 ข้อ 2 ทดสอบเลข) (12 ข้อ 1 ทดสอบเลข)
Li et al. (2006)	2,000	รวม 20 ข้อ (5 ข้อ : 4 Testlet)
Fukuhara&Kamata (2011)	1,000	รวม 42 ข้อ (6 ข้อ : 7 Testlet)
Eckes (2014)	2,859/2,214	รวม 25 ข้อ (7 ข้อ/8 ข้อ/10ข้ออย่างละ 1 Testlet)
Geramipour (2021)	4200/3220/4500	รวม 10 ข้อ (5 ข้อ : 2 Testlet)
Tonekaboni (2021)	2625	รวม 34 ข้อ (7 ข้อ 2 ทดสอบเลข/6 ข้อ

นักการศึกษา	ขนาดตัวอย่าง	ขนาดแบบสอบ
		1 เทสต์เลข)
Patarapichayatham & Kamata (2017)	1000	รวม 54 ข้อ (6 เทสต์เลข 9 ข้อ)

จากตาราง 16 นักการศึกษากำหนดขนาดเทสต์เลขที่หลากหลายรูปแบบ โดยเทสต์เลขมีขนาดอยู่ระหว่าง 2 ถึง 7 เทสต์เลข มีจำนวนข้อคำถามต่อ 1 เทสต์เลขอยู่ระหว่าง 5 - 12 ข้อ และมีจำนวนข้อรวม อยู่ระหว่าง 10 ถึง 54 ข้อ

2.5 ความยาวของข้อความและประเภทข้อคำถามที่ใช้ในแบบสอบ

การอ่านตามเนื้อเรื่องเป็นส่วนที่สำคัญที่สุดของแบบสอบเพื่อวัดความเข้าใจ ข้อคำถามการอ่านตามเนื้อเรื่องจะวัดความสามารถของนักเรียนในการอ่าน ทำความเข้าใจ และตีความ ข้อความอ่านเดี่ยวหรือสองข้อที่เกี่ยวข้อง เนื้อหาที่นำมาอาจเป็นเรื่องสั้น นิยาย สารคดี หรืออื่นๆไม่ว่าจะเป็นนิยายหรือสารคดี

2.5.1 ความยาวของข้อความในบทอ่านภาษาอังกฤษ

ความยาวของข้อความ (Length of passage) (จุฬาทิวเตอร์, 2558) จะมีทั้งข้อความแบบสั้นและข้อความแบบยาว สำหรับข้อความแบบสั้นในละตอนจะมีความยาวของบทอ่านหนึ่งย่อหน้า ซึ่งโดยทั่วไปเหมาะสมที่จะถามคำถาม 2 ข้อ สำหรับข้อความแบบยาวจะมีหลายย่อหน้าและมีคำถามจำนวนหลายข้อ ทั้งนี้จำนวนข้อคำถามขึ้นอยู่กับความยาวของข้อความ ข้อความทั้งสองประเภทมาพร้อมกับคำถามหลายประเภท ตัวอย่างเช่น ประเภทของคำถามคำศัพท์เป็นฐานที่ถามถึงคำพ้องความหมายของคำ คำถามเฉพาะบรรทัด และการจัดประโยคหรือย่อหน้าในเนื้อเรื่อง คำถามประเภทที่ยากที่สุดคือการขอให้คุณตอบเกี่ยวกับลักษณะทั่วไปของเนื้อเรื่องโดยรวม คำถามเกี่ยวกับข้อความทั้งสองประเภทจะจัดเรียงตามลำดับของข้อความที่พวกเขาพูดถึง ไม่ได้เรียงตามลำดับความยาก คำถามแรกอาจยากที่สุด คำถามสุดท้ายอาจง่ายที่สุด นอกจากนี้ คำถามเกี่ยวกับข้อความสั้นๆ ก็ไม่ง่ายเช่นกัน ทั้งข้อความแบบสั้นและแบบยาวอาจมีระดับความยากเท่ากัน ข้อความที่จับคู่หรือการเปรียบเทียบสองข้อความที่เกี่ยวข้องกันมีความยากกว่า ข้อความที่ถูกจับคู่เปรียบเทียบอาจเป็นแบบสั้นหรือแบบยาวก็ได้ หากเป็นข้อความคู่แบบสั้น ๆ ควรใช้คำถาม 4 ข้อ หากเป็นข้อความที่จับคู่แบบยาวๆ จำนวนคำถามควรเพิ่มมากขึ้น ทั้งนี้ขึ้นอยู่กับความยาวของข้อความ

2.5.2 ประเภทข้อคำถามที่ใช้ในแบบสอบ

คำถามเกี่ยวกับการอ่าน (Type of questions) โดยทั่วไปแบ่งออกเป็น 3 ประเภท (จุฬาทิวเตอร์, 2558)

1) คำศัพท์ (Vocabulary) หรือที่รู้จักกันในทางเทคนิคว่าคำศัพท์ในคำถามตามบริบท จะถามคุณเกี่ยวกับคำที่อยู่ในเนื้อเรื่อง เช่น จากเรื่องคำว่า 'heading' (บรรทัดที่ 5) หมายถึงข้อใด เป็นต้น คำศัพท์ในคำถามตามบริบทเหมาะสำหรับผู้สอบ เพราะเป็นข้อความที่ตรงไปตรงมาและตอบได้รวดเร็ว

2) คำถามเฉพาะเจาะจง (Specific question) คือคำถามที่มีคำตอบอยู่ในส่วนหนึ่งส่วนใดอยู่ในข้อความหนึ่งเท่านั้น โดยที่คำถามจะถามถึงบางสิ่งบางอย่างไม่น้อยกว่าหนึ่งคำ แต่ยังคงอยู่ในขอบเขตที่แคบและเป็นสิ่งที่มุ่งเน้น ได้แก่ คำถามระบุเฉพาะในบรรทัดหรือในแต่ละย่อหน้า คำถามระบุเฉพาะส่วนหนึ่งของข้อมูลที่ไม่ได้ ระบุแน่นอนด้วยหมายเลขบรรทัด ผู้สอบต้องค้นหาข้อมูลเอง โดยการอนุมาน เชื่อมโยง หรือสรุปเกี่ยวกับข้อมูลเฉพาะในบทความ

3) สารสำคัญของเรื่อง (Main Idea) เป็นคำถามเกี่ยวกับสาระสำคัญหรือใจความสำคัญของเรื่อง เป็นคำถามการอ่านที่ยากที่สุดในการทดสอบ และมักใช้เวลาในการตอบนานที่สุด เพราะวัดความเข้าใจเนื้อเรื่องโดยรวม

2.6 รูปแบบข้อสอบแบบทดสอบหลายตัวเลือก

ในด้านรูปแบบนั้นข้อสอบชุดทดสอบที่มีทั้งที่เป็นปรนัยและอัตนัยจึงทำให้การประเมินทางการศึกษาในทางปฏิบัตินั้นมีการสร้างแบบสอบมาตรฐานที่มีลักษณะทดสอบแบบใช้กากขิ้น โดยส่วนใหญ่ใช้รูปแบบข้อสอบแบบหลายตัวเลือก (Multiple choice: MC) ซึ่งมีความเหมาะสมและเป็นประโยชน์เมื่อใช้ในการประเมินกับคนกลุ่มขนาดใหญ่ (Large scale assessment) การให้คะแนนมีความเป็นปรนัย ตรวจสอบให้คะแนนได้ง่ายประหยัดเวลา และครอบคลุมเนื้อหา ผู้สอบจะเห็นบทความและมีกลุ่มข้อความ ที่เกี่ยวข้องกับบทความ ในการเรียงลำดับข้อความจะเรียงจากง่ายไปหายาก ซึ่งการที่มีข้อความเป็นชุดแบบนี้ จะสอดคล้องกับลักษณะของข้อสอบตามทฤษฎีตอบสนองข้อสอบ (Item Response Theory: IRT)

การทดสอบ TOFEL ด้านการอ่านเพื่อความเข้าใจ (Reading comprehension) ข้อสอบในส่วนนี้ จะมีลักษณะเป็นชุดข้อสอบ โดยกลุ่มข้อสอบประมาณ 6 ถึง 12 ข้อจะใช้บทความ (Passage) ที่มีความยาวประมาณ 250 ถึง 400 คำ เป็นสิ่งเร้าร่วมกันเพื่อให้ผู้สอบตอบข้อสอบ ซึ่งการตอบภายใน ชุดข้อสอบถูกหรือไม่นั้นไม่ได้ขึ้นอยู่กับความสามารถของผู้สอบเพียงอย่างเดียวแต่ขึ้นอยู่กับ ความเข้าใจในการตีความหรือการแปลความจากสิ่งเร้าและอาจจะเป็นไปได้ว่าผู้สอบเกิดการเรียนรู้จากการตอบข้ออื่น ๆ ในชุดข้อสอบนั้นๆ ดังนั้นการตอบข้อสอบก่อนหน้าอาจส่งผลต่อ การตอบในข้อถัดไป ถ้าผู้สอบเข้าใจสิ่งเร้าได้ถูกต้อง โอกาสในการตอบข้อสอบถูกต้องสูง ส่งผลให้ การตอบข้อต่อไปถูกต้องสูงด้วย แต่ถ้ามองในทางกลับกัน ถ้าผู้สอบไม่เข้าใจสิ่งเร้า แล้วทำข้อสอบข้อแรกผิด โอกาสในการตอบข้อต่อไปก็อาจจะผิดด้วย

จาง (Jang, 2005, p. 79) กล่าวว่า ความยาวของย่อหน้าและข้อความและระดับความถี่ของคำศัพท์มีความคล้ายคลึงกันมากในข้อความที่ว่าข้อความนั้นแตกต่างกันตามโครงสร้างองค์กรที่เป็นข้อความ รายละเอียดคำอธิบายโครงสร้างองค์กรที่เป็นข้อความแสดงไว้ด้านล่างความถี่ของคำที่รายงานนั้นอิงตาม Brown Corpus with Concapp สำหรับ Windows ซึ่งสร้างขึ้นในเครื่องมือสร้างโปรไฟล์ข้อความความถี่ของคำจากเว็บไซต์ (<http://www.edict.com.hk/textanalyser/>)

Tonekaboni & Ravand (2021, p. 65) การศึกษาแบบทดสอบการอ่านเพื่อความเข้าใจภาษาอังกฤษของ โตนิกาโบนี ประกอบด้วยบทอ่าน 3 บท ได้แก่ บทอ่านภาษาอังกฤษเรื่องแรกเป็นเรียงความ 5 ย่อหน้า 461 คำ ประกอบด้วยข้อคำถาม 7 ข้อ แนะนำการศึกษาวิจัย 2 เรื่องที่ทำในสาขาวิชาจิตวิทยา โดยเน้นที่การเห็นคุณค่าในตนเองและความพึงพอใจในชีวิตในย่อหน้าแรก บทอ่านที่สองประกอบด้วยสี่ย่อหน้า 417 คำ เป็นเรื่องของเด็กออทิสติก ความต้องการและวิธีการรักษา ประกอบด้วยคำถาม 7 ข้อ เรื่องที่ 2 นี้เป็นการที่ผู้อ่านแก้ปัญหา/วิธีแก้ไข เรื่องที่ 3 เป็นบทความห้าย่อหน้าจำนวน 426 คำอภิปรายประสิทธิภาพของการทดสอบความฉลาดในการวัดโครงสร้างที่พวกเขาอ้างว่าจะวัด ความคิดหลักส่วนใหญ่จะนำเสนอในตอนต้นของแต่ละย่อหน้า ยกเว้นวรรคแรกที่แนวคิดหลักถูกนำเสนอในประโยคสุดท้ายในย่อหน้า ประกอบด้วย 6 คำถาม รวมข้อสอบทดสอบทั้งหมดทั้งสิ้น 34 ข้อ และมีแบบอัตนัยเขียนตอบอีกด้วย ตาราง 17

ตาราง 17 แสดงความยาวของบทอ่านภาษาอังกฤษที่พบในงานวิจัย 3 เรื่อง

อ้างอิง	จำนวนคำในบทอ่าน	จำนวนข้อคำถาม
TOFEL	250-400	6-12 ข้อ
Jang (2005)	609-658	11-12 ข้อ
Tonekaboni & Ravand (2021)	417-461	6-7 ข้อ

สรุปได้ว่า ความยาวของข้อความที่เหมาะสมกับผู้เรียนในระดับพื้นฐานควรมีจำนวนคำอยู่ระหว่าง 250-450 คำ และสามารถตรวจสอบจำนวนคำด้วยโปรแกรมคอมพิวเตอร์ โปรไฟล์ข้อความความถี่ของคำจากเว็บไซต์ <http://www.edict.com.hk/textanalyser/> จำนวนข้อคำถามขึ้นอยู่กับความยาวของข้อความ สำหรับข้อความแบบยาวจะมีหลายย่อหน้าและมีคำถามจำนวนหลายข้อ ทั้งนี้จำนวนข้อคำถามขึ้นอยู่กับความยาวของข้อความ

งานวิจัยครั้งนี้ ผู้วิจัยกำหนดบทอ่านภาษาอังกฤษ (Passages) โดยกำหนดขอบเขตเนื้อหาอิงเป้าหมายตามกรอบอ้างอิงความสามารถทางภาษา CEFR ที่มีความยาวประมาณ

250 ถึง 450 คำ ทั้งนี้จำนวนข้อคำถามขึ้นอยู่กับความยาวของข้อความ ซึ่งควรมีจำนวนข้อสอบน้อยๆ เนื่องจากการลดจำนวนข้อสอบลงก็สามารถลดการชี้นำคำตอบซึ่งกันและกันของข้อสอบลงด้วย นอกจากนี้ เนื่องจากข้อสอบหนึ่งข้อต้องใช้ความสามารถหลายอย่างในการตอบทำให้ข้อสอบมีความยากสูงและมีความซับซ้อนซึ่งจะลดการชี้นำคำตอบระหว่างข้อได้อีกระดับหนึ่ง รูปแบบข้อสอบแบบหลายตัวเลือก (Multiple choice: MC) ซึ่งมีความเหมาะสมและเป็นประโยชน์เมื่อใช้ในการประเมินกับคนกลุ่มขนาดใหญ่ การให้คะแนนมีความเป็นปรนัย ตรวจสอบให้คะแนนได้ง่ายประหยัดเวลา และครอบคลุมเนื้อหา

2.7 ประเภทของข้อสอบทดสอบ

ทดสอบ (testlet test) สามารถแบ่งได้ เป็น 2 ประเภทตามวิธีการจัดกลุ่มข้อสอบ (จกมล บัวแก้ว, 2560, หน้า 14) คือ ใช้สิ่งเร้าร่วมกันและไม่ใช้สิ่งเร้าร่วมกัน

1) กลุ่มข้อสอบที่ใช้สิ่งเร้าร่วมกัน เช่น แบบทดสอบทางภาษา แบบทดสอบการอ่านเพื่อทำความเข้าใจ แบบทดสอบเชิงสถานการณ์ เป็นต้น

2) กลุ่มข้อสอบที่ใช้เนื้อหาเดียวกัน แต่ไม่ต้องใช้สิ่งเร้าร่วมกัน เช่น แบบทดสอบย่อยที่ประกอบด้วยข้อสอบเดี่ยว (single item)

ในปัจจุบันแบบสอบมาตรฐาน อาทิ PISA หรือแบบสอบระดับชาติ เช่น O-NET พบว่ามีการนำข้อสอบแบบทดสอบมาใช้ในการทดสอบระดับชาติในรายวิชาต่าง ๆ อาทิ ภาษาอังกฤษ วิชาภาษาไทย วิชาคณิตศาสตร์

ทั้งรูปแบบข้อสอบปรนัยและข้อสอบอัตนัยที่มีลักษณะของทดสอบเป็นข้อสอบเชิงสถานการณ์

3. ตัวอย่างข้อสอบที่มีลักษณะทดสอบและงานวิจัยที่เกี่ยวข้อง

3.1 ตัวอย่างข้อสอบที่มีลักษณะทดสอบ

ตัวอย่างข้อสอบที่มีลักษณะทดสอบในรายวิชาภาษาไทยระดับชั้นมัธยมศึกษาปีที่ 3 โดยข้อคำถามที่ 11 และ 12 ใช้บทสนทนาเป็นสิ่งเร้าร่วมกันในการตอบข้อสอบ ผู้สอบจะสามารถตอบข้อสอบได้ถูกต้อง โดยอาศัยข้อมูลจากบทสนทนาที่กำหนดให้ในแบบสอบ โดยข้อสอบทั้ง 2 ข้ออาจมีการพึ่งพาคำตอบกัน เช่น ผู้สอบสามารถตอบข้อสอบถูกโดยอาศัยร่องรอยจากอีกข้อหนึ่ง เป็นต้น

อ่านบทสนทนาต่อไปนี้แล้วตอบคำถามข้อ ๑๑ และข้อ ๑๒

นักเรียน ๑ แสดงความเห็น การวาดรูปของนักเรียน ๒ ดังนี้

นักเรียน ๑	ไม่ไหว ไม่ไหว วาดแบบนี้ใช้ไม่ได้ รวดค้ำเพื่องค้ำก้นมันเหมือนรูดค้ำเพื่องจริง ๆ มากกินไป
นักเรียน ๒	วาดให้เหมือนของจริง ไม่ได้หรือ
นักเรียน ๑	การวาดให้เหมือนของจริง ไคร ๆ ก็ทำได้ ถ้าฉันจะวาดรูดค้ำเพื่อง ฉันคงจะวาดให้เหมือนซาเลาปาลีแกลงมากกว่า
นักเรียน ๒	แล้วทำไมต้องเป็นอย่างนั้นด้วยละ
นักเรียน ๑	ทำไมนะหรือ ก็เพราะการวาดแค่สิ่งที่ตามองเห็น จะไม่มีค่าอะไรเลยนะสิ รูปที่ลึกลับคือรูปที่วาดในสิ่งที่ใจมองเห็นต่างหาก

๑๑. ข้อใดกล่าวได้แย้งกับบทสนทนาข้างต้น

๑. การวาดรูปให้ดีต้องวาดให้เหมือนจริง และมองเห็นชัดเจน
๒. การวาดรูปให้เหมือนของจริงมาก ๆ เป็นสิ่งที่ทุกคนทำได้
๓. รูปที่วาดจะมีคุณค่าสูงขึ้น ต้องวาดรูปตามที่ใจของผู้วาดมองเห็น
๔. ผู้วาดรูปที่พยายามวาดรูปให้เหมือนของจริงเป็นการวาดรูปตามที่ตามองเห็น

๑๒. บทสนทนาข้างต้นสนับสนุนเรื่องใด

๑. พัฒนาการ
๒. นันทนาการ
๓. จินตนาการ
๔. ศิลปวิทยาการ

แบบสอบฉบับนี้ มีสิ่งเร้าร่วมกันเป็นบทสนทนา ใช้ตอบข้อคำถาม 2 ข้อ

ภาพ 29 ข้อสอบแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนรายวิชาภาษาไทยระดับชั้นมัธยมศึกษาปีที่ 3 (สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน), 2563)

ต่อไปเป็นตัวอย่างข้อสอบที่มีลักษณะทดสอบวัดผลสัมฤทธิ์ทางการเรียนรายวิชาภาษาอังกฤษระดับชั้นมัธยมศึกษาปีที่ 3 โดยข้อคำถามที่ 16, 17 และ 18 ใช้บทอ่านเป็นสิ่งเร้าร่วมกันในการตอบข้อสอบ

Reading 3: Items 16 -18

Situation: Read the news article.

(1) Europe's first underwater restaurant has recently opened in Lindesnes, Norway. The restaurant is called 'Under', which has a double meaning. Not only is the restaurant underwater, but the word 'under' also means 'wonder' in Norwegian.

(2) The restaurant offers visitors the chance to eat seafood while looking out at life under the ocean. At the bottom of the building is the main dining room, which rests on the sea floor and can seat 40 people. This dining room has a huge window that is 36 feet (11 meters) wide and 11 feet (3.4 meters) high. At night, special lighting will be used to attract fish and other sea creatures in front of the window.

(3) 'Under' will not just be used by people looking for a meal. When the restaurant is not open, it will be used by scientists researching sea life. The restaurant owners will work with scientists to make sure that they use sea creatures in a way that has little effect on the undersea environment. The restaurant is suitable for the scientists to study small sea creatures over a long period of time.

16. According to the news article, which sentence is **TRUE** about the underwater restaurant?

1. It is the first underwater restaurant in the world.
2. The restaurant will also be used to catch fish.
3. Diners can watch the marine life pass by as they eat.
4. More than 40 visitors can seat in the main dining room.

17. What is the second paragraph mainly about?

1. The design of the restaurant
2. The protection of marine life
3. The location of the restaurant
4. The history of the restaurant

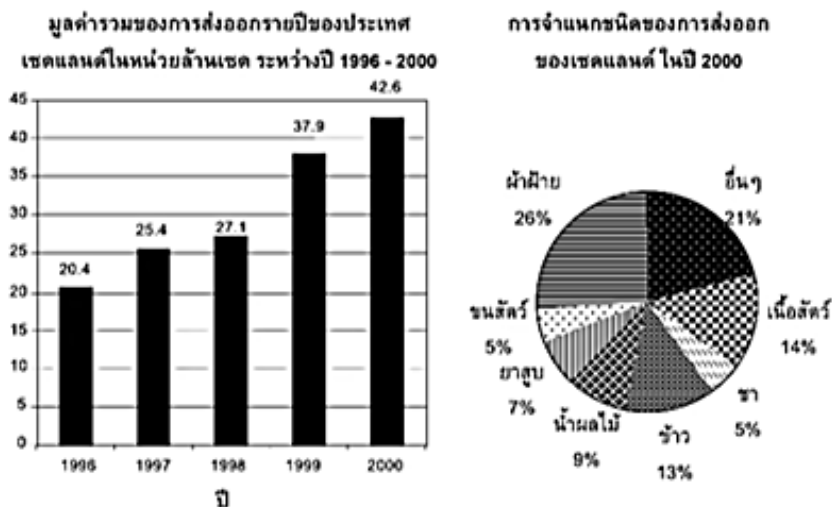
18. Which of the following can be inferred about the underwater restaurant?

1. Its name means 'puzzle' in Norwegian language.
2. Its owners are concerned about underwater environment.
3. It can be used as a place to create a new species of fish.
4. It is a suitable place to keep fish and other water animals.

สิ่งนี้เป็นบทความจำนวน 3 ย่อหน้า ตอบข้อคำถาม 3 ข้อ คือ 16-18

ภาพ 30 ข้อสอบแบบทดสอบเลทวิชาภาษาอังกฤษชั้นมัธยมศึกษาปีที่ 3
(สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) , 2563)

2) กลุ่มข้อสอบที่ใช้เนื้อหาเดียวกัน แต่ไม่ต้องใช้สิ่งเร้าร่วมกัน เช่น แบบทดสอบย่อยที่ประกอบด้วยข้อสอบเดี่ยว (single item)



คำถามที่ 3 : การส่งออก	M438Q01 - 0 1 9
ในปี 1998 มูลค่ารวมการส่งออกของประเทศเขตแดนเป็นเงินเท่าไร (หน่วยเป็นล้านเซต)	
คำตอบ:	
คำถามที่ 4 : การส่งออก	M438Q02
มูลค่ารวมการส่งออกน้ำตาลไม้จากประเทศเขตแดนในปี 2000 เป็นเท่าไร	
<ol style="list-style-type: none"> 1. 1.8 ล้านเซต 2. 2.3 ล้านเซต 3. 2.4 ล้านเซต 4. 3.4 ล้านเซต 5. 3.8 ล้านเซต 	

ภาพ 31 ข้อสอบเชิงสถานการณ์ข้อสอบ PISA วิชาคณิตศาสตร์

จากตัวอย่างข้อคำถามที่นำเสนอในข้างต้น แบบสอบทดสอบเหล่านี้ทั้งที่เป็นปรนัยและอัตนัย จึงทำให้การประเมินทางการศึกษาในทางปฏิบัตินั้นมีการลดหรือขยายโครงสร้างการตอบสนองทดสอบเหล่านี้มาใช้มากขึ้นและ นำมาใช้ในการสร้างแบบสอบมาตรฐานแบบดั้งเดิม โดยส่วนใหญ่มีรูปแบบข้อสอบแบบหลายตัวเลือก (Multiple choice: MC) ซึ่งมีประโยชน์มากในการประเมินกับคนกลุ่มขนาดใหญ่ (Large scale assessment) ได้แก่ 1) ผู้ประเมินไม่มีอคติ เมื่อตรวจข้อสอบแบบหลายตัวเลือก หรือที่เรียกว่ามีความเป็นปรนัย 2) การทดสอบหนึ่งครั้งสามารถใช้ข้อสอบแบบเลือกตอบได้หลายข้อ และครอบคลุมเนื้อหา 3) การให้คะแนนมีความเป็นปรนัย และจำนวนข้อสอบที่มากเป็นการเพิ่ม ความเชื่อมั่นและความตรงของแบบสอบ และ 4) การให้คะแนนของข้อสอบแบบหลายตัวเลือก สามารถทำได้ง่ายและประหยัดเวลา (Chen, 2014, p.11, อ้างถึงใน จงกล บัวแก้ว, 2560, หน้า 16)

3.2 ตัวอย่างข้อความที่มุ่งวัดหลายคุณลักษณะ

Edward Mustang President
Whispering Pines Inn and Resort
P.O. Box 65
Upper River, New Brunswick Canada
Dear. Mustang,

I am writing in regard to my recent stay at Whispering Pines. I have spent my annual summer vacation there for the past four or five years and have always enjoyed it. The comfortable accommodations and delicious menu are a big attraction for me. This year, however, the resort seemed to be lacking in the area of customer service. I enjoyed my three daily meals that came with my room. However, when on the last day of my stay I decided to try out the inn's high tea, I had a disappointing experience. The food was delicious, but the waitress was sullen and rude. Also, this year I decided to take golf lessons instead of my usual tennis lessons. I am a complete beginner and the instructor had no patience with me. He yelled at my mistakes and mad me feel very uncomfortable. Despite these issues, I am not thinking about vacationing elsewhere. I plan to return to whispering Pines and may even try another golf lesson. However, Whispering Pines is a high-quality resort, and I thought you should know about these things.

Sincerely,
Lia Kim

Lia Kim
1165 Putnam Avenue
Croton, NY
Dear Ms.Kim

I was very sorry to hear about your recent unpleasant experience at the Whispering Pines Inn and Resort. As you know from your previous stays at Whispering Pines, we do everything possible to ensure the comfort of our guests and we are widely know for our excellent accommodations and five-star menu. I sincerely regret the problems you had with your meal and your instructor. I will be in contact with the manager of the inn to discuss these issue. In the meantime, please accept the enclose coupon. It entitles you to the same special meal you enjoyed on your last day, and I am sure your next time you will have a better experience. I am glad to hear that you plan on being our guest again. You may be interested to know that in addition to golf and tennis, next year we will be adding a system of hiking trails and an indoor pool. We look forward to seeing you again at Whispering Pines.

Sincerely
Edward Mustang
President

Read two emails above and answer the items 1-4

Q1 : Why did Ms. Lia write the letter?

To complain about some employees ✓

To praise the accommodations

To ask about the menu

To make reservations

(A) Ms. Lia wrote to tell Mr. Mustang about her problems with a sullen waitress and impatient golf instructor. Choice (B) is mentioned but is not the reason for the letter. Choice (C) repeats the word menu. Choice (D) is associated with the context of a vacation resort.

ข้อ (A) ถูกต้อง คุณ Lia เขียนถึงคุณมัสแตงเกี่ยวกับปัญหาของเธอกับพนักงานเสิร์ฟที่บูดบึ้ง และครูสอนกอล์ฟที่ใจร้อน มีการกล่าวถึงตัวเลือก (B) แต่ไม่ใช่สาเหตุของจดหมาย ตัวเลือก (C) ทำซ้ำเมนูค่า ตัวเลือก (D) เกี่ยวข้องกับบริบทของรีสอร์ทตากอากาศ

ข้อคำถามแบบทดสอบเลขที่ 4 ข้อตอบสนองสิ่งเร้าเดียวกันคือ อีเมลล์ทั้งสองฉบับ ข้อคำถามดังกล่าวต้องการเหตุผลว่าทำไมคุณเลียจึงเขียนอีเมลล์ถึงมัสแตง ซึ่งผู้ตอบต้องมีความเข้าใจในเนื้อหาของอีเมลล์ทั้งสองฉบับ (ความรู้คำศัพท์ / ไวยากรณ์) รู้สาระสำคัญของจดหมายอย่างถูกต้อง (ระบุนใจความสำคัญ) จากนั้นทำการเชื่อมโยงและสรุปอ้างอิงในสิ่งที่ไม่ได้ระบุชัดเจน เกี่ยวกับเหตุผลที่เขียน (สรุปอ้างอิง)

Q2. How often does Ms. Lia visit Whispering Pines?

- Every week
- Every month
- Every year ✓
- Every four or five years

ข้อ (C) ถูกต้อง เพราะคุณ Lia ใช้เวลาวันหยุดประจำปีหรือรายปีของเธอที่ Whispering Pines ตัวเลือก (A), (B) ไม่ได้กล่าวถึง ซอยส์ (D) สับสนกับจำนวนปีที่เธอมาเยี่ยมรีสอร์ท

Q3. What can Ms. Lia get with the coupon Mr. Mustang sent?

- High tea ✓
- Three daily meals
- A golf lesson
- A room at the inn

ข้อ (A) ถูกต้องเพราะคุณมัสแตงบอกว่าคุณเลียสามารถเพลิดเพลินกับอาหารมื้อพิเศษแบบเดียวกับที่เธอชอบในวันสุดท้ายที่รีสอร์ท ซึ่งก็คือไฮที ตัวเลือก (B) (C) และ (D) คือทุกสิ่งที่มีที่รีสอร์ทแต่ไม่ใช่สิ่งที่ใช้คูปอง ผู้สอบต้องมีความเข้าใจเกี่ยวกับวัฒนธรรมเจ้าของภาษาและเชื่อมโยงได้

Q4. What will Ms. Lia probably do on her next summer vacation?

Hike

Give up golf

Return to Whispering Pines ✓

Go to another resort

ข้อ (C) ถูกต้องเพราะ คุณ Lia ใช้เวลาทุกวันหยุดฤดูร้อนที่ Whispering Pines มาหลายปีแล้ว และเธอบอกว่าเธอมีแผนจะกลับไปที่นั่นอีกครั้ง ดังนั้นนั่นอาจเป็นสิ่งที่เธอจะทำในวันหยุดครั้งต่อไป Choice (A) เป็นกิจกรรมที่ริสอร์ทมีให้ แต่คุณ Lia ไม่เคยพูดถึงเลย ตัวเลือก (B) และ (D) คือสิ่งที่เธอบอกว่าจะไม่ทำ

3.3 ข้อดีและข้อจำกัดของเทสต์เลท

กระบวนการคิดขั้นสูงควรเป็นเป้าหมายที่สำคัญของการจัดการศึกษา และยังพบเครื่องมือประเมินความคิดขั้นสูงน้อยมาก ซึ่งใช้วัดความรู้ความสามารถหรือทักษะทางปัญญาที่ซับซ้อน ในประเด็นนี้ เวินเนอร์ และคณะแสดงทัศนะไว้ว่า ด้วยลักษณะเฉพาะของเทสต์เลทอาจสามารถพัฒนาเพื่อใช้แก้ปัญหาดังกล่าวได้ด้วยเหตุผล 4 ประการ (Wainer et al., 2000; DeMars, 2006 cited in Chen, 2014, p.12 จงกล บัวแก้ว, 2560) ดังนี้

1) การรวมกันของเทสต์เลทในแบบทดสอบจะยังคงความเป็นแบบสอบหลาย ตัวเลือกทั้งหมด เช่น การให้คะแนนที่มีความเป็นปรนัย การครอบคลุมเนื้อหาที่ออกสอบ

2) เทสต์เลทประกอบด้วยกลุ่มของข้อคำถามที่มีสิ่งเร้าเดียวกัน เป็นการรวมเนื้อหาในการทดสอบเพิ่มขึ้นและ ช่วยลดความกังวลเกี่ยวกับการใช้ข้อคำถามที่เป็นอิสระกันหรือคำถามเดียว

3) เทสต์เลทจะมีข้อคำถามหลายข้อที่ใช้เนื้อหาที่เป็นสิ่งเร้าเดียวกัน เป็นการเพิ่มประสิทธิภาพในการพัฒนาแบบสอบ และดีสำหรับการใช้เวลาการทำข้อสอบของผู้เข้าสอบ

4) การใช้ชุดข้อสอบที่มีความซับซ้อน และมีความสัมพันธ์กันภายในแบบสอบทำให้แบบสอบที่สร้างขึ้นมีความยืดหยุ่นมากขึ้นในการเขียนข้อสอบที่มีการประเมินต่างกันเกี่ยวกับลักษณะและระดับขั้นของการวัดพุทธิพิสัย ดังนั้นข้อสอบแบบเทสต์เลทจึงได้รับการพิจารณาในการใช้จริงมากขึ้น และมีความเหมาะสมกับการวัด ทักษะในระดับสูงที่ติกว่า

การที่ข้อสอบชี้คำตอบซึ่งกันและกันในทฤษฎีการทดสอบแบบดั้งเดิมถือว่าเป็นลักษณะของข้อสอบที่ไม่ดี อย่างไรก็ตามทฤษฎีการทดสอบแบบดั้งเดิมจะมองความคลาดเคลื่อนเป็นรายฉบับ จึงยังไม่สามารถแก้ปัญหาในประเด็นนี้ได้อย่างชัดเจน แต่ในทางตรงข้าม ทฤษฎีการตอบสนองข้อสอบที่ได้รับการพัฒนาภายหลังมุ่งที่จะประมาณค่าพารามิเตอร์ของข้อสอบแต่ละข้อ เช่น พารามิเตอร์ความยาก(b) พารามิเตอร์อำนาจจำแนก (a) โอกาสการเดาได้ (c) จึงถือว่า IRT สามารถวิเคราะห์ราย

ข้อได้ และมีการวัดความคลาดเคลื่อนรายข้อทำให้ทฤษฎีการวัดแบบดั้งเดิมมีความคลาดเคลื่อนในการวัดมากกว่าทฤษฎีการตอบสนองข้อสอบ

3.4 ความเป็นอิสระ (Local Independence)

ความเป็นอิสระ หมายถึง เมื่อควบคุมคุณลักษณะหรือความสามารถของบุคคลให้คงที่แล้ว ผลการตอบข้อคำถามในข้อคำถามแต่ละคู่ จะต้องมีความเป็นอิสระจากกันทางสถิติ หรือกล่าวอีกนัยหนึ่งได้ว่า คุณลักษณะแฝงภายในบุคคล ต้องเป็นปัจจัยเดียวที่ทำให้การตอบข้อคำถามของผู้ทดสอบมีความแตกต่างกัน ความแตกต่างของการตอบข้อคำถามต้องไม่เกิดขึ้นจากความสัมพันธ์รูปแบบอื่น หากพิจารณาอย่างผิวเผิน แนวคิดเรื่องความเป็นอิสระของข้อคำถามอาจดูขัดแย้งกับแนวคิดเรื่องความสอดคล้องภายในของข้อคำถามตามทฤษฎีการทดสอบแบบดั้งเดิม ที่ระบุว่าข้อคำถามที่มีความสัมพันธ์กันมาก แสดงให้เห็นว่ามาตรวัดนั้นมีความเที่ยงในระดับสูง อย่างไรก็ตามเมื่อพิจารณาอย่างละเอียดจะพบว่า แนวคิดตามทฤษฎีทั้งสองนั้น มิได้ขัดแย้งกัน เพราะความเป็นอิสระของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบนั้น เป็นความอิสระของข้อคำถามที่เกิดขึ้นภายหลังการควบคุมพารามิเตอร์บุคคล และพารามิเตอร์ข้อสอบที่มีอิทธิพลโดยตรงต่อการตอบข้อคำถามในโมเดลแล้ว (สุชาติ สกลกิจรุ่งโรจน์, 2558, หน้า 27-28)

การพึ่งพาคำตอบกันของข้อสอบในรูปแบบของชุดทดสอบเหล่านั้น ทำให้ข้อสอบแต่ละข้อไม่เป็นอิสระจากกัน (Local Item Dependence: LID) ซึ่งเป็นการละเมิดข้อตกลงเบื้องต้นที่สำคัญของ IRT สองประการคือ ข้อสอบต้องมุ่งวัดลักษณะเดียวกัน (Unidimensionality) และมีความเป็นอิสระระหว่างข้อคำถาม (Independence) (ศิริชัย กาญจนวาสี, 2555, หน้า 77) นอกจากนี้ข้อสอบแบบทดสอบเหล่านี้ยังมีการวัดคุณลักษณะหลายด้านในข้อเดียวกันซึ่งสอดคล้องกับทฤษฎีการตอบสนองแบบพหุมิติ (Multidimensional Item Response Theory: MIRT) ที่ไม่มีข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกมิติและความเป็นอิสระระหว่างข้อสอบ ซึ่งการละเมิดข้อตกลงเบื้องต้นที่กล่าวมานั้นส่งผลต่อการประมาณค่าความเที่ยงของข้อสอบจะเกินจริงประมาณ 10 - 15 % เนื่องจากการตอบถูกของแต่ละข้อไม่เป็นอิสระจากกัน (Sireci et al., 1991, p. 238) แบบสอบที่มีลักษณะของทดสอบเหล่านี้จะทำให้มีแหล่งความแปรปรวนและปัจจัยเนื่องจากทดสอบเพิ่มขึ้นซึ่งหากไม่คำนึงถึงอิทธิพลของทดสอบเหล่านี้จะมีผลให้ประมาณค่าความเที่ยง (Reliability) ความคลาดเคลื่อนมาตรฐาน (Standard Error) และอำนาจจำแนกไม่ถูกต้อง (Christine, 2012) เพื่อให้การใช้โมเดล IRT ประมาณค่าได้ถูกต้องจะต้องเพิ่มความยาวของแบบสอบเป็น 2 เท่าเพื่อชดเชยค่าความเที่ยงที่ประมาณเกินจริงไป (Sireci et al., 1991)

3.5 แนวทางในการแก้ปัญหาความไม่เป็นอิสระกัน

จากปัญหาความไม่เป็นอิสระในการตอบข้อสอบนี้ มีนักวิจัยได้เสนอ แนวทางในการแก้ปัญหาไว้ ดังนี้ (อริสฟา เตหลิม, 2559, หน้า 49-50; จงกล บัวแก้ว, 2560, หน้า 19-20; พิสิฐ พิณจสกุล, 2563, หน้า 79)

1. นำการวิเคราะห์ของคะแนนแบบหลายค่ามาประยุกต์ใช้ (Polytomous scoring) โดยการรวมข้อสอบที่ตรวจให้คะแนนแบบสองค่าในแบบสอบทดสอบให้เป็นข้อสอบหนึ่งข้อที่มีการตรวจให้คะแนนมากกว่าหนึ่งค่า (เช่น Samejima's, 1969; graded response model, Bock's 1972; nominal response model, and Muraki's, 1992; generalized partial credit model) ซึ่งการทำลักษณะเช่นนี้จะมิข้อจำกัดเกี่ยวกับสารสนเทศของข้อสอบที่เป็นรายข้อจะไม่สามารถ วิเคราะห์ได้

2. การกำหนดให้ชุดข้อสอบถือเป็นหน่วยในการวิเคราะห์ด้วย โดยคำนวณคะแนนที่ ตอบถูกในแต่ละชุดข้อสอบแฝงในชุดข้อสอบ (เช่น Lee & Frisbie, 1999; Lee et al., 2001; Sireci, Thissen, & Wainer, 1991; Thissen, Steinberg, & Mooney, 1989; Wainer & Thissen, 1996) ซึ่งวิธีนี้จะทำให้ข้อสอบที่อยู่ระหว่างชุดข้อสอบมีความเป็นอิสระในการตอบ แต่จะไม่เป็นอิสระในการตอบภายในชุดข้อสอบทดสอบเหล่านั้นๆ

3. โมเดลทฤษฎีการตอบสนองทดสอบ (Testlet response theory models: TRT) โดยคำนึงถึงอิทธิพลของความไม่เป็นอิสระในการตอบ ด้วยการเพิ่มส่วนที่เป็น Random Effect ($Y_{id(j)}$) เข้าไปในสมการของโมเดลการตอบสนองข้อสอบแบบ IRT ซึ่งเป็นการผ่อนคลายข้อตกลงที่เกี่ยวกับความเป็นอิสระระหว่างข้อสอบ

4. ในการวิเคราะห์ทฤษฎีการตอบสนองข้อสอบแบบทดสอบ (TRT) เพื่อแก้ปัญหาความเป็นเอกพันธ์ (Unidimensional) และความเป็นอิสระจากกันของข้อคำถาม (Local Independence) โดย เป็นการใช้รูปแบบโมเดลของทฤษฎีการตอบสนองข้อสอบแบบพหุมิติแบบ 3 พารามิเตอร์ (3PL MIRT) เนื่องจากข้อตกลง เบื้องต้นของทฤษฎีการตอบสนองข้อสอบแบบพหุมิติสามารถที่จะวัดได้พร้อมกันหลายมิติและแต่ละมิติมีความสัมพันธ์กันได้

3.6 งานวิจัยที่เกี่ยวข้องกับการพัฒนาข้อสอบแบบทดสอบ

Eckes & Baghaei (2015, pp. 85-98) ได้ศึกษาวิจัยการใช้แบบสอบทดสอบเพื่อวัดความสามารถทางภาษาอังกฤษ นำเสนอการใช้ทฤษฎีการตอบสนองข้อสอบแบบทดสอบที่จะศึกษาความไม่เป็นอิสระต่อกันในข้อสอบแบบ C-test หรือเติมในช่องว่างซึ่งเป็นข้อสอบวัดความสามารถทางภาษา อิงตามทฤษฎีการตอบสนองข้อสอบแบบทดสอบ โดยมีการใช้โปรแกรม SCORIGHT ในการประมาณค่า 2 พารามิเตอร์ (2-PL TRT) โดยใช้กลุ่มตัวอย่าง 4,708 คนในการที่จะทำข้อสอบชุดทดสอบแบบ C-test จำนวน 8 บทความ ผลพบว่าขนาดอิทธิพลของทดสอบเหล่านั้นยังไม่มาก มีความแปรปรวนอยู่ระหว่าง 0.09 ถึง 0.18 และการประมาณค่าพารามิเตอร์ออกมาเป็นรูปแบบการให้คะแนนแบบมากกว่าสองค่า

(polytomous IRT) และทฤษฎีการตอบสนองข้อสอบแบบปกติ (IRT) ที่ละเว้นข้อตกลงเบื้องต้นเรื่องความเป็นอิสระของข้อสอบ (local independence) จากการประมาณค่าความเป็นอิสระจากกันแล้ว จึงสรุปได้ว่าข้อตกลงเบื้องต้นนี้ไม่เป็นปัญหาสำหรับข้อสอบชุด C-test แบบทดสอบหลาย

Masoud Geramipour (2021; p. 1-23) ทำการศึกษาเพื่อแก้ปัญหาความไม่เป็นอิสระกันของข้อสอบทดสอบหลายข้อสอบที่วัดความสามารถการอ่านเพื่อความเข้าใจ โดยใช้ Rasch testlet Model และ bi-factor model เป็นโมเดลการวัดสองแบบที่สามารถจัดการกับความไม่เป็นอิสระกันของข้อสอบ (LID) ในการประเมินมิติของความเข้าใจในการอ่านการทดสอบ การศึกษานี้มีวัตถุประสงค์เพื่อใช้โมเดลการวัดกับข้อมูลการตอบสนองรายการจริงของการทดสอบความเข้าใจในการอ่าน EFL ของอิหร่านและเปรียบเทียบความถูกต้องของปัจจัยสองส่วนโมเดลและพารามิเตอร์ข้อสอบที่สอดคล้องกันด้วยโมเดล Rasch แบบเอกมิติและ โมเดล Rasch แบบพหุมิติ ข้อมูลที่รวบรวมจากส่วนความเข้าใจในการอ่าน EFL ของการสอบเข้ามหาวิทยาลัยแห่งชาติอิหร่านตั้งแต่ปี 2016 ถึง 2018 ขึ้นสูงต่างๆโปรแกรม R ถูกใช้เพื่อให้พอดีกับ Rasch เอกมิติ โมเดล Rasch แบบพหุมิติ และ โมเดล bi-factor ซึ่งสำรวจและโมเดล bi-factor เจริญย่น จากนั้น ประมาณค่าพารามิเตอร์ข้อสอบ และระบุอิทธิพลของทดสอบ (Testlet effects) นอกจากนี้ดัชนีการวัด (G²) และความสัมพันธ์ของพารามิเตอร์ข้อสอบสำหรับโมเดลต่างๆถูกคำนวณ ผลการศึกษาพบว่า อิทธิพลทดสอบมีขนาดเล็กและไม่มีผลใดๆต่อการวัดสำหรับการทดสอบการอ่าน EFL ทั้งหมด นอกจากนี้ bi-factor model ยังเหนือกว่าโมเดลอื่นๆ ในเงื่อนไขของความพอดี (Model Fit) ในขณะที่โมเดล bi-factor เจริญย่นมีความตรงเชิงโครงสร้างของการทดสอบความเข้าใจในการอ่าน EFL อย่างไรก็ตามพารามิเตอร์ความยากของข้อสอบในโมเดล Rasch มีความสอดคล้องกันมากกว่าโมเดล bi-factor การศึกษามีนัยสำคัญสำหรับวิธีการจัดการกับความไม่เป็นอิสระกันของข้อสอบ (LID) และมีติในการประเมินความเข้าใจในการอ่านโดยอ้างอิงจากการทดสอบ EFL

ตอนที่ 4 มโนทัศน์เกี่ยวกับการให้ข้อมูลย้อนกลับ

การให้ข้อมูลย้อนกลับหรือผลการกระทำแก่ผู้เรียนเป็นสิ่งสำคัญ มีนักวิชาการหลายท่านได้ให้ข้อมูลที่น่าสนใจเกี่ยวกับความหมาย ความสำคัญและหลักการให้ข้อมูลย้อนกลับไว้ดังนี้

1. ความหมายและความสำคัญของข้อมูลย้อนกลับ

1.1 ความหมายของข้อมูลย้อนกลับ

ข้อมูลย้อนกลับ หมายถึง การให้ข้อมูลเกี่ยวกับผลการตอบหรือผลการกระทำแก่ผู้เรียนผ่านข้อความ เสียง รูปภาพ ภายหลังจากการกระทำกิจกรรมแต่ละครั้ง เพื่อให้ผู้เรียนทราบความก้าวหน้าของตนเองเนื่องจากข้อมูลย้อนกลับจะแสดงให้เห็นให้ผู้เรียนทราบว่าตนเองประสบผลสำเร็จในการเรียน หรือเข้าใจในเนื้อหาและสามารถแก้ปัญหาได้มากขึ้นเพียงใด ซึ่งผู้เรียนสามารถแสดงความ

คิดเห็นหรือแสดงออกต่อผลที่ได้รับ ถ้าหากไม่พอใจผลที่ได้รับก็จะแก้ไขปรับปรุงจุดบกพร่อง ให้เกิดการเรียนรู้ได้ดียิ่งขึ้น และเป็นการสร้างแรงจูงใจในการเรียนให้สูงขึ้น (Gagne ,1977; Bardwell, 1981; Rowntree, 1981; Kalish, 1984; Kazdin, 1984 : 104-105; Cater, 1984; Cohen, 1985; Cyboran, 1995; สุกัญญา นิมานันท์, 2531; นางฮัซมีนี ยูโซะ, 2555; จรรยา ศิรินิล 2559 หน้า 23; จุฑาภรณ์ มาสันเทียะ 2560) โดยวัตถุประสงค์และเป้าหมายของการให้ข้อมูลป้อนกลับนั้นไม่ใช่เพื่อการตัดสิน (Judgment) หรือลงโทษ (Punish) บุคคล แต่เป็นการกระตุ้นให้ผู้รับมีการเรียนรู้ เติบโต และเปลี่ยนแปลงของพฤติกรรมที่สำคัญที่มีผลกระทบจากผลการปฏิบัติงานที่เฉพาะเจาะจง และนำเสนอในแนวทางที่บุคคลสามารถยอมรับและปฏิบัติได้ (Armstrong, 2006, อ้างถึงใน ชลิตา ชาญวิจิตร, 2561, หน้า 292)

1.2 ความสำคัญของข้อมูลย้อนกลับ

การให้ข้อมูลย้อนกลับ (Feedback) ถือว่าเป็นหัวใจสำคัญของการประเมินความก้าวหน้าในการเรียนของผู้เรียน (Hattie & Timberley, 2007) และเป็นองค์ประกอบสำคัญในการทำให้นักเรียนเกิดการเรียนรู้ มีประโยชน์ในการเปลี่ยนแปลงพฤติกรรมในการเรียนการสอน อย่างไรก็ตามครูต้องคำนึงถึงวิธีการและเวลาที่ให้ข้อมูลจึงจะเกิดผลสำเร็จในการจัดการเรียนการสอน (ฮัซมีนี ยูโซะ, 2555) ข้อมูลย้อนกลับมีผลกระทบในด้านบวกและด้านลบโดยการให้ข้อมูลย้อนกลับจะให้แก่ผู้เรียนภายหลังจากที่ได้ตอบสนองต่อการเรียนรู้ในครั้งแรก เมื่อได้ให้ข้อมูลย้อนกลับแก่ผู้เรียนในมิติความรู้ ทักษะต่างๆ ที่ได้เรียนไปแล้วโดยเฉพาะอย่างยิ่งหากเป็นการให้ข้อมูลย้อนกลับเพื่อชี้แจงความผิดพลาดที่เกิดขึ้นจะทำให้ผู้เรียนเกิดความรู้ความเข้าใจที่ดียิ่งขึ้น (Kuthayy, 1977) ข้อมูลย้อนกลับเป็นสารสนเทศที่ส่งเสริมสนับสนุนความสามารถของแต่ละบุคคล มีวัตถุประสงค์หลัก คือ การลดช่องว่างระหว่างความรู้ความสามารถในปัจจุบัน และความรู้ความสามารถตามเป้าหมายที่ครูผู้เรียนหรือผู้ปกครองต้องการโดยมีการสร้างเป้าหมายร่วมกันตั้งแต่เริ่มต้น (พินดา วราสุนันท์. 2556 ก: 134, อ้างถึงใน อพันธ์วี พูลพุทธา, 2560 หน้า 9-18) และยังสามารถลดความแตกต่างระหว่างความรู้ ความเข้าใจในสภาพปัจจุบันที่เป็นอยู่และสภาพที่ต้องการให้เกิดขึ้นได้อีกด้วย (Hattie & Timberley, 2007) ซึ่งการที่จะลดความแตกต่างหรือช่องว่างระหว่างสภาพปัจจุบันและสภาพที่ต้องการให้เกิดขึ้นโดยผ่านการให้ข้อมูลย้อนกลับนั้น หากเป็นการแก้ปัญหาที่ยากหรืองานที่มีความท้าทาย และมีเป้าหมายที่คาดหวังไว้ชัดเจน ผู้เรียนจะมีความพยายามมากกว่าการปฏิบัติกิจกรรมเดิมซ้ำๆหรืองานที่ง่ายเกินไป (Kluger & De Nisi, 1996) ผู้เรียนยังสามารถพัฒนาทักษะในการติดตามข้อผิดพลาดของตนเองจากข้อมูลย้อนกลับของตนเอง เพื่อให้ไปถึงยังเป้าหมายที่ตั้งไว้ นอกจากนี้ผู้เรียนยังสามารถแสวงหากลยุทธ์ที่ดีขึ้นเพื่อให้งานหรือการเรียนรู้จากการได้รับข้อมูลเพิ่มเติมเพื่อพวกเขาจะสามารถแก้ปัญหาปัญหาหรือใช้เทคนิคที่มีในกำกับ ดูแลตนเอง ครูผู้สอนมีบทบาทในการช่วยลดความแตกต่างหรือช่องว่างระหว่างความรู้ ทักษะในปัจจุบัน และความรู้ ทักษะที่ต้องการให้เกิดขึ้นกับ

ผู้เรียน ซึ่งหมายถึงครูอาจจะต้องช่วยผู้เรียนในการกำหนดเป้าหมายการเรียนรู้ที่มีความชัดเจนมากยิ่งขึ้น ซึ่งเป้าหมายที่มีความชัดเจน (Specific Goals) นั้นจะดีกว่าเป้าหมายที่ไม่เฉพาะเจาะจง เพราะส่วนใหญ่เน้นไปที่ความสนใจของผู้เรียนและเน้นการให้ข้อมูลย้อนกลับที่มีข้อมูลสารสนเทศเกี่ยวกับการนำไปใช้แก้ไขปัญหามากกว่าเป้าหมายทั่วไปที่ไม่เฉพาะเจาะจง (จุฑาภรณ์ มาสันเทียะ, 2560, หน้า 28-29)

การให้ข้อมูลย้อนกลับเป็นการให้ข้อมูลผ่านข้อความ เสียง หรือรูปภาพ หลังจากการตอบสนอง ของผู้เรียนเพื่อให้ผู้เรียนได้ทราบว่าผลการตอบถูกต้องหรือไม่ อีกทั้งยังทำให้ทราบความก้าวหน้าของตนเอง นำไปสู่การทำความเข้าใจ แก้ไขปรับปรุงจุดบกพร่องในกรณีที่ไม่ดี และเป็นการส่งเสริมแรงจูงใจ แก่ผู้เรียน การให้ข้อมูลย้อนกลับถือว่าเป็นหัวใจสำคัญของการประเมินความก้าวหน้าในการเรียนของผู้เรียน วัตถุประสงค์หลักของการให้ข้อมูลย้อนกลับ คือ การลดช่องว่างระหว่างความสามารถในปัจจุบัน กับความสามารถที่ปรารถนาของผู้เรียน (ไพฑูริย์ เอื้อบุญประดิษฐ์, 2563, หน้า 69)

การให้ข้อมูลย้อนกลับเป็นแรงจูงใจ (motivation) ให้ผู้เรียนทราบความสำเร็จของตนเองอาจด้วยคำพูด หรือท่าทาง และเป็นแรงจูงใจกระตุ้น ให้ผู้เรียนเกิดการกระทำ หรือแสดงพฤติกรรมโต้ตอบ เพื่อให้ได้มาซึ่งความภาคภูมิใจในการกระทำของตนในครั้งต่อไป อีกทั้งการให้ข้อมูลย้อนกลับยังช่วยกำหนดพฤติกรรม (regulate behavior) เพราะข้อมูลย้อนกลับให้รายละเอียดที่เป็นประโยชน์ต่อการเรียนรู้ ผู้เรียนจึงสามารถปรับปรุงพฤติกรรม ของตนไปสู่วิธีการที่ถูกต้องหรือใกล้เคียงได้และการให้ข้อมูลย้อนกลับทำให้ผู้เรียนทราบผลการเรียน ของตนเองตลอดเวลาในขณะที่เรียน และได้เรียนไปตามความสามารถของตนเอง เป็นการเสริมให้ผู้เรียนมีความมั่นใจมากขึ้น ช่วยให้ผู้เรียนมีความพยายามและตั้งใจในการเรียนครั้งต่อไป และทำให้ความวิตกกังวลลดน้อยลง ช่วยปรับปรุงการเรียนรู้ของผู้เรียนเพื่อไม่ให้กระทำผิดซ้ำอีก และสามารถแก้ไขข้อบกพร่องของตนเองได้ ทำให้มีทัศนคติที่ดีต่อวิชาที่เรียนและการเรียน (สุกัญญา นิมานันท์, 2533; สุริรัตน์ คน, 2545; วรลาภ ดาราพาณิชย์, 2549, พรพิมล รอดเคราะห์, 2550, อ่างถึงใน ฮัซมีนี ยูโซะ, 2555, หน้า 16)

ศิลปชัย สุวรรณธาดา (2548, หน้า 37 - 39, อ่างถึงใน พลอยไพลิน นิลกรรณ, 2552, หน้า 10) ได้สรุปบทบาทสำคัญของผลย้อนกลับต่อการเรียนรู้ไว้ 3 ประการ ดังนี้

1) ผลย้อนกลับแก้ไขความผิดพลาด (Correction) ผลย้อนกลับจะเป็นข้อมูลที่บอกให้ผู้เรียนรู้ว่าการแสดงทักษะของตนอยู่ห่างจากเกณฑ์ มาตรฐานหรือเป็นการแสดงทักษะที่ถูกต้องมากน้อยเพียงใด ผู้เรียนจะได้นำข้อมูลเหล่านั้นมาแก้ไขการกระทำของตนเองในครั้งต่อไป ความผิดพลาดที่เกิดขึ้นครั้งก่อนๆ จะถูกขจัดออกไปจนกระทั่ง การแสดงทักษะนั้นถูกต้องถึงเกณฑ์ หรือมาตรฐานที่วางไว้

2) ผลย้อนกลับทำหน้าที่เสริมแรง (Reinforcement) ในที่นี้หมายถึงการที่ผลย้อนกลับกระตุ้นให้ผู้เรียนรักษาการแสดงทักษะที่ถูกต้องคงเส้นคงวาอยู่เสมอ การที่ครูบอกผู้เรียนว่าถูกต้องแล้ว ดีแล้วหรือการได้คะแนนเต็มจะช่วยให้ผู้เรียนตั้งใจที่จะ แสดงทักษะให้ถูกต้องสม่ำเสมอ เป็นผลให้การเรียนรู้ทักษะนั้นมีการเปลี่ยนแปลงค่อนข้างถาวร การ แสดงทักษะก็จะมีประสิทธิภาพมากขึ้น พร้อมทั้งจะเรียนรู้ในระดับสูงต่อไป

3) ผลย้อนกลับทำหน้าที่เป็นแรงจูงใจ (Motivation) จะกระตุ้นให้ผู้เรียนมีความพยายามมากขึ้น ขยันฝึกซ้อมมากขึ้น จนกระทั่งมีทักษะเป็นไป ตามเกณฑ์ที่กำหนดไว้ การที่ครูแจ้งให้ผู้เรียนทราบว่าได้ 8 คะแนน จาก 10 คะแนน จะเป็นข้อมูล บอกว่าตนเองอยู่ห่างจากจุดหมายปลายทางเพียง 2 คะแนน เท่านั้น ข้อมูลนี้จะช่วยกระตุ้นให้ผู้เรียน พยายามฝึกหัดมากขึ้นเพื่อที่จะได้บรรลุจุดหมายปลายทางที่วางไว้

Race (2005, อ้างถึงใน ฮัสมินี ยูโซะ, 2555, หน้า 16) กล่าวไว้ว่า บทบาทการให้ข้อมูลย้อนกลับทำให้ผู้เรียนนำไปใช้ประโยชน์ในการเรียนรู้มี 4 ประการ ได้แก่

1) การให้ข้อมูลย้อนกลับเป็นตัวช่วยให้ผู้เรียนได้บรรลุถึงงานที่ได้รับมอบหมายเพราะว่า สิ่งที่ผู้เรียนทำหรือแสดงออกมานั้นอาจจะยังไม่ถูกต้องหรือยังมีข้อ ผิดพลาดซึ่งอาจเป็นเพราะผู้เรียนยังไม่เข้าใจในส่วนของเนื้อหาหรือคำถามนั้นๆ

2) การให้ข้อมูลย้อนกลับเป็นการแสดงถึงความจำเป็นในการเรียนรู้ โดยที่ผู้เรียนจะต้องมีความพยายามในการเฝ้หาความรู้ เพื่อจะได้ประสบความสำเร็จ และผู้สอนเองก็ต้องชี้แจงและเน้นถึงความสำคัญของแต่ละเนื้อหาให้ ผู้เรียนได้ทราบถึงเหตุผลที่จำเป็นต้องเรียน

3) การให้ข้อมูลย้อนกลับเป็นการเพิ่มความกระตือรือร้น ของผู้เรียนในการเรียนรู้และให้ผู้เรียนรู้สึกต้องการที่จะเรียนรู้และขวนขวายมากขึ้น และผู้สอนเองก็ควรทำให้ผู้เรียนสนใจในทุกๆ ประเด็นเนื้อหา

4) การให้ข้อมูลย้อนกลับเป็นตัวสนับสนุนให้ผู้เรียนได้พัฒนาการเรียนภาษาในทุกๆ ทักษะจากการแสดงความคิดเห็นและแก้ไขข้อผิดพลาดที่ผู้สอนได้จัดให้ผู้เรียนจากงานที่ผู้เรียนได้ทำ ผู้เรียนจำเป็นต้องได้รับแรงจูงใจจากผู้สอนในการพัฒนาการเรียนรู้อาษาด้วย

2. ประเภทของการให้ข้อมูลย้อนกลับ

มีนักการศึกษาหลายท่านได้สรุปประเภทของการให้ข้อมูลย้อนกลับไว้หลายรูปแบบ ได้แก่ แบ่งประเภทตามลักษณะสาระเนื้อหา แบ่งประเภทโดยใช้เกณฑ์ และแบ่งประเภทตามรูปแบบ ดังนี้

2.1 แบ่งประเภทตามลักษณะสาระเนื้อหาของข้อมูลย้อนกลับที่แสดงให้ ผู้เรียนทราบหลังจากตอบคำถามในบทเรียน แบ่งเป็น 3 ประเภท ดังนี้

1) ข้อมูลย้อนกลับแบบจูงใจ หรือเสริมแรง (motivational feedback) เป็นการให้ข้อมูลย้อนกลับในรูปของการชมเชย หรือการให้ รางวัล เพื่อเสริมแรงให้ผู้เรียนเกิดความคงทนในการจำและสร้างกำลังใจในการเรียนในลำดับต่อไป การให้ข้อมูลย้อนกลับในลักษณะนี้ไม่มีการลงโทษต่อคำตอบที่ผิด

2) ข้อมูลย้อนกลับเชิงข้อมูล (informational feedback) เป็นการให้ข้อมูลย้อนกลับในรูปข้อความ ซึ่งอาจเป็นการแนะนำแนวทางใน การตอบคำถามให้ถูกต้อง หรืออธิบายว่า คำตอบที่ถูกต้องนั้นถูกต้องอย่างไร หรือคำตอบที่ผิดนั้นผิด อย่างไร การใช้รูปแบบการให้ข้อมูลย้อนกลับแบบนี้ใช้ได้ทั้งภายหลังการตอบถูกและหลังการตอบ ผิด การให้ข้อมูลย้อนกลับเชิงข้อมูลนี้ แบ่งตามลักษณะข้อมูลได้ 2 ลักษณะ ดังนี้ ลักษณะแรก คือ ให้ข้อความโดยบอกว่า ถูก / ไม่ถูก (Correct / incorrect message) เป็นการบอกว่าถูกหรือไม่ถูกโดย ไม่มีรายละเอียดอื่น ๆ ลักษณะถัดมา คือ การแสดงคำตอบที่ถูกต้อง (presentation of correct answers) เป็นการให้ข้อมูลย้อนกลับที่บอกคำตอบที่ถูกต้องเท่านั้น

3) การให้ข้อมูลย้อนกลับแบบอธิบาย รายละเอียด (explanatory feedback) เป็นการอธิบายรายละเอียดคำตอบของข้อที่ถูก และข้อที่ผิด แบ่งได้ 2 ลักษณะ คือ การให้ข้อมูลย้อนกลับแบบอธิบาย (explanative feedback) เป็นการอธิบาย คำตอบของข้อที่ถูก และข้อที่ผิด และการให้ข้อมูลย้อนกลับแบบชี้แนะ (directive feedback) เป็นการให้ข้อมูลย้อนกลับที่มีการแนะนำวิธีการแก้ปัญหา หรือวิธีการที่ได้คำตอบที่ถูกต้อง ซึ่งส่วนใหญ่จะใช้ข้อมูลย้อนกลับชนิดนี้หลังการตอบที่ผิดในครั้งแรก เพื่อให้มีความพยายามตอบให้ถูกในครั้งต่อไป (Kulhavy and Stock, 1989; Demsey and Wager, 1988; สุกัญญา นิมานันท์, 2533, อ้างถึงใน ฮัซมีนี ยูโซะ, 2555, หน้า 17)

2.2 แบ่งประเภทของการให้ข้อมูลย้อนกลับโดยใช้เกณฑ์ในการแบ่ง (De Cecco, 1968: 255, อ้างถึงใน จรรยา ศิรินิล, 2559, หน้า 28-29) ดังนี้

1) แบ่งตามเวลาที่ให้ข้อมูลย้อนกลับ ประกอบด้วย

1.1 การให้ข้อมูลย้อนกลับแบบทันที (immediate informative feedback) ทำได้โดยให้ นักเรียนรู้ผลการกระทำอย่างทันทีทันใดว่าการกระทำนั้นถูกหรือผิด เหมาะสมหรือไม่ เนื่องจากถ้า ต้องการให้การเสริมแรงเป็นไปอย่างมีประสิทธิภาพแล้วจะต้องจัดสิ่งเร้าที่เป็นตัวเสริมแรงให้ในเวลาใกล้เคียงกับการตอบสนองที่ได้แสดงออกไปแล้ว

1.2 การให้ข้อมูลย้อนกลับแบบล่าช้า (delayed informative feedback) เป็นการให้ผู้เรียนรู้ผลการกระทำหลังจากตอบสนองแล้ว และเว้นระยะไว้ช่วงหนึ่ง

2) แบ่งตามวิธีการที่ให้ข้อมูลย้อนกลับ ประกอบด้วยการให้ข้อมูลย้อนกลับที่เป็นคำพูด รูปภาพ เสียงในเทปบันทึกเสียง และการให้ข้อมูลย้อนกลับที่เป็นคำเฉลย

2.3 แบ่งประเภทของข้อมูลย้อนกลับตามรูปแบบ (Shute, 2008: 1-2) มี 2 รูปแบบ คือ

1) ข้อมูลย้อนกลับเพื่อความก้าวหน้า (Formative Feedback) หมายถึง การให้ข้อเสนอแนะเพื่อส่งเสริมสนับสนุนความสามารถของผู้เรียนที่ไม่ใช่การประเมินเพื่อมุ่งตัดสินวัตถุประสงค์หลัก คือ การส่งเสริมการเรียนรู้หรือความสามารถโดยให้สารสนเทศที่ถูกต้องตรงกับเป้าหมายและทักษะและเชื่อมโยงกับการวินิจฉัยและบุคลิกภาพของผู้เรียน

2) ข้อมูลย้อนกลับแบบสรุปผลรวม (Summative Feedback) คือข้อมูลย้อนกลับที่ผู้เรียนได้รับหลังจากการสอบแล้ว

2.4 การให้ข้อมูลย้อนกลับแบบประเมินผล (evaluative feedback) เป็นการตัดสินผลของการกระทำนั้นๆ (Tunstall and Gipps, 1996, อ้างถึงใน ฮัซมีนี ยูโซะ, 2555, หน้า 17) ดังนี้

1) การให้ข้อมูลย้อนกลับทางบวก (positive feedback) นั้นหมายถึงการให้ผู้เรียนได้ทราบผลการตอบสนองของตน หรือสะท้อนให้เห็นส่วนดี ซึ่งจะไม่ก่อให้เกิดปัญหากับผู้เรียน (พรรณี ช.เจนจิต, 2545, 80, อ้างถึงใน จรรยา ศิรินิล, 2559, หน้า 28) การให้ข้อมูลย้อนกลับในเชิงบวก มีความแตกต่างการพูดเยินยอ (Praise) หรือ คำชม เนื่องจากการเยินยอเป็นการพูดชื่นชมแต่ไม่ได้บอกถึงจุดดีหรือพฤติกรรมที่กระทำว่าดีอย่างไร แต่การให้ข้อมูลย้อนกลับในเชิงบวก คือ การบรรยาย/อธิบายถึงพฤติกรรมต่างๆ ของผู้ที่ได้รับข้อมูลย้อนกลับว่ามีพฤติกรรมที่ดีอย่างไร สิ่งใดเป็นการกระทำที่ถูกต้องเหมาะสม เพื่อให้ผู้ที่ได้รับข้อมูล ย้อนกลับเกิดแรงจูงใจที่จะปฏิบัติพฤติกรรมนั้นอย่างต่อเนื่อง และควรที่จะให้อย่างต่อเนื่องและชัดเจน แต่ถ้าหากผู้สอนไม่ให้ข้อมูลย้อนกลับในเชิงบวกแก่ผู้เรียนอาจส่งผลให้พฤติกรรมนั้นๆ ไม่ได้รับการปฏิบัติอย่างต่อเนื่อง เพราะไม่ทราบว่าเป็นพฤติกรรมที่ถูกต้องเหมาะสม หรือท้อแท้ หหมดกำลังใจในการปฏิบัติเนื่องจากทำแล้วไม่มีใครรับรู้ไม่มีใครเห็น หรือไม่แน่ใจว่าพฤติกรรมนั้นๆ สมควรกระทำหรือไม่

2) การให้ข้อมูลย้อนกลับทางลบ (negative feedback) เป็นการให้ผู้เรียนได้ทราบ ผลการตอบสนองของตนหรือสะท้อนให้เห็นจุดบกพร่องซึ่งเป็นปัญหาของผู้เรียน (พรรณี ช.เจนจิต, 2545 , หน้า 80, อ้างถึงใน จรรยา ศิรินิล, 2559, หน้า 28) เป็นการอธิบายพฤติกรรมของตัวบุคคลที่จำเพาะ เจาะจงว่าผู้เรียนทำเรื่องนั้นๆ ที่ไม่ได้ตามที่คาดหวังหรือตามที่ควรจะเป็น ว่ามีความไม่เหมาะสม อย่างไร โดยกล่าวเฉพาะพฤติกรรมเรื่องนั้นๆ เพียงอย่างเดียว มิใช่ตัวตนของผู้เรียน พร้อมทั้งบอก แนวทางที่จำเพาะในการปรับปรุงแก้ไข เพื่อก่อให้เกิดการปรับปรุงในทางที่ดีขึ้น ไม่ปฏิบัติพฤติกรรม นั้น ซึ่งการให้ข้อมูลย้อนกลับในเชิงลบ แตกต่างจากการตำหนิหรือดุ ซึ่งตำหนิ (Criticism) เป็นการพูดที่ตำหนิตัวบุคคล โดยที่ไม่บอกทางออก แก่ผู้เรียนว่าจะต้องไปปรับปรุงแก้ไขอย่างไร ทำให้ผู้เรียนสูญเสียความมั่นใจในตนเอง นอกจากนี้การละเลยการให้ข้อมูลย้อนกลับในเชิงลบ จะทำให้พฤติกรรมที่ไม่ถูกต้องคงอยู่ต่อไป ผู้เรียนจะไม่ทราบ ถึงข้อบกพร่อง และเกิดความเข้าใจผิดว่า

สิ่งที่ปฏิบัติคือสิ่งที่ถูกต้องอีก (จกมลวรรณ มุสิกทอง และ พิจิตรา เล็กดำรงกุล, 2556, อ้างถึงใน จุฑาภรณ์ มาสันเทียะ, 2560, หน้า 32-33)

2.5 รูปแบบการให้ข้อมูลย้อนกลับที่จำแนกตามแหล่งที่มาของข้อมูลย้อนกลับ

จากการสังเคราะห์งานวิจัยที่เกี่ยวข้องกับรูปแบบของการให้ข้อมูลย้อนกลับที่จำแนกตามแหล่งที่มาของข้อมูลย้อนกลับ สามารถจำแนกออกเป็น 4 ประเภท (ไพฑูริย์ เอื้อบุญประดิษฐ์, 2563, หน้า 74-75) ดังนี้

1) การให้ข้อมูลย้อนกลับโดยผู้สอน ผู้สอนมักจะเป็นผู้ให้ข้อมูลย้อนกลับส่วนใหญ่แก่ผู้เรียน ตัวอย่างเช่นในงานวิจัย ที่มีการให้ข้อมูลย้อนกลับของผู้สอนเกี่ยวกับความถูกต้องของภาษา Form-Focused Corrective Feedback (FFCF) ซึ่งมี 2 รูปแบบ คือ Direct Error Correction คือ การแก้ไขคำตอบให้แก่ผู้เรียน และ Metalinguistic Feedback คือ การให้คำอธิบายเกี่ยวกับไวยากรณ์ทางภาษาแล้วให้ผู้เรียน แก้ไขคำตอบเอง โดยรูปแบบการให้คำอธิบายเกี่ยวกับไวยากรณ์ทางภาษาที่แตกต่างกันนั้น ส่งผลให้มีระดับความสามารถทางภาษาที่แตกต่างกันด้วย โดยจากการสัมภาษณ์นักเรียนส่วนใหญ่ คิดเป็น 80% พบว่า ต้องการได้รับการให้ข้อมูลย้อนกลับจากผู้สอน (Diab, 2015, อ้างถึงใน ไพฑูริย์ เอื้อบุญประดิษฐ์, 2563)

2) การให้ข้อมูลย้อนกลับโดยเพื่อน การให้ข้อมูลย้อนกลับโดยเพื่อน คือ การใช้แหล่งข้อมูลของผู้เรียนและการปฏิสัมพันธ์ กับแหล่งต่างๆ ที่ได้เรียนเพื่อแสดงความคิดเห็นต่อกระบวนการเขียนงาน ซึ่งการให้ข้อมูลย้อนกลับโดยเพื่อนนั้นจัดว่าเป็นความสัมพันธ์ทางสังคม การพัฒนาทางเขาวนปัญญา ตลอดจนเป็นการก่อให้เกิด ความก้าวหน้าทางความคิดสร้างสรรค์ (Gülsah Çinar Yastibas, 2015) การให้ข้อมูลย้อนกลับโดยเพื่อนสำหรับการประเมินความก้าวหน้าในการเรียน เป็นวิธีการที่เหมาะสมกับนักเรียนในระดับมัธยมศึกษาและอุดมศึกษา โดยในปัจจุบันได้มีการนำระบบออนไลน์ มาประยุกต์ใช้ร่วมกับการให้ข้อมูลย้อนกลับโดยเพื่อน ทำให้เป็นวิธีการที่มีประสิทธิภาพมากยิ่งขึ้น โดยเฉพาะอย่างยิ่งในบริบทการประเมินงานเขียนผ่านระบบออนไลน์โดยเพื่อน โดยมีการศึกษา ถึงระดับของรายละเอียดของข้อมูลที่เพื่อนได้ให้ข้อมูลย้อนกลับซึ่งแตกต่างกัน โดยการให้ข้อมูล แบบอธิบายรายละเอียดนั้นดีที่สุด (Mario Gielen, 2015) นอกจากนี้ การให้ข้อมูลย้อนกลับโดยเพื่อน แก่ผู้เรียนในการประเมินงานเขียนช่วยให้ผู้เรียนมีความมั่นใจมากขึ้น ส่งผลให้ผู้เรียนมีระดับ ความวิตกกังวลในงานเขียนลดลง (Gülsah Çinar Yastibas, 2015, อ้างถึงใน ไพฑูริย์ เอื้อบุญประดิษฐ์, 2563)

3) การให้ข้อมูลย้อนกลับโดยตนเอง การให้ข้อมูลย้อนกลับโดยตนเอง คือ การแก้ไขข้อบกพร่องหรือจุดผิดพลาดในงานเขียน ด้วยตนเอง โดยจากการศึกษาวิจัยการให้ข้อมูลย้อนกลับในการเขียนโดยผู้สอนและโดยตนเองนั้น พบว่ากรณีที่ผู้เรียนแก้ไขงานเขียนด้วยตนเองนั้น ผู้เรียนมีความคาดหวังเพื่อที่จะให้ได้คะแนนสูงขึ้น แต่เมื่อเปรียบเทียบผลการให้ข้อมูลย้อนกลับโดยครู

และโดยตนเองนั้นจะพบว่า การให้ข้อมูลย้อนกลับ โดยตนเองจะมีข้อผิดพลาดมากกว่าข้อมูลย้อนกลับที่ได้รับจากครูผู้สอน (Diab, 2015, อ้างถึงใน ไพฑูรย์ เอื้อบุญประดิษฐ์, 2563)

4) การให้ข้อมูลย้อนกลับโดยระบบออนไลน์หรือระบบอัตโนมัติ การให้ข้อมูลย้อนกลับโดยระบบออนไลน์หรือระบบอัตโนมัติ คือ การให้ข้อมูลย้อนกลับ ที่ให้แก่ผู้เรียนโดยโปรแกรมคอมพิวเตอร์ภายหลังจากทำแบบฝึกหัดหรือคำถาม ซึ่งในปัจจุบัน ได้มีการนำการให้ข้อมูลย้อนกลับโดยโปรแกรมคอมพิวเตอร์มาประยุกต์ใช้ในการเรียนรู้ อย่างกว้างขวาง โดยเฉพาะอย่างยิ่งในสถานการณ์ที่มีการให้การบ้านหรือแบบฝึกหัดออนไลน์ ที่ผู้เรียนสามารถคัดลอกคำตอบจากระบบอินเทอร์เน็ตมาตอบได้ จึงมีการสร้างโปรแกรม เพื่อตรวจจับการคัดลอกคำตอบจากระบบอินเทอร์เน็ตมาใช้ เพื่อป้องกันการคัดลอกคำตอบ ทั้งนี้ เมื่อตรวจพบเปอร์เซ็นต์การคัดลอกเกินจากเกณฑ์ การให้ข้อมูลย้อนกลับจะเกิดขึ้นทันที พร้อมทั้ง แจ้งถึงส่วนที่มีการคัดลอกคำตอบ โดยผลการวิจัยพบว่า ภายหลังจากการใช้โปรแกรมการให้ข้อมูลย้อนกลับ แบบอัตโนมัติเพื่อตรวจจับการคัดลอกคำตอบแล้ว นักเรียนมีพฤติกรรมการคัดลอกคำตอบ จากอินเทอร์เน็ตมาตอบในข้อสอบหรือแบบฝึกหัดลดน้อยลง (Akçapinar, 2015, อ้างถึงใน ไพฑูรย์ เอื้อบุญประดิษฐ์, 2563)

จากที่กล่าวมาสรุปได้ว่าการให้ข้อมูลย้อนกลับเป็นการแจ้งสารสนเทศที่เกี่ยวข้องกับผลการเรียนหรือผลปฏิบัติงานซึ่งเป็นส่วนหนึ่งในกระบวนการประเมินความก้าวหน้าในการเรียนของผู้เรียนโดยมีวัตถุประสงค์เพื่อกระตุ้นเสริมแรงและสร้างแรงจูงใจในการเรียนแก่ผู้เรียนเกิดการเรียนรู้พัฒนา และเปลี่ยนแปลงพฤติกรรมโดยในการศึกษาค้นคว้าครั้งนี้เป็นการให้สารสนเทศที่นำไปใช้ในการแก้ไขจุดอ่อนหรือข้อผิดพลาดที่เกี่ยวกับความสามารถด้านการอ่านเพื่อความเข้าใจภาษาอังกฤษ อีกทั้งครูผู้สอนยังสามารถนำไปใช้ในออกแบบกิจกรรมการเรียนรู้เพื่อพัฒนาผู้เรียนให้บรรลุผลสำเร็จทางการเรียนได้อีกด้วย หลักการให้ข้อมูลย้อนกลับคือการให้สารสนเทศในเรื่องใดเรื่องหนึ่งโดยเฉพาะเจาะจง ซึ่งในการศึกษาค้นคว้าครั้งนี้เป็นสารสนเทศย้อนกลับเกี่ยวกับความสามารถการอ่านเพื่อความเข้าใจของผู้เรียนแต่ละบุคคล ใช้ภาษาเข้าใจง่าย สั้นได้ใจความ มีความชัดเจนนำเสนอด้วยถ้อยคำในเชิงบวกระบุพฤติกรรมที่ผู้เรียนสามารถนำไปปฏิบัติได้จริง มีการกำหนดทิศทางให้ผู้เรียนสามารถ ไปถึงเป้าหมายได้ด้วยตัวผู้เรียนเอง และเป็นการให้ข้อมูลย้อนกลับแบบทันทีทันใด เพื่อให้ผู้เรียนเกิดแรงจูงใจในการเรียนและไม่ล้าซึมประเด็น โดยมีการบันทึกข้อมูลย้อนกลับไว้อย่างชัดเจนทุกครั้ง

ดังนั้นในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยกำหนดประเภทข้อมูลย้อนกลับตามรูปแบบ (Shute, 2008: 1-2) ที่เป็นการให้ข้อมูลย้อนกลับเพื่อความก้าวหน้า (Formative Feedback) โดยให้ข้อเสนอแนะเพื่อส่งเสริมสนับสนุนความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของผู้เรียนมีจุดมุ่งหมายในการส่งเสริมการเรียนรู้หรือความสามารถโดยให้สารสนเทศที่ถูกต้องตรงกับคุณลักษณะหรือทักษะที่ทำการวินิจฉัย และให้สารสนเทศย้อนกลับแบบสรุปผลรวม (Summative Feedback) โดยเป็นการให้ผลการวินิจฉัยหลังจากทำแบบทดสอบแล้ว ผู้วิจัยใช้หลักการให้ข้อมูลย้อนกลับที่มีการ

อธิบายประกอบเนื้อหาซึ่งมีผลต่อการเรียนรู้ของผู้เรียนมากทำให้ผู้เรียนสามารถทำความเข้าใจได้ดีกว่า และส่งผลต่อการพัฒนาการเรียนรู้ของผู้เรียนได้ดีกว่าการให้ข้อมูลย้อนกลับ ที่เพียงแต่บอกว่าถูกหรือผิดโดยไม่มีคำอธิบายใด ๆ และผู้ให้ข้อมูลป้อนกลับนำเสนอแนวทางที่ผู้เรียนสามารถยอมรับและนำไปปฏิบัติได้ (Krumboltz and Bouawitz, 1962, อ้างถึงใน ทิพนีพา อรุณวิภาส, 2552, 11; De Cecco, 1968, pp. 284-294)

3. แนวทางการให้ข้อมูลย้อนกลับ

ข้อมูลย้อนกลับที่ให้นั้นต้องมีความชัดเจนเฉพาะเจาะจง มีความตรงไปตรงมา ไม่ยาวจนเกินไป และเสนอแนะถึงความถูกต้องของสิ่งที่ต้องทำสามารถนำไปปฏิบัติได้จริง และเป็นข้อมูลที่แท้จริงของพฤติกรรมที่เกิดขึ้นมีการกำหนดทิศทางให้ผู้เรียนสามารถ ไปถึงเป้าหมายได้ หลีกเลี่ยงการให้ข้อมูลที่แสดงอารมณ์ ให้สารสนเทศในเชิงบวกและมุ่งเน้นการพัฒนาความสามารถที่เปรียบเทียบเฉพาะในตัวบุคคลไม่เปรียบเทียบกับบุคคลอื่นๆ โดยผู้ให้ข้อมูลย้อนกลับต้องเป็นที่เชื่อถือ และได้รับการยอมรับในแง่ของการมีความตั้งใจในการให้ข้อมูล ในส่วนของผู้รับข้อมูลย้อนกลับจะต้องยืนยันเกี่ยวกับการเปลี่ยนแปลงที่เกิดขึ้นสำหรับตนเองได้อย่างชัดเจน ข้อมูลย้อนกลับต้องทำให้เกิดความวิตกกังวลด้วย เพื่อเป็นการจูงใจไปสู่การเปลี่ยนแปลง เพราะหากผู้ได้รับข้อมูลย้อนกลับได้รับข้อมูลที่ตรงกับความคิดเห็นของตนเองเสมอไปแล้ว การเปลี่ยนแปลงพฤติกรรมก็จะไม่เกิดขึ้น ผู้ให้ข้อมูลย้อนกลับต้องทำให้เห็นว่าข้อมูลย้อนกลับเป็นสิ่งสำคัญและจำเป็นจะต้องยอมรับเพื่อนำไปสู่การปรับปรุงไม่ว่าผลย้อนกลับจะเป็นทางบวกหรือทางลบ ต้องมีรูปแบบที่ชัดเจนสำหรับการเปลี่ยนแปลง หลังจากได้รับข้อมูลย้อนกลับ (Tuckman, 1976 อ้างถึงใน ฮัซมีนี ยูโซะ, 2555, หน้า 20-21; สมโภชน์ เอี่ยมสุภชาติ, 2549 หน้า 202-203, อ้างถึงใน จรรยา ศิรินิล 2559, หน้า 26)

คลิฟฟอร์ด (Clifford, 1981) ได้สรุปว่าหลักการให้ข้อมูลย้อนกลับในลักษณะจำเพาะเจาะจง (specific feedback) จะทำให้ผู้เรียนปรับปรุงการเรียนรู้ได้ดีกว่าการให้ข้อมูลย้อนกลับในลักษณะการให้ข้อมูลย้อนกลับในลักษณะรายละเอียด (detailed feedback) โดยมีการบันทึกข้อมูลย้อนกลับไว้อย่างชัดเจนทุกครั้ง ทรุสคอตต์ (Truscott, 2007) ให้ความเห็นว่า ในการแก้ไขจุดอ่อนทางภาษานั้นผู้เรียนควรได้รับการสนับสนุนส่งเสริมจากครูผู้สอนให้เกิดการปรับปรุงพัฒนาแก้ไขจุดอ่อนทางภาษาด้วยตัวผู้เรียนเอง สำหรับระยะเวลาที่เหมาะสมในการให้ข้อมูลย้อนกลับมี 2 แบบคือ ในช่วงแรกของการเรียนรู้ในสิ่งที่มีความซับซ้อนมากซึ่งเป็นกิจกรรมที่ใหม่และยาก การให้ข้อมูลย้อนกลับควรให้แบบทันทีทันใด แต่หากต้องการส่งเสริมกระบวนการเรียนรู้ ในเรื่องที่เรียนรู้แล้วโดยสามารถลงมือปฏิบัติงานด้วยความรู้ที่มีควรให้ข้อมูลย้อนกลับแบบภายหลัง ซึ่งทั้งสองแบบมีข้อดีต่างกัน คือ ข้อมูลย้อนกลับแบบทันทีทันใด (Immediately Feedback) ก่อให้เกิดแรงจูงใจในการทำงานและไม่ลืมประเด็น ในส่วนของข้อมูล ย้อนกลับแบบย้อนหลัง (Delay Feedback) นักเรียนได้รับเมื่อเวลา

ผ่านไปหลังจากทำหน้าที่เสร็จแล้วจะช่วยลดเหตุการณ์ที่ไม่พึงประสงค์ (Clifford, 1981; Shute, 2008: 15, อ้างถึงใน จรรยา ศิริสิน 2559, หน้า 26)

4. การพัฒนาโปรไฟล์สารสนเทศจากการวินิจฉัย

Skill profile คือ รูปแบบของข้อมูลสารสนเทศที่ระบุความหมายหรือคำจำกัดความของความรู้และความสามารถที่จำเป็นสำหรับสิ่งที่มุ่งศึกษาได้อย่างแม่นยำ นอกจากนี้ยังเป็นเครื่องมือวินิจฉัยและวิธีตรวจสอบการเรียนรู้ที่มีประสิทธิภาพ การสร้าง Skill profile จะทำการจัดกลุ่มคุณลักษณะหรือทักษะประเภทเดียวกันหรือเกี่ยวข้องกัน นอกจากนี้ระดับความรู้ความเข้าใจ ความถูกต้อง และความชัดเจนของคำอธิบายของแต่ละทักษะมีความจำเป็นในการประเมินที่ถูกระบุโดยผู้เชี่ยวชาญ Jang (2009, p.39) ระบุความสำคัญของ Skill profile ของการวินิจฉัยความสามารถในการอ่านเพื่อความเข้าใจ 2 ประการเกี่ยวกับลักษณะพิเศษของ skill profiles คือการรายงานข้อมูลความสามารถ และการใช้ประโยชน์จากข้อมูลที่ได้จากการวินิจฉัยสำหรับนักเรียนและครู สอดคล้องกับ Alderson (2007) ซึ่งเสนอว่า หัวใจสำคัญในการวินิจฉัยต้องเป็นการให้ข้อเสนอแนะที่สามารถนำไปใช้งานได้จริงทั้งแก่ผู้เรียนเองหรือผู้วินิจฉัย

4.1 ประโยชน์ของ Skill Profile

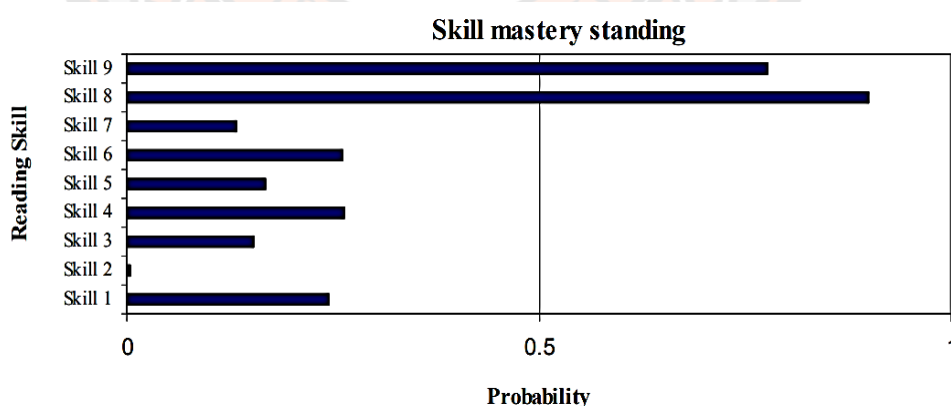
Skill profile ที่ถูกสร้างมาอย่างดีเมื่อตั้งค่าแล้วช่วยในการตรวจสอบเพื่อให้เกิดความแน่ใจว่าการฝึกอบรมหรือการสอนสอดคล้องกับวัตถุประสงค์ที่แท้จริงได้อย่างถูกต้อง ทักษะคุณลักษณะใดบ้างที่มีความรอบรู้และคุณลักษณะใดบ้างที่ไม่รอบรู้ เพื่อสร้างแนวทางที่เหมาะสมในการเสริมสร้างให้ผู้เรียนมีคุณลักษณะที่จำเป็นในความสามารถนั้น ๆ ได้อย่างแม่นยำ Skill profile ไม่ใช่คำอธิบายผลงาน แต่เป็นคลังความรู้และทักษะที่ผู้เรียนจำเป็นต้องมีเพื่อปฏิบัติงานในขอบเขตเนื้อหาที่กำหนด Skill profile สามารถระบุความหมายหรือคำจำกัดความของความรู้และความสามารถที่จำเป็นสำหรับสิ่งที่มุ่งศึกษาได้อย่างแม่นยำ นอกจากนี้ยังเป็นเครื่องมือวินิจฉัยและวิธีตรวจสอบการเรียนรู้ที่มีประสิทธิภาพ การสร้าง Skill profile จะทำการจัดกลุ่มคุณลักษณะหรือทักษะประเภทเดียวกันหรือเกี่ยวข้องกัน นอกจากนี้ระดับความรู้ความเข้าใจ ความถูกต้อง และความชัดเจนของคำอธิบายของแต่ละทักษะมีความจำเป็นในการประเมินที่ถูกระบุโดยผู้เชี่ยวชาญ

4.2 ตัวอย่าง Skill Profile

สารสนเทศจากการวินิจฉัยถูกรายงานในรูปแบบแตกต่างกัน นักวิจัยหลายท่านได้พัฒนาบัตรรายงานขึ้น โดยมีชื่อเรียกแตกต่างกัน อาทิ DiagnOsis Scoring Reports (Jang, 2005, pp.155-157) Pattern Occurrence Probabilities (Ravand, 2013, p.24) A Diagnostic Score Report (Gierl, M.J. et al., (2008, p.40) Component Skills Profile (DELTA, 2012) Skill Profile (Li, 2015; เกศรินทร์ สมราช, 2022; วรางคณา ก้อมน้อย, 2022) Skill Mastery Profile

(Javidanmehr & Sarab, 2019) ในการศึกษาครั้งนี้ผู้วิจัยใช้คำ “Skill Profile” (Li, 2015; เกศรินทร์ สมราช, 2022; วราภรณ์ ก้อมน้อย, 2022)

Jang (2005) ได้พัฒนา DiagnOsis Scoring Reports เพื่อรายงานสารสนเทศจากการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจ ภาพ 25 แสดงตัวอย่างที่สมบูรณ์ของบัตรรายงาน DiagnOsis Scoring Reports (Jang, 2005, pp.155-157) บัตรรายงานถูกจัดทำขึ้นสำหรับ Margo หลังจากการทดสอบวินิจฉัยล่วงหน้า การวัดประกอบด้วยสองหน้า หน้าแรกให้ข้อมูลการวินิจฉัยเป็นรายบุคคลสำหรับ Margo ในสามองค์ประกอบ หน้าที 2 นำเสนอคำอธิบาย 9 ทักษะพร้อมพร้อมรายการทดสอบที่เกี่ยวข้องกับแต่ละทักษะ โดยสังเขปดัชนีข้อมูลการวินิจฉัย (Diagnostic Information Index: DII) เพื่อแจ้งนักเรียนและครูว่ารายการทดสอบแตกต่างกันตามอำนาจการวินิจฉัยและเพิ่มเติมเพื่อช่วยให้พวกเขาฝึกการทดสอบโดยใช้รายการที่มีข้อมูลในการวินิจฉัยมากขึ้น ธงบนด้านซ้ายมือระบุทักษะที่ต้องปรับปรุง เครื่องหมายคำถามระบุว่าความเชี่ยวชาญทักษะไม่ได้ถูกกำหนด แต่ละองค์ประกอบอธิบายไว้ด้านล่างองค์ประกอบแรกของ “ทบทวนคำตอบ” ดังรูปที่ 17.1 คือมีวัตถุประสงค์เพื่อให้โอกาสในการทบทวนคำตอบสำหรับคำถามและคำตอบที่ถูกต้องพร้อมกับข้อมูลการให้คะแนน มีตัวเลือกหลักหลายตัวที่ใช้ใน DiagnOsis นอกจากนี้ ระบุความยากของข้อคำถามสำหรับแต่ละข้อโดยใช้หนึ่งในสามหมวดหมู่: 'ง่าย', 'ปานกลาง' และ 'ยาก' ความยากของข้อสอบถูกกำหนดโดยพารามิเตอร์ความยากข้อสอบจาก BILOG โดยใช้ข้อมูลการทดสอบภาคสนาม LanguEdge รายการ 'ง่าย'หมวดหมู่รวมทั้งหมด 12 รายการซึ่งมีการประมาณค่าพารามิเตอร์ b อยู่ในช่วงต่ำสุด -2.7 ถึงสูงสุด -0.53



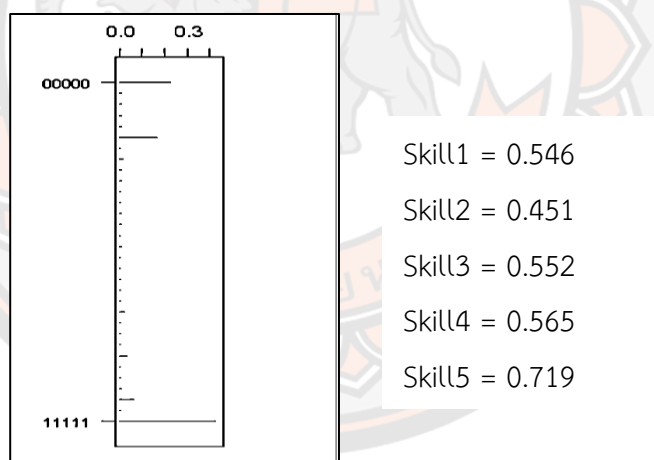
ภาพ 32 แสดงบัตรรายงาน DiagnOsis Scoring Reports

(Jang, 2005, pp.158)

แถบแนวนอนแต่ละอันแสดงถึงการประมาณค่าความน่าจะเป็นของความรอบรู้สำหรับแต่ละทักษะพื้นที่สีเทาตรงกลางพล็อตแสดงถึงพื้นที่ที่ไม่แตกต่างระหว่างรอบรู้หรือไม่รอบรู้ ถ้าบาร์ตก

ลงมาบริเวณนี้ ไม่สามารถกำหนดความรอบรู้ของนักเรียนได้ ในตัวอย่างที่แสดงในรูป 29-1 ผู้สอบรอบรู้ทักษะ 8 และ 9 นั่นคือ SUM และ MCF แต่เขาจำเป็นต้องปรับปรุงทักษะอื่น ๆ ตามที่แถบ 'สั้น' ระบุ

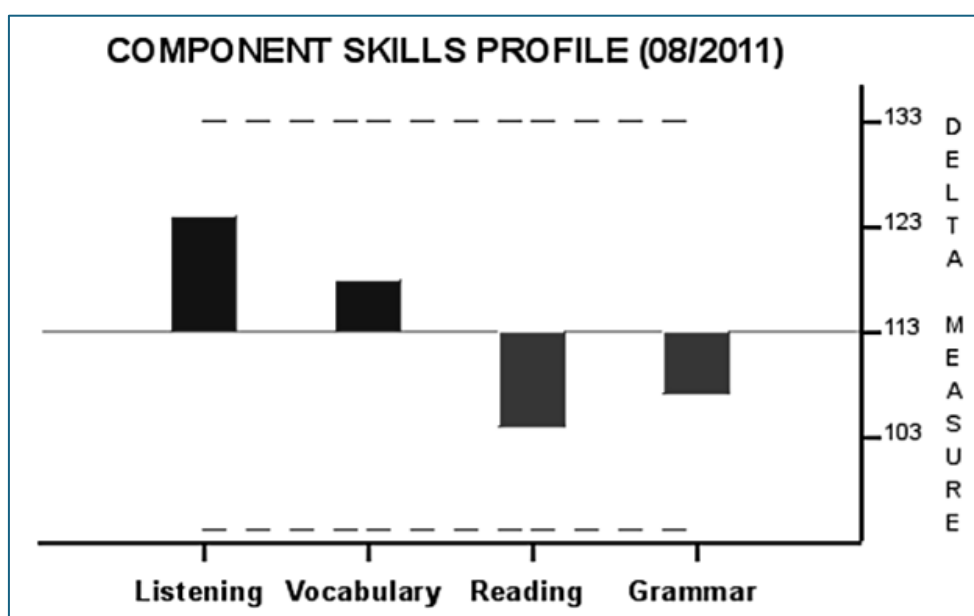
ราวันด์ (Ravand, 2013, p. 24) พัฒนาข้อมูลสารสนเทศการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษโดยเรียกว่า Pattern Occurrence Probabilities แสดงรูปแบบความรอบรู้สองรูปแบบที่มีความน่าจะเป็นสูงสุดคือ skill profile1, $\alpha_1 = [00000]$ และ skill profile 32, $\alpha_{32} = [11111]$ แสดงว่าในการศึกษาครั้งนี้ประมาณร้อยละ 42 ของผู้ตอบแบบสอบถามถูกจัดอยู่ในชั้นแฝง 32 ในฐานะเป็นผู้รอบรู้ของทักษะทั้งหมดและประมาณร้อยละ 22 ถูกจัดอยู่ในชั้นแฝง 1 ในฐานะผู้ไม่รอบรู้ ผู้ตอบแบบสอบถามจำนวนมากจัดอยู่ในประเภทที่มีโปรไฟล์ "Flat" กล่าวคือ ไม่รอบรู้คุณลักษณะใด (00000) หรือรอบรู้ทุกคุณลักษณะ (11111) อย่างไม่อย่างหนึ่ง แทนที่จะเป็นเฉพาะการรวมกันของคุณลักษณะ คำสั่งนี้ยังให้ความน่าจะเป็นของทักษะแสดงให้เห็นว่าผู้สอบมีความรอบรู้ในทักษะที่ 1 ประมาณร้อยละ 54 และผู้สอบมีความรอบรู้ในทักษะที่ 5 ร้อยละ 71 (Ravand, 2013, p. 24)



ภาพ 33 ตัวอย่าง Pattern Occurrence Probabilities การอ่านภาษาอังกฤษ
(Ravand, 2013, p.24)

ต่อมาในปี ค.ศ. 2012 DELTA ได้พัฒนา Component Skills Profile (DELTA, 2012, pp.74-75) สำหรับรายงานความสามารถทางภาษาอังกฤษ 4 ด้าน ได้แก่ การฟัง ความรู้คำศัพท์ การอ่าน และความรู้ทางไวยากรณ์ โดยตัวอย่างข้อคำถามทั้งหมดได้รับการปรับเทียบโดยใช้ข้อมูลการทดสอบจริง ความยากของข้อคำถามแต่ละข้อส่งผลอย่างถูกต้องต่อความสามารถโดยรวมของนักเรียน หรือมาตรฐาน DELTA ทักษะย่อยทางภาษาอังกฤษถูกวัดด้วยข้อคำถาม โดยข้อคำถามแต่ละข้อจะถูก

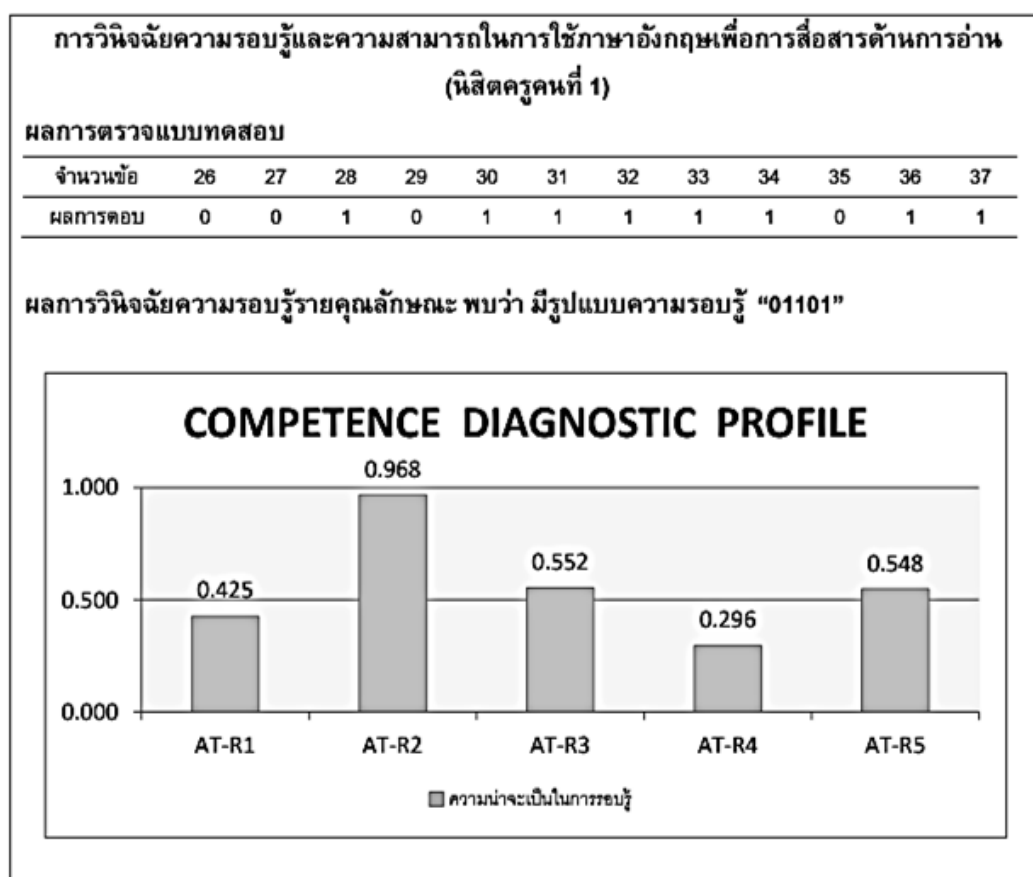
บันทึกไว้ในรายงานความสามารถของผู้สอบ DELTA คำชี้แจงเกี่ยวกับจุดแข็งและจุดอ่อนของนักเรียน ในทักษะย่อยของภาษานั้นขึ้นอยู่กับความยากของข้อคำถามที่สัมพันธ์กับของความรู้ กล่าวอีกนัยหนึ่งข้อคำถามที่เป็นของระดับความยากที่ต่ำกว่าความสามารถของนักเรียนคือระดับที่นักเรียนจะคาดว่าจะตอบถูก หากตอบไม่ถูกแสดงว่าจุดอ่อนที่เป็นไปได้ในทักษะย่อยนั้น เมื่อนักเรียนได้รับรายงานของพวกเขา พวกเขาขอแนะนำให้ใช้ผ่านการเรียนรู้อิสระหรือด้วยความช่วยเหลือของครูของพวกเขา ครูได้รับการบรรยายสรุปเกี่ยวกับวิธีการช่วยให้นักเรียนเข้าใจรายงานและให้คำแนะนำในการดำเนินการที่เหมาะสม ดังภาพ 29-3



ภาพ 34 Component Skills Profile ความสามารถภาษาอังกฤษ 4 ด้าน
(DELTA, 2012, pp. 74-75)

ทางฝั่งนักวัดผลในประเทศได้ทำการพัฒนา Skill profile เช่นกัน ในปี พ.ศ. 2563 เกศรินทร์ สมราช (2563) พัฒนา Skill profile สารสนเทศคุณลักษณะความสามารถภาษาอังกฤษเพื่อการสื่อสาร 4 ทักษะ การฟัง การพูด การอ่าน และการเขียน ในที่นี้ผู้วิจัยขอยกตัวอย่างเฉพาะการอ่าน คำอธิบายผลการวินิจฉัย ผลการวินิจฉัยพบว่า นิสิตครูคนที่ 1 มีความรอบรู้ในคุณลักษณะการอ่านที่ AT-R2, AT-R3 และ AT-R5 คือ มีความเข้าใจในการใช้ภาษาอังกฤษค้นหาข้อมูลเอกสารที่เกี่ยวข้อง กับการดำเนินชีวิต มีความเข้าใจในการอ่านภาษาอังกฤษได้เข้าใจ รู้เนื้อหาประเด็นสำคัญในบทความ ข่าวสารทางการศึกษา และมีความเข้าใจคำภาษาอังกฤษที่อธิบายเหตุการณ์ สถานการณ์ ความรู้สึก เจตจำนงและตอบกลับเป็นข้อความได้แต่ควรได้รับการพัฒนาในคุณลักษณะที่ AT-R1 และ AT-R4 คือ ควรฝึกอ่านภาษาอังกฤษ ค้นหาความหมาย ทำความเข้าใจในข้อความที่สนใจเกี่ยวข้องกับ

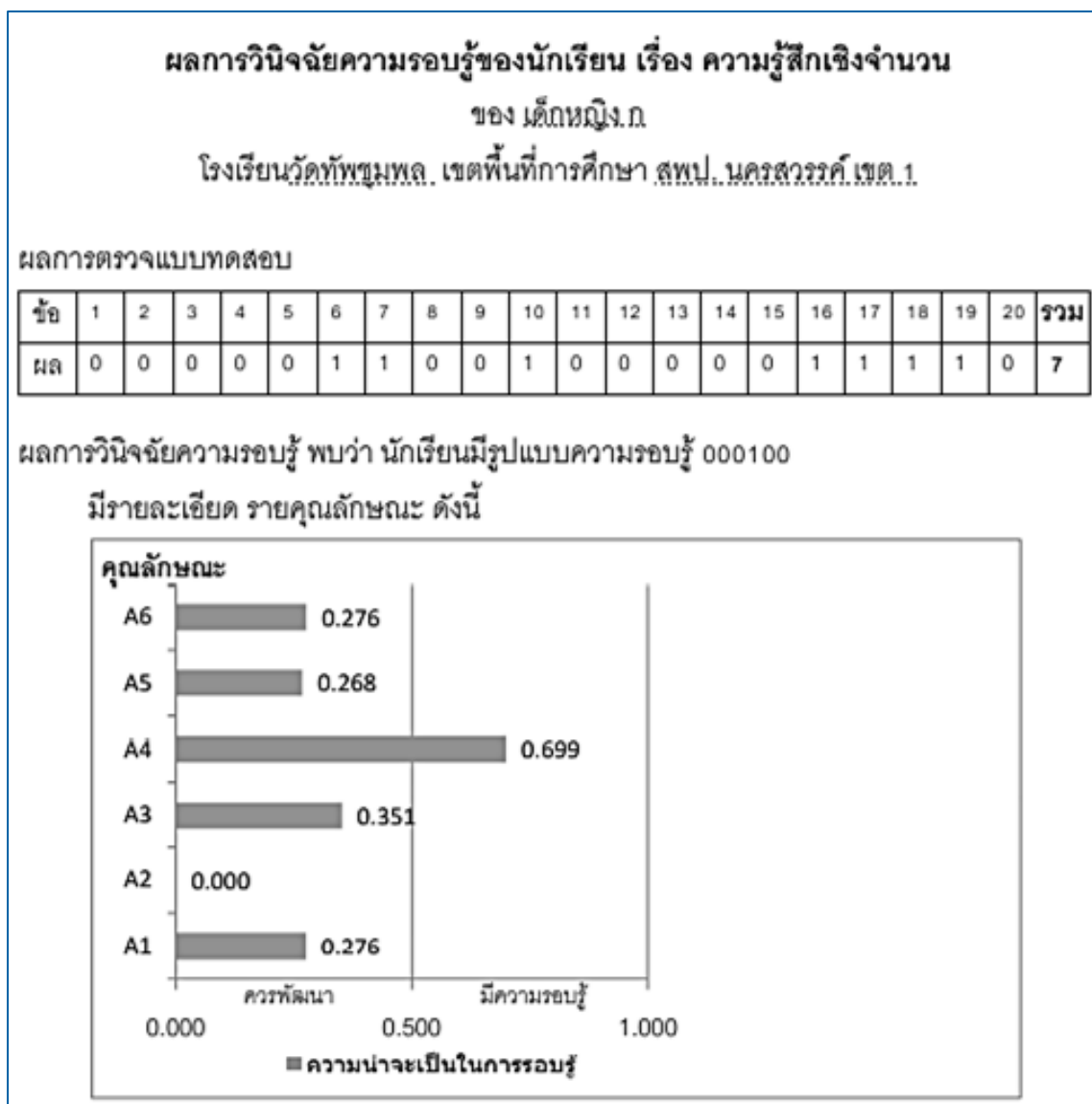
ชีวิตประจำวันและการทำงาน และฝึกอ่านภาษาอังกฤษค้นหาความหมาย ทำความเข้าใจเพื่อปฏิบัติ ตามคำแนะนำคำชี้แจงบนฉลากสินค้าที่เขียนเป็นภาษาอังกฤษได้ถูกต้องไม่ก่อให้เกิดอันตราย



ภาพ 35 Skills Profile ความสามารถภาษาอังกฤษเพื่อการสื่อสารด้านการอ่าน
(เกศรินทร์ สมราช, 2563)

นอกจากความสามารถทางภาษาอังกฤษ ผู้วิจัยได้นำตัวอย่าง Skill profile ความสามารถทางคณิตศาสตร์ อาทิจ วรางคณา ก้อมน้อย (2563) ได้พัฒนา Skill profile รายงานผลการวินิจฉัยความรู้สึกเชิงจำนวนของนักเรียนระดับชั้นประถมศึกษาปีที่ 3 วิชาคณิตศาสตร์เรื่องความรู้สึกเชิงจำนวน ประกอบด้วย 6 คุณลักษณะ โดยนำเสนอ เป็น 2 หัวข้อย่อย คือ ผลคะแนนของนักเรียน และผลการวินิจฉัยความรอบรู้ของนักเรียน

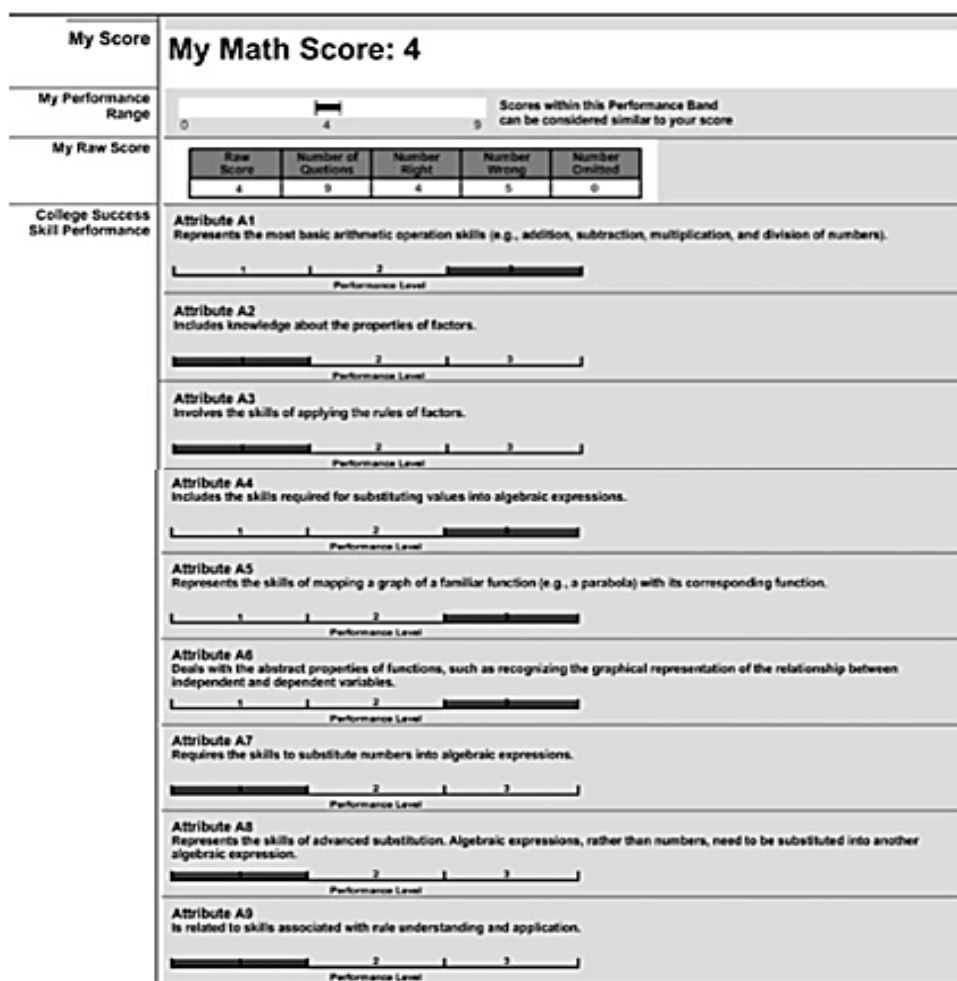
ตัวอย่าง Skill profile ความสามารถทางคณิตศาสตร์



ภาพ 36 ตัวอย่าง Skill profile วิชาคณิตศาสตร์เรื่องความรู้สึกเชิงจำนวน
 (วารงคณา ก้อมน้อย, 2563)

ผลการวินิจฉัยพบว่า เด็กหญิง ก มีความรอบรู้ในคุณลักษณะที่ 4 คือ นักเรียนมีความเข้าใจผลของการดำเนินการของจำนวน และสามารถใช้ประโยคสัญลักษณ์ที่สมมูลกัน เพื่อประโยชน์ในการหาคำตอบโดยไม่ต้องคำนวณได้แต่ควรได้รับการพัฒนาในคุณลักษณะอื่นที่เหลือ

การศึกษาของ Gierl, M.J. et al., (2008, p. 40) ได้พัฒนา A Diagnostic Score Report คุณลักษณะ 9 ข้อของความสามารถทางวิชาคณิตศาสตร์ ผู้สอบผู้ไม่มีความรอบรู้ในทักษะที่เกี่ยวข้องกับคุณลักษณะ 5 ประการ



ภาพ 37 ตัวอย่าง A Diagnostic Score Report
พัฒนาโดย Gierl, M.J. et al., (2008, p.40)

ผู้สอบมีความรอบรู้คุณลักษณะ A1,A4, A5, A6 ระดับประสิทธิภาพสามระดับที่ถูกเลือกสำหรับการรายงานคุณลักษณะความรอบรู้หรือไม่รอบรู้ (ค่าความน่าจะเป็นของความสามารถอยู่ระหว่าง 0.00 ถึง 0.35), ความรอบรู้บางส่วน (ค่าความน่าจะเป็นของความสามารถอยู่ระหว่าง 0.36 ถึง 0.70) และความรอบรู้ (ค่าความน่าจะเป็นของความสามารถอยู่ระหว่าง 0.71 ถึง 1.00) อีกช่วงหนึ่งของค่าความน่าจะเป็นของความสามารถถูกเลือกในการตัดสินระดับความรอบรู้ ในการดำเนินงานสถานการณ์การทดสอบ ระดับประสิทธิภาพสำหรับแต่ละสถานะความรอบรู้ที่จะเป็นไปได้

ถูกกำหนดโดยใช้กระบวนการที่เป็นมาตรฐาน ผลคะแนนรายงานเผยว่าผู้เข้าสอบมีความรอบรู้ใน 4 คุณลักษณะอย่างชัดเจน A1(การคำนวณทางคณิตศาสตร์ขั้นพื้นฐาน) A4 (ทักษะที่จำเป็นสำหรับการแทนที่ค่าเป็นนิพจน์พีชคณิต), A5 (ทักษะของการทำแผนที่กราฟของฟังก์ชันที่คุ้นเคยกับฟังก์ชันที่สอดคล้องกัน) และ A6 (คุณสมบัตินามธรรมของฟังก์ชัน).

การศึกษาของไพฑูรย์ เอื้อบุญประดิษฐ์ (2563) เกี่ยวกับการพัฒนาแบบสอบวินิจฉัยการรู้เรื่องคณิตศาสตร์ที่มีการให้ข้อมูลย้อนกลับสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีลำดับชั้นของคุณลักษณะ นำเสนอรายงานการวินิจฉัย

การรู้เรื่องคณิตศาสตร์ (Math Literacy) คะแนนของชั้น 18		ชื่อ (นางสาว ก.) โรงเรียน (A)	
คุณลักษณะ	ความรอบรู้		
	รอบรู้	ไม่รอบรู้	
ภาพรวม: การรู้เรื่องคณิตศาสตร์			
ความสามารถในการระบุประเด็นทางคณิตศาสตร์ของปัญหาในชีวิตจริง			
ความสามารถในการแปลงปัญหาให้อยู่ในรูปของภาษาทางคณิตศาสตร์			
ความสามารถในการใช้หลักการและกระบวนการทางคณิตศาสตร์ในการแก้ปัญหา			
ความสามารถในการตีความ การประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์			

ข้อแนะนำ

นักเรียนมีข้อบกพร่องเกี่ยวกับการรู้เรื่องคณิตศาสตร์ โดยอาจมีความเข้าใจที่คลาดเคลื่อนหรือมีข้อบกพร่องที่ควรแก้ไขเพื่อให้การรู้เรื่องคณิตศาสตร์มีความถูกต้องสมบูรณ์ ด้วยการเพิ่มเติมข้อเสริมใน 3 คุณลักษณะ ได้แก่

- 1) การแปลงปัญหาให้อยู่ในรูปของภาษาทางคณิตศาสตร์ นักเรียนควรได้รับการฝึกเกี่ยวกับการอธิบายความสัมพันธ์ระหว่างบริบทของปัญหากับภาษาสัญลักษณ์ที่จำเป็นต้องใช้ในการแสดงเชิงคณิตศาสตร์ เพื่อให้สามารถแปลงปัญหาจากสถานการณ์ในชีวิตจริงให้อยู่ในขอบเขตของคณิตศาสตร์ โดยอาจนำเสนอสถานการณ์เชิงคณิตศาสตร์โดยใช้ตัวแปร สัญลักษณ์ เครื่องหมายแทน แผนภาพ แบบจำลองมาตรฐานที่เฉพาะสม
- 2) การใช้หลักการและกระบวนการทางคณิตศาสตร์ในการแก้ปัญหา นักเรียนควรได้รับการฝึกเกี่ยวกับการคำนวณ การแก้สมการ การสกัดข้อมูลทางคณิตศาสตร์จากตารางและกราฟ การจัดการกับรูปร่างและรูปทรง การวิเคราะห์หาค่า และ การสร้างแบบจำลองของสถานการณ์ปัญหา เพื่อให้สามารถใช้หลักการและกระบวนการทางคณิตศาสตร์ในการแก้ปัญหาได้ดียิ่งขึ้น
- 3) การตีความ การประยุกต์ใช้และการประเมินผลลัพธ์ทางคณิตศาสตร์ นักเรียนควรได้รับการฝึกเกี่ยวกับการสะท้อนวิธีแก้ปัญหา ผลลัพธ์ หรือข้อสรุปทางคณิตศาสตร์ แล้วตีความออกมาในบริบทของปัญหาโลกชีวิตจริง และตัดสินใจว่าผลลัพธ์ที่ได้เป็นเหตุเป็นผลและเข้ากับบริบทของปัญหาหรือไม่

ภาพ 38 แสดงตัวอย่างแบบรายงานคะแนนเชิงวินิจฉัย

(ไพฑูรย์ เอื้อบุญประดิษฐ์ ,2563)

จากภาพ 38 สรุปได้ว่า นางสาว ก. มีข้อบกพร่องเกี่ยวกับการรู้เรื่องคณิตศาสตร์ ควรได้รับการพัฒนาใน 3 คุณลักษณะ ได้แก่ 1) การแปลงปัญหาให้อยู่ในรูปของภาษาทางคณิตศาสตร์ 2) การใช้หลักการและกระบวนการทางคณิตศาสตร์ในการแก้ปัญหา และ 3) การอนุมานการประยุกต์ใช้ และการประเมินผลลัพธ์ทางคณิตศาสตร์

ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง

5.1 งานวิจัยในประเทศ

กุลพร พูลสวัสดิ์ (2556) ทำการศึกษาพัฒนาวิธีวินิจฉัยความสามารถในการใช้ภาษาอังกฤษ เพื่อการสื่อสารสำหรับนักเรียนมัธยมศึกษาปีที่ 3 วัดอุปประสงค์ 1) เพื่อพัฒนาวิธีการประเมินวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสาร 2) เพื่อตรวจสอบคุณภาพของวิธีการประเมินเชิงวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสาร 3) เพื่อวิเคราะห์รูปแบบของพฤติกรรมการใช้ภาษาอังกฤษเพื่อการสื่อสารที่บกพร่องของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 กลุ่มตัวอย่าง เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 127 คนผู้เชี่ยวชาญด้านภาษาอังกฤษและผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา จำนวน 12 คน การวิจัยแบ่งออกเป็น 6 ขั้นตอน คือ 1) การทบทวนวรรณกรรม 2) การระบุคุณลักษณะที่ต้องการวินิจฉัยและสร้างเครื่องมือวินิจฉัย 3) การนำเครื่องมือวินิจฉัยไปทดลองใช้และสำรวจรูปแบบพฤติกรรมการใช้ภาษาอังกฤษเพื่อการสื่อสารที่บกพร่อง 4) การสร้างแบบตรวจสอบรายการพฤติกรรมและสร้าง Q-Matrix 5) การวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสาร และ 6) การศึกษาคุณภาพของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นและให้ข้อมูลวินิจฉัยย้อนกลับแก่ผู้สอบ ผลการวิเคราะห์คุณภาพของแบบสอบวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารตามทฤษฎีการทดสอบแบบดั้งเดิม พบว่ามีค่าความยาก 0.17 ถึง 0.92 ค่าอำนาจจำแนก 0.12 ถึง 0.97 ค่าความเที่ยงเท่ากับ 0.95 วิเคราะห์สารสนเทศข้อสอบตามทฤษฎีการตอบสนองข้อสอบ พบว่ามีค่าพารามิเตอร์ความยาก -5.8 ถึง 1.98 และค่าพารามิเตอร์อำนาจจำแนก 0.18 ถึง 11.11 คุณภาพของแบบสอบถามความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารสำหรับนักเรียนในเรื่องความคิดเห็นความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารของนักเรียนมีค่า แอลฟาของครอนบาค .952 และ คุณภาพของแบบสอบถามในส่วนของปัญหาที่พบของนักเรียนมีค่า แอลฟาของครอนบาค .801 การวิเคราะห์รูปแบบพฤติกรรมการใช้ภาษาอังกฤษเพื่อการสื่อสารที่บกพร่อง พบว่า นักเรียนมีความบกพร่องในองค์ประกอบที่ 1 ความสามารถด้านไวยากรณ์ ด้านการใช้คำศัพท์มากที่สุด ร้อยละ 55.17 รองลงมา คือ การสร้างประโยคที่สมบูรณ์ ร้อยละ 34.10 นักเรียนมีพฤติกรรมการใช้ภาษาอังกฤษบกพร่องในองค์ประกอบที่ 2 ความสามารถด้านภาษาศาสตร์สังคมการใช้สำนวนรูปแบบภาษาที่เหมาะสมมากที่สุด ร้อยละ 99.7 และยังพบว่า วิธีวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารที่พัฒนาขึ้นให้ผลวินิจฉัยที่

ละเอียด ระบุรูปแบบพฤติกรรมการใช้ภาษาอังกฤษบกพร่องของนักเรียนให้ผลการประเมินที่มีความเที่ยง และความตรง

ทิพย์ ขำอยู่ (2557) ทำการศึกษาวิจัย เรื่อง ผลการวินิจฉัยทักษะการอ่านภาษาอังกฤษของนักศึกษาระดับปริญญาตรี โดยวิธีลำดับชั้นคุณลักษณะและการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ วัตถุประสงค์ เพื่อวินิจฉัยความรอบรู้ในทักษะการอ่านภาษาอังกฤษของนักศึกษาระดับปริญญาตรี โดยวิธีลำดับชั้นคุณลักษณะและการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ มี 5 ขั้นตอน คือ 1) กำหนดลำดับชั้นคุณลักษณะเนื้อหา 2) พัฒนาชุดข้อสอบ 3) ตรวจสอบคุณภาพของชุดข้อสอบ 4) จัดทำคลังข้อสอบสำหรับใช้ทดสอบบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 5) วินิจฉัยความรอบรู้ของนักศึกษาจากแบบแผนการตอบข้อสอบที่สังเกตได้ กลุ่มตัวอย่าง เป็นนักศึกษาระดับปริญญาตรี จำนวน 1,763 คน สุ่มตัวอย่างแบบมีระบบ วิเคราะห์ข้อมูลด้วยสถิติพื้นฐาน F-test ผลการวิจัย การจำแนกคุณลักษณะเนื้อหา 6 คุณลักษณะด้วยข้อสอบ 760 ข้อ 1) ชุดข้อสอบคุณลักษณะที่ 1 Part of Speech จำนวน 160 ข้อ 2) ชุดข้อสอบเนื้อหาคุณลักษณะที่ 2 Using Context Clues จำนวน 120 ข้อ 3) ชุดข้อสอบเนื้อหาคุณลักษณะที่ 3 Word Formation จำนวน 120 ข้อ 4) ชุดข้อสอบเนื้อหาคุณลักษณะ ที่ 4 Identify Topic and Main Idea จำนวน 120 ข้อ 5) ชุดข้อสอบคุณลักษณะที่ 5 Reading for Details จำนวน 120 ข้อ และ 6) ชุดข้อสอบคุณลักษณะที่ 6 Making Inferences จำนวน 120 ข้อรวมทั้งสิ้น 760 ข้อ สรุปผลการวิจัย ได้ค่าอำนาจจำแนกข้อสอบตั้งแต่ 0.35 ถึง 2.49 ค่าความยากตั้งแต่ -1.67 ถึง 2.47 ค่าการเดาข้อสอบตั้งแต่ 0.02 ถึง 0.29 ค่าความเที่ยงตั้งแต่ 0.82 ถึง 0.95

สุมาลี มีสกุล,องอาจ นัยพัฒน์, ชูศักดิ์ ชัมภลลิขิต, และสุวิมล กฤษชฤทธานนท์ (2558) ทำการศึกษาวิจัย เรื่อง การศึกษาเพื่อเปรียบเทียบประสิทธิภาพของโมเดลวินิจฉัยเชิงจำแนกในการวินิจฉัยทักษะความสามารถทางพุทธิปัญญา การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของโมเดลที่นำมาใช้ในการวินิจฉัย ทักษะความสามารถทางพุทธิปัญญา ในการแก้ปัญหา เรื่องสมการเชิงเส้นตัวแปรเดียวของนักเรียนระดับชั้นมัธยมศึกษาตอนต้นรวมทุกระดับชั้น กลุ่มนักเรียนที่มีระดับ ชั้น เพศ แผนการเรียน แตกต่างกันวิธีดำเนินการวิจัยโดยนำแบบทดสอบที่ใช้ในการประเมินเพื่อวินิจฉัยทักษะความสามารถทางพุทธิปัญญาของนักเรียนระดับชั้นมัธยมศึกษาตอนต้นที่ใช้แก้ปัญหาเรื่องสมการเชิงเส้นตัวแปรเดียว ซึ่งพัฒนาขึ้นด้วยวิธีการตามแนวคิดของ Embretson's Cognitive Design System Approach (CDS Framework) ไปใช้ทดสอบกับนักเรียนจำนวน 1,214 คน วิเคราะห์ข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของโมเดลการวินิจฉัยเชิงจำแนก คือ G-DINA Model และ DINA Model ตามตัวแปรปัจจัยที่ส่งผลต่อความแตกต่างความสามารถในการใช้ทักษะทางพุทธิปัญญาเพื่อการหาคำตอบของปัญหา ผลการวิจัยปรากฏว่า G-DINA Model เป็นโมเดลที่ให้ค่าดัชนีความสอดคล้องของโมเดลในเชิงสัมบูรณ์และค่าดัชนีความสอดคล้องของโมเดลในเชิงสัมพัทธ์

ต่ำกว่าค่าที่ได้จาก DINA Model ในทุกเงื่อนไขการวิเคราะห์ แสดงให้เห็นว่า G-DINA Model เป็นโมเดลที่ให้ประสิทธิภาพในการวิเคราะห์ข้อมูลเชิงการวินิจฉัยดีกว่า DINA Model ทั้งการวิเคราะห์กับกลุ่มนักเรียนระดับชั้นมัธยมศึกษาตอนต้นทุกระดับชั้น ที่ใช้วิเคราะห์กลุ่มของนักเรียนที่มี ชั้น เพศ และแผนการเรียน ต่างกัน

สยามรัก สว่างศรี (2560) สร้างแบบทดสอบวินิจฉัยทางการเรียนกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ เรื่อง ความสามารถด้านไวยากรณ์วิชาภาษาอังกฤษ (English Grammar Ability) ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 และเพื่อวินิจฉัยความสามารถด้านไวยากรณ์วิชาภาษาอังกฤษ (English Grammar Ability) ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 โดยใช้การสอบผ่านคอมพิวเตอร์ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2555 สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 27 จำนวน 1,271 คน ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอนเครื่องมือที่ใช้ในการวิจัย ประกอบด้วย 1) แบบทดสอบเพื่อสำรวจข้อบกพร่องทางการเรียน เป็นแบบทดสอบชนิดเติมคำ จำนวน 80 ข้อ ใช้ทดสอบกับกลุ่มทดลอง 155 คน เพื่อหาข้อบกพร่อง 2) แบบทดสอบวินิจฉัยทางการเรียนกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ เรื่อง ความสามารถด้านไวยากรณ์วิชาภาษาอังกฤษ ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 เป็นชนิดเลือกตอบ 4 ตัวเลือก จำนวน 80 ข้อ ใช้ทดสอบครั้งที่ 1 กับกลุ่มทดลอง จำนวน 147 คน เพื่อหาคุณภาพข้อสอบรายข้อตามทฤษฎี การทดสอบแบบดั้งเดิม หาค่าความยาก ค่าอำนาจจำแนกรายข้อ และค่าความเชื่อมั่น 3) แบบทดสอบวินิจฉัยทางการเรียนกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ เรื่องความสามารถด้านไวยากรณ์วิชาภาษาอังกฤษ ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 60 ข้อ ใช้ทดสอบครั้งที่ 2 กับกลุ่มทดลอง จำนวน 599 คน เพื่อหาคุณภาพข้อสอบรายข้อตามทฤษฎีการตอบสนองข้อสอบ 4) แบบทดสอบวินิจฉัยทางการเรียนกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ เรื่อง ความสามารถด้านไวยากรณ์วิชาภาษาอังกฤษ ของนักเรียนชั้นมัธยมศึกษา ปีที่ 2 จำนวน 40 ข้อ เพื่อบรรจุไว้ในโปรแกรมคอมพิวเตอร์ ใช้ในการวินิจฉัยข้อบกพร่องทางการเรียนของนักเรียนชั้นมัธยมศึกษาปีที่ 2

ศุภามณ จันทรสกุล, และสุกัญญา บุญศรี (2560) ทำการศึกษาด้วยการใช้โมเดล DINA และโมเดล G-DINA เพื่อวินิจฉัยความรอบรู้ทางการพยาบาล โมเดลวินิจฉัยจัดประเภทเป็นวิธีการที่สนใจการประเมินวินิจฉัยความรอบรู้ทางการพยาบาล บทความมีจุดประสงค์ เพื่อนำเสนอวิธีการประเมินวินิจฉัยโดยใช้โมเดลทางสถิติในการวินิจฉัยและใช้วิธีการทางสถิติ ประมาณค่าคุณลักษณะทักษะต่าง ๆ ที่ต้องการวัดเพื่อจัดประเภทผู้สอบ โมเดลสถิติในการวินิจฉัยจัดประเภทได้แก่ G-DINA Model เป็นโมเดลแบบชดเชยได้ การตอบถูกในข้อถามนั้นโดยทักษะบางทักษะสามารถถูกชดเชยได้ด้วยทักษะที่สูงกว่า และ DINA Model เป็นโมเดลแบบไม่ชดเชยซึ่งผู้สอบจะตอบถูกเมื่อมีทักษะครบทุกทักษะในข้อคำถามนั้น การประเมินประสิทธิภาพของโมเดลพิจารณาจากค่าสถิติความสอดคล้อง โดยพิจารณาค่าดัชนีความสอดคล้องของโมเดลเชิงสัมบูรณ์และเชิงสัมพัทธ์

พิศิษฐ ตัณฑวณิช (2560) การพัฒนา “แบบสอบแนวคิดพื้นฐานการวิเคราะห์ข้อสอบ” ตามแนวทางทฤษฎีการพัฒนาแบบสอบวินิจฉัยโดยวิธีการทางพุทธิปัญญา (การศึกษาระยะที่ 2) เพื่อ (1) ศึกษาความสัมพันธ์สหสัมพันธ์ของข้อสอบใน “แบบสอบแนวคิดพื้นฐานการวิเคราะห์ข้อสอบ” ที่สร้างขึ้นตามแนวทางการพัฒนาแบบสอบวินิจฉัยโดยวิธีการทางพุทธิปัญญา เมื่อวิเคราะห์ในรูปแบบของทฤษฎีดั้งเดิม รูปแบบ The DINA Model The G-DINA Model และ The rRUM Model (2) ศึกษาตัวประกอบ (Factor) ของโอกาสความน่าจะเป็นในการสอบผ่านในแต่ละคุณลักษณะย่อยที่ปรากฏในแบบสอบและกับคะแนนเกณฑ์เมื่อวิเคราะห์ในรูปแบบของโมเดลทั้ง 3 แบบ (3) วิเคราะห์จำนวนค่าความถี่ของโอกาสความน่าจะเป็นในการสอบผ่านและเป็นผู้รอบรู้ในคุณลักษณะย่อยต่าง ๆ ของผู้สอบในแบบสอบที่ระดับคะแนนรวมจากต่ำสุดไปสูงสุด ผลการวิจัยพบว่า 1) จากผลการศึกษาค่าสัมพันธ์สหสัมพันธ์ของค่าประจำข้อสอบพบว่ากลุ่มค่าความยาก/ค่าอำนาจจำแนกในแนวทฤษฎีดั้งเดิมและการวิเคราะห์โดยโมเดลทั้ง 3 แบบ มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ทำให้นักวิชาการมีความมั่นใจและเห็นว่าแนวทางการวิเคราะห์ข้อสอบแบบใหม่ ผลที่ได้ไม่ต่างกัน อีกทั้งโมเดลการวิเคราะห์ข้อสอบทั้ง 3 โมเดล ต่างก็ให้ผลการวิเคราะห์เป็นไปในทิศทางเดียวกันทั้งสิ้น 2) จากการศึกษาค่าตัวประกอบของโอกาสความน่าจะเป็นในการเป็นผู้รอบรู้ในคุณลักษณะย่อย ในแบบสอบที่พัฒนาขึ้นและกับคะแนนเกณฑ์ด้วยโมเดลทั้ง 3 แบบ พบว่า ตัวแปรทั้งหมดรวมกันวัดองค์ประกอบเพียงองค์ประกอบเดียว ก่อให้เกิดความมั่นใจในด้านความตรงของผลที่ได้ในการตัดสินระดับการรอบรู้รายบุคคลจากโปรแกรมการวิเคราะห์ข้อมูล มีประเด็นปัญหาที่ควรจะต้องพิจารณาคือการกำหนดความใหญ่เล็กของคุณลักษณะใน Q-Matrix ในการวิจัยครั้งนี้ขอบเขตเนื้อหาสาระในการสร้างข้อสอบเป็นกรอบเนื้อหาที่ค่อนข้างแคบ ซึ่งอาจมีผลทำให้ค่าโอกาสความน่าจะเป็นในแต่ละคุณลักษณะและกับตัวแปรเกณฑ์รวมตัวกันเกิดเป็นองค์ประกอบเดียว โดยมีประเด็นน่าสงสัยทางวิชาการว่า หากในขอบเขตของเนื้อหาที่ค่อนข้างกว้างมากขึ้น คุณลักษณะต่าง ๆ ที่มุ่งวัดอาจมีการกระจายตัวมากกว่าและแตกต่างกันมากกว่า ผลการวิจัยที่ได้ก็อาจจะแตกต่างกันออกไปได้

อรอนงค์ อุ้นแก้ว และคณะ (2562) ทำการศึกษาวิจัย เรื่องการสร้างแบบทดสอบวินิจฉัยความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 โรงเรียนสังกัดกรมส่งเสริมการปกครองท้องถิ่น จังหวัดสมุทรสาคร วัดดูประสงค์ 1) เพื่อสร้างแบบทดสอบวินิจฉัยความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษ 2) เพื่อหาคุณภาพด้านความตรงตามเนื้อหา ความยาก อำนาจจำแนก ความตรงตามสภาพ ความเที่ยง 3) เพื่อหาเกณฑ์การวินิจฉัยและสร้างคู่มือการใช้แบบทดสอบ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 โรงเรียนสังกัดกรมส่งเสริมการปกครองท้องถิ่น จังหวัดสมุทรสาคร จำนวน 643 คน สุ่มตัวอย่างแบบแบ่งชั้น เครื่องมือคือ แบบทดสอบวินิจฉัยความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษ ชนิด

เลือกตอบ 4 ตัวเลือก จำนวน 5 ฉบับ ฉบับละ 20 ข้อ รวม 100 ข้อ ผลการวิจัยพบว่า แบบทดสอบวินิจฉัยทั้ง 5 ฉบับ ได้แก่ ฉบับที่ 1 ความเข้าใจในเนื้อหา 20 ข้อ ฉบับที่ 2 การเชื่อมโยงองค์ความรู้ 20 ข้อ ฉบับที่ 3 การสะท้อนกลับของคำตอบ 20 ข้อ ฉบับที่ 4 การอนุมาน 20 ข้อ และฉบับที่ 5 การวิเคราะห์ประเมินค่าเนื้อหา 20 ข้อ ค่าความตรงตามเนื้อหา ค่า IOC ตั้งแต่ 0.80 – 1.00 ค่าความยาก ตั้งแต่ 0.41 – 0.76 ค่าอำนาจจำแนกใช้สูตรของ Brennan มีค่าตั้งแต่ 0.51 – 0.78 ค่าความตรงตามสภาพใช้สูตรสหสัมพันธ์ของ Spearman มีค่าตั้งแต่ 0.72 – 0.87 ที่ระดับนัยสำคัญทางสถิติที่ระดับ 0.01 ส่วนค่าความเที่ยงใช้สูตรของ Livingston มีค่าตั้งแต่ 0.92 – 0.93 สำหรับคะแนนจุดตัด หาโดยวิธีของ Angoff แต่ละฉบับ เท่ากับ 12 คะแนน จากคะแนนเต็มฉบับละ 20 คะแนน

แคทรียา แสงดาวเทียน (2562) การเปรียบเทียบประสิทธิภาพการประเมินวินิจฉัยเชิงพุทธิปัญญา เรื่องสมการเชิงเส้นตัวแปรเดียว ระหว่างโมเดลดีไอเอ็นโอ (DINO model) โมเดลดีไอเอ็นเอ (DINA model) และโมเดลจี-ดีไอเอ็นเอ (G-DINA model) เมื่อเปรียบเทียบประสิทธิภาพของโมเดลที่นำมาใช้ในการวินิจฉัยทางพุทธิปัญญาโดยพิจารณาจากดัชนีความสอดคล้องของโมเดลในเชิงสัมพัทธ์ โดยพิจารณาจากค่า -2Loglikelihood ($-2LL$), Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC) และ Root Mean Square Error of Approximation (RMSEA) โมเดลที่มีประสิทธิภาพมากที่สุดคือ G-DINA model รองลงมาอันดับที่สองคือ DINA model และโมเดลที่มีประสิทธิภาพน้อยที่สุดคือ DINO model

วรางคณา ก้อมน้อย (2563) ทำการพัฒนาแบบทดสอบสร้างและตรวจสอบคุณภาพของแบบทดสอบเชิงวินิจฉัย ทางพุทธิปัญญา เรื่องความรู้สึกลงใจจำนวนสำหรับนักเรียนระดับชั้นประถมศึกษาตอนต้น และเพื่อ วินิจฉัยความรู้สึกลงใจจำนวนของนักเรียนระดับชั้นประถมศึกษาปีที่ 3 จำนวน 1220 คน ตัวแปรที่ศึกษา คือ รูปแบบความรอบรู้ของนักเรียน ผลการวิจัย พบว่าแบบทดสอบ ๓ มี ค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับคุณลักษณะที่มุ่งวัด อยู่ในช่วง .50 ถึง .82 มีความ ตรงตามเกณฑ์สัมพัทธ์ เมตริกซ์ควิปรีมีความเหมาะสมกับข้อมูล มีค่าดัชนีอำนาจจำแนกรายข้อ อยู่ในช่วง .204 – 1.000 รายฉบับ เท่ากับ .606 มีค่าความตรงเชิงจำแนกรายฉบับ เท่ากับ .734 รายคุณลักษณะ ในช่วง .901 - .964 มีค่าความเที่ยงเชิงจำแนกรายฉบับ เท่ากับ .652 รายคุณลักษณะอยู่ในช่วง .842 - .949 และผลการตรวจสอบความสอดคล้องของโมเดล พบว่า ค่า RMSEA2 เท่ากับ .021และผลการวินิจฉัยพบว่า นักเรียนมีรูปแบบความรอบรู้ 19 แบบ ดังนั้น หากต้องการพัฒนานักเรียนได้ตรงจุดบกพร่อง หน่วยงานที่สนใจควรสร้างชุดกิจกรรมหรือจัดกิจกรรม การเรียนรู้ที่สอดคล้องกับ 19 รูปแบบความรอบรู้ของนักเรียน

เกศรินทร์ สมราช (2563) ทำการศึกษาเรื่องการพัฒนาแบบสอบวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารของนิสิตครู โดยประยุกต์ใช้ G-DINA Model กลุ่มตัวอย่าง คือ นิสิตครู (หลักสูตรสี่ปี) ที่กำลังศึกษาอยู่ในภาคการศึกษาที่ 2 ปีการศึกษา 2563 จำนวน 566 คน เก็บข้อมูล

ทางระบบออนไลน์ วินิจฉัยด้วยโมเดลการวินิจฉัยประเภท Generalization DINA (G-DINA MODEL) ผลการวิจัย พบว่า 1) แบบสอบวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารสำหรับ นิสิตครู ที่พัฒนาขึ้น มีลักษณะเป็นแบบสอบปรนัยเลือกตอบ 4 ตัวเลือก จำนวน 49 ข้อ เป็นไปตาม กรอบมาตรฐาน CEFR ในระดับ B1 และเกณฑ์ความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารที่ ครูสุภากำหนด ผลการวิเคราะห์ประสิทธิภาพความสอดคล้องเชิงสัมพันธ์ด้วย G-DINA MODEL ผ่าน เกณฑ์ความสามารถทั้ง 4 ด้าน 2) โมเดลสอดคล้องกับคะแนนข้อมูลเชิงประจักษ์ สามารถนำไปใช้ในการ วินิจฉัยความสามารถของผู้สอบได้ และผลการวิเคราะห์ความถูกต้องของการวินิจฉัย พบว่า แบบทดสอบที่พัฒนาขึ้นมีคุณภาพด้านความถูกต้องของการวินิจฉัย ผ่านเกณฑ์ทุกด้านและทุก คุณลักษณะย่อย มีค่าพารามิเตอร์การเดา (Guessing Parameter) และค่าพารามิเตอร์ความสะเพร่า (Slipping Parameter) ผ่านเกณฑ์ ค่าดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ (RMSEA) ในระดับต่ำ สามารถวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารของนิสิตครู แยกตามความสามารถแต่ละด้านที่นำเสนอในรูปกลุ่มแฝง (Latent Class) ของผู้สอบที่มีความ สามารถในด้านต่างๆ ได้ตามคุณลักษณะ 3) ผลการพัฒนา PROFILE พบว่า PROFILE สามารถ ให้ข้อมูลสารสนเทศที่ได้จากการวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารของนิสิต ครูทั้ง 4 ด้านแยกจากกัน ประกอบด้วย คะแนนดิบจากการทดสอบแบบข้อต่อข้อ (1/0) แบบแผนการ มีคุณลักษณะในความสามารถด้านนั้นๆ ตามกลุ่มแฝง (Latent Class) ของผู้สอบโดยสามารถบอกค่า ความน่าจะเป็น (Probability) ของการรอบรู้คุณลักษณะนั้น รวมถึงสารสนเทศที่เป็นความสามารถที่ ผู้สอบผู้นั้นสามารถปฏิบัติได้ และข้อบกพร่องที่ผู้สอบควรจะต้องทำการปรับปรุง

ไพฑูริย์ เอื้อบุญประดิษฐ์ (2563) ทำการศึกษาเกี่ยวกับการพัฒนาแบบสอบวินิจฉัยการรู้ เรื่องคณิตศาสตร์ที่มีการให้ข้อมูลย้อนกลับสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีลำดับขั้นของ คุณลักษณะนำเสนอรายงานการวินิจฉัย โดยมีตัวอย่างวิจัย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 956 คน ที่ได้จากการสุ่มตัวอย่างแบบหลายขั้นตอน เครื่องมือวิจัยเป็นแบบสอบวินิจฉัยการรู้เรื่อง คณิตศาสตร์ตรวจสอบคุณภาพของเครื่องมือ ด้านความตรงเชิงโครงสร้างด้วยการวิเคราะห์ องค์ประกอบเชิงยืนยันและด้านความเที่ยงด้วยวิธีสัมประสิทธิ์แอลฟา ของครอนบาค วิเคราะห์ข้อมูล โดยใช้ความถี่ ร้อยละ และสถิติทดสอบไคสแควร์คำนวณคะแนนเชิงวินิจฉัย โดยประยุกต์ใช้เครือข่าย เบย์เซียน และตรวจสอบคุณภาพของการวินิจฉัยด้วยค่าสัมประสิทธิ์ความสอดคล้องของ แคปปา ผลการวิจัย 1) แบบสอบที่พัฒนาขึ้นประกอบด้วย 8 สถานการณ์สถานการณ์ละ 4 คำถามย่อย รวม 32 ข้อ ประกอบด้วย 4 คุณลักษณะ ได้แก่ การระบุประเด็นทางคณิตศาสตร์ของปัญหาในชีวิตจริง การแปลงปัญหาให้อยู่ในรูปของภาษาทางคณิตศาสตร์ การใช้หลักการและกระบวนการทาง คณิตศาสตร์ในการแก้ปัญหา และการอนุมานการประยุกต์ใช้ และการประเมินผลลัพธ์ทาง คณิตศาสตร์ 2) แบบสอบที่พัฒนาขึ้นมีความตรงเชิงโครงสร้าง ($\text{Chi-square} = 334.770, p = 0.918,$

GFI = 0.931, AGFI = 0.902, RMSEA = 0.000) มีค่าสัมประสิทธิ์แอลฟาของครอนบาคตั้งแต่ 0.763 ถึง 0.842 สามารถวินิจฉัยข้อบกพร่องเกี่ยวกับการรู้เรื่องคณิตศาสตร์ของผู้ตอบ และ 3) ผลการวินิจฉัยการรู้เรื่องคณิตศาสตร์ ในภาพรวมพบว่ามีนักเรียนที่มีข้อบกพร่องจำนวน 378 คน คิดเป็นร้อยละ 56.17 และมีนักเรียนที่ไม่มีข้อบกพร่อง จำนวน 295 คน คิดเป็นร้อยละ 43.83 โดยคุณลักษณะที่มีนักเรียนที่มีข้อบกพร่องมากที่สุด คือ การใช้หลักการ และกระบวนการทางคณิตศาสตร์ในการแก้ปัญหา ซึ่งมีนักเรียนที่มีข้อบกพร่องจำนวน 420 คน คิดเป็นร้อยละ 62.41 และคุณลักษณะที่มีนักเรียนที่มีข้อบกพร่องน้อยที่สุด คือ การระบุประเด็นทางคณิตศาสตร์ของปัญหาในชีวิตจริง ซึ่งมีนักเรียนที่มีข้อบกพร่องจำนวน 273 คน คิดเป็นร้อยละ 40.56

5.2 งานวิจัยในต่างประเทศ

Lingyun Gao (2006) เถาได้ทำการศึกษาพัฒนาการทดสอบด้วยโมเดลของกระบวนการทางพุทธิปัญญา (CDA) ที่ตั้งสมมติฐานไว้ภายใต้ประสิทธิภาพการทดสอบการอ่านที่เรียกว่า MELAB ทำการวิเคราะห์คะแนนจาก การประเมินความสามารถทางภาษา หรือ Michigan English Language Assessment Battery (MELAB) สองแบบโดยกระบวนการตรวจสอบสามเสา ได้แก่ (1) ทบทวนโมเดลทางทฤษฎีของกระบวนการอ่านของผู้ที่ใช้ภาษาอังกฤษเป็นภาษาที่สอง (L2) และสร้างคุณลักษณะความสามารถในการอ่านระดับ L2 (2) วิเคราะห์ความต้องการทางพุทธิปัญญาของ MELAB ทำโดยรวบรวมรายงานการคิดออกเสียง (Think aloud) ของผู้เข้าสอบขณะทำข้อการทดสอบเมื่อตอบคำถามถูกต้อง และ (3) ตรวจสอบความสัมพันธ์ระหว่างกระบวนการทางพุทธิปัญญา และตัวบ่งชี้เชิงประจักษ์ของความยากข้อสอบโดยใช้การวัดตามความรู้ความเข้าใจโมเดลการถอดแบบต้นไม้ (TBR) ผลการศึกษาทำให้ทราบการตรวจสอบความถูกต้องของการทดสอบการอ่าน MELAB การสร้างข้อสอบ และการวางรากฐานสำหรับ MELAB ให้เป็นมาตรฐานวินิจฉัย ส่งเสริมการรวมกันของจิตวิทยาการรับรู้และการประเมินในด้านการทดสอบภาษาที่สอง/ภาษาต่างประเทศ

Jang (2009) ศึกษาการประเมินเชิงวินิจฉัยทางพุทธิปัญญาความสามารถความเข้าใจการอ่านโดยใช้ภาษาอังกฤษเป็นภาษาที่ 2 โดยมีวัตถุประสงค์ 1) เพื่อประเมินความตรงของการนำผลการประเมินเชิงวินิจฉัยทางพุทธิปัญญาไปใช้กับแบบสอบความสามารถความเข้าใจการอ่านภาษาที่ 2 ที่ไม่ใช่แบบสอบวินิจฉัย 2) เพื่อให้สารสนเทศความท้าทายและเงื่อนไขของวิธีการประเมินเชิงวินิจฉัยทางพุทธิปัญญา การศึกษาครั้งนี้เน้นที่ความเชื่อถือได้ของโมเดลฟิวชั่น ลักษณะผล ทักษะการอ่าน และความสามารถในการวินิจฉัยข้อสอบ วิธีดำเนินการแบ่งออกเป็น 3 ระยะ 1) กำหนดทักษะความเข้าใจในการอ่าน 9 ทักษะ ด้วยการคิดออกเสียงทำการวิเคราะห์ข้อสอบและ แบบสอบ 2) ตรวจสอบลักษณะโปรไฟล์ทักษะที่ประมาณค่าโดยโมเดลฟิวชั่น ซึ่งมีความสอดคล้องกับข้อมูลจากแบบสอบ LanguEdge™ 3) ประเมินสารสนเทศเชิงวินิจฉัยจากห้องเรียน โดยผู้เรียนจะทำการทดสอบเชิงวินิจฉัย ก่อนและหลังเรียน นอกจากนี้ยังใช้การสัมภาษณ์ สังเกตชั้นเรียน และวิธีการสำรวจด้วย เพื่อ

ใช้ในการให้ข้อมูลย้อนกลับ กลุ่มตัวอย่าง 2,703 คน ที่สอบ Langu Edge ในปี 2002 และ ผู้เรียน 28 คน ครู 3 คน ในหลักสูตรเตรียมสอบ TOEFL เพื่อประเมินรายงานเชิงวินิจฉัยระยะที่ 3 เครื่องมือที่ใช้คือ แบบสอบ Langu Edge ใช้ข้อมูล ทฤษฎีกับโมเดลฟิวชั่น ในระยะที่ 3 และใช้ในการประเมินเชิงวินิจฉัยทั้งก่อน - หลังเรียนในระยะที่ 3 2) แบบสอบถามประเมินตนเองของผู้สอบ Langu Edge จะถามผู้เรียนเกี่ยวกับระดับ ความมั่นใจในการอ่าน การเขียน และการฟัง 3) แบบสอบถามผู้เรียนในการรายงานเชิง ประเมินผลการศึกษาเสนอแนะว่าวิธีการประเมินเชิงวินิจฉัยทางพุทธิปัญญาสามารถให้สารสนเทศจากการวินิจฉัยระดับสมรรถนะการอ่านได้ละเอียดมากกว่าการให้คะแนนแบบดั้งเดิม มีหลักฐานเชิงประจักษ์ที่สนับสนุนความเชื่อถือได้ของกระบวนการ โปรไฟล์ทักษะ โมเดลมี ความสอดคล้องกับข้อมูล แต่เมื่อโมเดลวินิจฉัยทางพุทธิปัญญาถูกแก้ไขปรับปรุงเพื่อพัฒนาแบบ สอบสำหรับวัตถุประสงค์ที่ไม่มุ่งวินิจฉัย มันอาจจะสูญเสียความสามารถในการวินิจฉัยบางอย่าง ไประหว่างข้อสอบกับระดับความยากที่สุดได้ ข้อสอบเหล่านี้ไม่จำเป็นต้องใช้การประเมิน สมรรถนะทักษะของผู้เรียน

Templin, & Henson (2006) ประยุกต์ใช้โมเดลไอเอ็นเอ ซึ่งเป็นโมเดลวินิจฉัยทางพุทธิปัญญามาใช้ในการประเมินทางจิตวิทยาและใช้ข้อมูลแนวโน้มการเสี่ยงโชคของผู้เรียนระดับอุดมศึกษา จำนวน 593 คน ด้วยแบบวัด GRI จำนวน 41 ข้อ เป็นมาตรฐานประมาณค่า 6 ระดับ จำแนกคุณลักษณะ 10 คุณลักษณะสร้างเป็น Q - matrix ทำการวิเคราะห์ด้วยโมเดลไอ เอ็นโอ ประมาณค่าพารามิเตอร์ด้วยวิธี Markov Chain Monte Carlo (MCMC) ผลสรุปได้ว่า โมเดลไอเอ็นโอ มีความเหมาะสมกับข้อมูลและให้ข้อมูลสารสนเทศการวินิจฉัยสำหรับแต่ละคุณลักษณะชุดของคุณลักษณะแต่ละบุคคลเป็นที่น่าพอใจ

De la Torre, J., (2011) ศึกษาและสรุปว่า G-DINA Model เป็นรูปแบบทั่วไปของโมเดล DINA MODEL แต่ สมมติฐาน มีความอ่อนคลายมากขึ้นในรูปแบบอิมิตัว G-DINA MODEL เทียบเท่ากับรุ่นทั่วไปอื่น ๆ สำหรับการวินิจฉัยทางพุทธิปัญญาตามฟังก์ชันลิงค์ ทางเลือก เมื่อข้อจำกัดที่เหมาะสมถูกนำไปใช้รูปแบบการวินิจฉัยองค์ความรู้ที่ใช้กันอย่างหลากหลาย (CDM) แสดงให้เห็นว่าเป็นรูปแบบทั่วไปของกรณีพิเศษ นอกเหนือจากการกำหนดแบบจำลองโมเดล G-DINA เป็นเฟรมเวิร์ก CDM ทั่วไป ยังมีส่วนประกอบสำหรับการประมาณค่าข้อสอบตามการออกแบบและเมทริกซ์น้ำหนักองค์ประกอบสำหรับเปรียบเทียบโมเดลข้อสอบจากการทดสอบ Wald เอกสารรายงานแสดงการประมาณค่าและการประยุกต์ใช้แบบจำลอง G-DINA MODEL เป็นกรอบการทำงานโดยใช้ข้อมูลจริงของ แบบจำลอง สรุปได้ด้วยการอภิปรายผลกระทบที่อาจเกิดขึ้นหลายประเด็นและเกี่ยวข้องกับกรอบการนำเสนอ

Li, Hongli (2011) ทำการศึกษาด้วยวิธีการวิเคราะห์แบบสอบวินิจฉัยทางพุทธิปัญญา โดยผู้สอบแต่ละคนจะได้รับโปรไฟล์ทักษะหลายมิติที่แสดงว่าเขา/เธอเป็นผู้เชี่ยวชาญหรือไม่ใช่ นาย

ของแต่ละทักษะ การวัดด้วยการทดสอบข้อเสนอนี้ การวินิจฉัยที่ละเอียดต่อการสอนและ การเรียนรู้สามารถมอบให้ครูและนักเรียนใช้ประโยชน์ได้ การศึกษานี้ตรวจสอบการวิเคราะห์ การวินิจฉัยองค์ความรู้ที่นำไปใช้กับแบบทดสอบการอ่านภาษาอังกฤษ การประเมินแบบเตอร์ Michigan (MELAB) Fusion Model (Hartz, 2002) ถูกใช้เพื่อประเมินโปรไฟล์ผู้เข้าสอบแต่ละคน การอ่าน Subskill ซึ่งเป็นพื้นฐานของการทดสอบ การอ่าน MELAB ด้วยข้อมูลที่รวบรวมจากหลาย แหล่งเช่น โปรโตคอลคิดตั้งและการจัดอันดับโดยผู้เชี่ยวชาญ Q-matrix ได้รับการพัฒนาเพื่อระบุ จำนวนย่อยที่จำเป็นสำหรับแต่ละรายการ Q-matrix นี้ผ่านการตรวจสอบด้วยแอปพลิเคชันของ Fusion Model ใช้ข้อมูลจากการทดสอบการอ่าน MELAB สี่ทักษะย่อยพบว่ารองรับการทดสอบ เช่น คำศัพท์ ไวยากรณ์ การแยกข้อมูลที่ชัดเจน การทำความเข้าใจข้อมูลโดยนัยโปรไฟล์การ เรียนรู้ทักษะที่ตรวจสอบได้ เป็นผลผลิตจากการวิเคราะห์วินิจฉัยองค์ความรู้ ประเด็นที่มีส่วนร่วม ในการวิเคราะห์ การวินิจฉัยองค์ความรู้การทดสอบการอ่านถูกกล่าวถึงและพื้นที่สำหรับการวิจัยในอนาคตก็สามารถนำมาใช้แนะนำได้

Hamdollah Ravand (2012) สำนวจความสามารถในการวินิจฉัยของแบบทดสอบ ความเข้าใจการอ่านซึ่งเป็นการทดสอบที่มีผลกระทบสำคัญสำหรับผู้สอบการผ่านมีผลประโยชน์ที่สำคัญ เช่น การจบการศึกษา โดยการประยุกต์ใช้โมเดล DINA การวิเคราะห์แบบดั้งเดิมของผลการทดสอบด้วยแบบจำลอง Item Response Theory (IRT) และทฤษฎีการทดสอบแบบคลาสสิก (CTT) ให้ค่าความสามารถที่มีมิติเดียว ซึ่งช่วยให้จัดอันดับนักเรียนตามความสามารถ การประเมินการวินิจฉัยทางปัญญา (CDA) มีวัตถุประสงค์เพื่อให้ข้อเสนอแนะในการวินิจฉัยโรคในรูปแบบที่ละเอียดการ รายงานความรอบรู้ด้านทักษะของผู้เรียน การศึกษาครั้งนี้ประยุกต์ใช้ DINA model (de la Torre & Douglas, 2004) ตรวจสอบระดับด้วยแบบสอบถามความเข้าใจในการอ่าน การทดสอบสามารถให้ข้อมูลที่ เป็นประโยชน์ในการวินิจฉัย จากการให้คะแนนโดยผู้เชี่ยวชาญ Q-matrix ที่มีห้าทักษะย่อยได้รับการ พัฒนาโดยใช้โมเดล DINA โปรไฟล์ทักษะความรอบรู้ของนักเรียนถูกสร้างขึ้นสอดคล้องสัมพันธ์กับการ วินิจฉัยสมรรถนะการอ่าน ในขณะเดียวกัน การศึกษาครั้งนี้แสดงให้เห็นอย่างชัดเจนถึงการประยุกต์ใช้ โมเดล DINA โดยใช้โปรแกรม R

H. Chen and J. Chen (2015, pp. 2-18) ได้ทำการศึกษาเกี่ยวกับการสำรวจ ความสัมพันธ์ทักษะความเข้าใจในการอ่านภาษาอังกฤษโดยใช้โมเดล G-DINA โดย Chen กล่าวว่า การประยุกต์ใช้โมเดลวินิจฉัยเชิงพุทธปัญญา (CDM) นับเป็นความท้าทายที่สำคัญซึ่งถูกกำหนดโดย การผันธรรมชาติของทักษะการอ่านเพื่อความเข้าใจ Chen ทำการวิเคราะห์โครงสร้างการอ่านของ PISA (PISA Reading Structure) เพื่อสำรวจความสัมพันธ์ของทักษะย่อยของการอ่านเพื่อความ เข้าใจ โดยการวิเคราะห์ข้อมูลการทดสอบความสามารถการอ่านเพื่อความเข้าใจของนักเรียนชาว อังกฤษระดับชั้นมัธยมศึกษาตอนต้นจำนวน 1,029 คน ใน 20 โปรแกรมสำหรับการประเมินการอ่าน

ภาษาอังกฤษของนักเรียนต่างชาติ โดยใช้โมเดล G-DINA การศึกษาครั้งนี้ได้จัดการสืบเสาะ 2 ครั้งในการสำรวจความสัมพันธ์ระหว่างทักษะความเข้าใจในการอ่าน 5 ทักษะย่อยที่กำหนดโดยผู้เชี่ยวชาญด้านภาษาอังกฤษ 6 คน กลุ่มทักษะ ประเภททักษะ และความสัมพันธ์ที่ขัดแย้งกันถูกค้นพบในการตรวจสอบความสัมพันธ์ระหว่างทักษะการอ่านเพื่อความเข้าใจที่เป็นส่วนหนึ่งของโครงสร้างความรู้ทางปัญญาของผู้สอบซึ่งให้ข้อมูลเชิงลึกเป็นลำดับขั้นการฝึกทักษะการอ่าน การตรวจสอบความสัมพันธ์ระหว่างทักษะการอ่านจับใจความในกระบวนการแก้ปัญหา ดำเนินการในระดับข้อสอบ การนิยามแบบแผนทักษะย่อยเป็นกลยุทธ์สำหรับข้อสอบแบบแก้ปัญหา การศึกษาครั้งนี้สาธิตว่าโมเดลชดเชย G-DINA รองรับคุณลักษณะของทักษะการอ่านเพื่อความเข้าใจและสามารถนำไปประยุกต์ใช้กับการทดสอบที่เกี่ยวข้องกับทักษะเชิงปฏิสัมพันธ์และลำดับขั้นที่สูง

Zhan, P. et.al (2015) The multidimensional testlet-effect cognitive diagnostic models การศึกษานี้เสนอ testlet-CDM สองชุด ชุดที่ 1 ปฏิบัติตามแนวทางการชดเชยและอีกวิธีหนึ่งปฏิบัติตามแนวทางที่ไม่ชดเชย: (1) CDM ผลการทดสอบหลายมิติชดเชย (C-MTECDM) ขึ้นอยู่กับการรวมกันของ LLM และ MTERM ในขณะที่ (2) CDM ผลการทดสอบหลายมิติที่ไม่ชดเชย (N-MTECDM) ขึ้นอยู่กับการรวมกันของ (logit) DINA model และ MTERM ตามลำดับ พารามิเตอร์โมเดลสามารถประมาณค่าได้โดยวิธี Bayesian ด้วยอัลกอริธึม Markov chain Monte Carlo (MCMC) ซึ่งถูกนำไปใช้กับโปรแกรม WinBUGS ซึ่งเป็น Free ware

Chalie Patarapichayatham and Akihito Kamata (2017, pp. 163-170) ทำการศึกษาความเหมาะสมของโมเดล bi-factor testlet ในการวิเคราะห์ข้อมูลรวมถึงมีความเหมาะสมในการประมาณค่าเมื่อแบบทดสอบเป็นแบบทดสอบตัวอย่างเช่นแบบทดสอบแบบบทความที่มีข้อคำถามที่เกี่ยวข้องภายใต้บทความนั้นๆ ความสามารถ การศึกษาในครั้งนี้มีวัตถุประสงค์เพื่อตรวจสอบคุณภาพการประมาณค่าความสามารถของโมเดล bi-factor testlet โดยมุ่งเน้นที่ความสัมพันธ์ระหว่างองค์ประกอบเฉพาะ (testlet factor) ผ่านการจำลองข้อมูลและข้อมูลจริง โมเดลหลักในการตรวจสอบคือ 1) โมเดล bi-factor testlet 2 พารามิเตอร์แบบดั้งเดิมที่ความสัมพันธ์ระหว่างองค์ประกอบเป็น “ศูนย์” 2) โมเดล bi-factor testlet 2 พารามิเตอร์แบบประยุกต์ที่ความสัมพันธ์ระหว่างองค์ประกอบ “ไม่เป็นศูนย์” ผลการศึกษาพบว่า ค่าความสัมพันธ์ระหว่างองค์ประกอบทั่วไป (general factors) และองค์ประกอบเฉพาะ (testlet factor) มีค่าเป็นบวกหมายความว่ามีความสัมพันธ์กันในเชิงบวกมีทิศทางเดียวกัน ในขณะที่ค่าความสัมพันธ์ระหว่างองค์ประกอบเฉพาะ (testlet factor) แต่ละองค์ประกอบเป็นลบทั้งหมด ค่าความสัมพันธ์ระหว่างองค์ประกอบเฉพาะ (testlet factor) และการประมาณค่าความสามารถมีค่าเป็นบวกหมายความว่ามีความสัมพันธ์กันในเชิงบวกมีทิศทางเดียวกัน นอกจากนี้ยังพบว่า ผลการวิเคราะห์ข้อมูลจริงมีความสอดคล้องกับผลการวิเคราะห์ด้วยข้อมูลจำลอง

Hamdollah Ravand (2018) ทำการศึกษาเปรียบเทียบประสิทธิภาพของโมเดลการวินิจฉัยทางปัญญา (CDM) 6 แบบเพื่อสำรวจความสัมพันธ์ระหว่างทักษะย่อย (Sub-skills) ในการทดสอบความเข้าใจในการอ่านภาษาอังกฤษ ด้วยเหตุนี้ จึงมีการวิเคราะห์คำตอบของผู้สอบประมาณ 21,642 คนต่อการทดสอบความเข้าใจในการอ่าน โมเดลทั้ง 6 โมเดลถูกเปรียบเทียบในแง่ของความพอดีของโมเดล (Model Fit) ทั้งในระดับการทดสอบและรายการ ความสอดคล้องและความแม่นยำในการจำแนกประเภท และสัดส่วนของแบบแผนความรู้ด้านทักษะ โมเดล DINA มีประสิทธิภาพในระดับดี

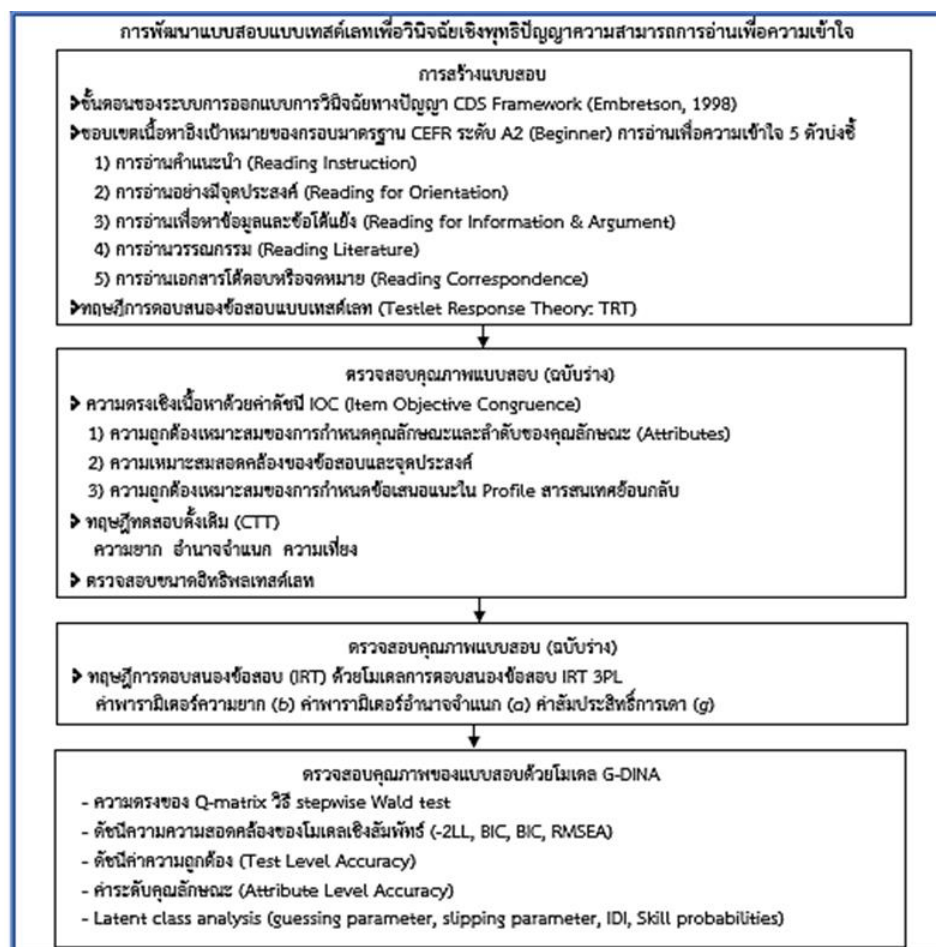
ตอนที่ 6 กรอบแนวคิดการวิจัย

การวิจัยในครั้งนี้เป็นการพัฒนาแบบสอบแบบทดสอบเพื่อวินิจฉัยเชิงพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 และออกแบบ Profile สารสนเทศผลการวินิจฉัยย้อนกลับความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษรายบุคคล โดยประยุกต์ใช้ G-DINA MODEL ผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องความสามารถด้านการอ่านเพื่อความเข้าใจภาษาอังกฤษ และเป้าหมายความสามารถผู้เรียนระดับเริ่มต้น (Beginner) ซึ่งเป็นผู้เรียนระดับการศึกษาภาคบังคับหรือระดับชั้นมัธยมศึกษาปีที่ 3 ตามกรอบอ้างอิงความสามารถทางภาษาอังกฤษ CEFR ระดับ A2 คือ เข้าใจบทอ่านสั้น ๆ และง่าย ที่ประกอบด้วยคำศัพท์ที่ผู้อ่านคุ้นเคย ประกอบด้วย 5 ตัวบ่งชี้ โดยมีคำอธิบายความสามารถด้านการอ่านภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ตามกรอบอ้างอิง CEFR ระดับ A2 ดังนี้ 1) การอ่านคำแนะนำ (Reading Instructions) สามารถเข้าใจคำแนะนำที่ใช้ภาษาต่างๆและขั้นตอนการใช้ เช่น วิธีใช้โทรศัพท์ วิธีกดเงินจากตู้ ATM หรือวิธีซื้อเครื่องดื่มจากตู้อัตโนมัติ 2) การอ่านอย่างมีจุดประสงค์ (Reading for Orientation) สามารถหาข้อมูลที่สำคัญในบทโฆษณาแผ่นพับ ใบปลิว หน้าเว็บ (web page) แคตตาล็อก ตารางเวลา ฯลฯ 3) การอ่านเพื่อหาข้อมูลและข้อโต้แย้ง (Reading Information & Argument) เป็นความสามารถเข้าใจประเด็นสำคัญ (main idea) ของเรื่องสั้น เหตุการณ์ หรือข่าวง่ายๆ เมื่อผู้อ่านมีความรู้ในเรื่องนั้นอยู่บ้าง เช่น ข่าวกีฬา หรือบุคคลที่มีชื่อเสียง 4) การอ่านวรรณกรรม (Reading Literature) สามารถเข้าใจประเด็นสำคัญของงานเขียนที่ใช้ภาษาต่างๆเกี่ยวกับชีวิตประจำวันที่มีสื่อหรือภาพประกอบ และ 5) การอ่านเอกสารโต้ตอบหรือจดหมาย (Reading Correspondence) สามารถเข้าใจข้อความสั้นๆ และง่ายจากเพื่อน เช่น อีเมล การสนทนาผ่านแอปพลิเคชัน แชท เป็นต้น ซึ่งผู้วิจัยจะนำตัวบ่งชี้ดังกล่าวไปใช้สำหรับการสร้างแบบสอบวินิจฉัยแบบทดสอบที่มีลักษณะเป็นแบบสอบปรนัย 4 ตัวเลือก มีการให้คะแนนแบบ 2 ค่า คือ ตอบถูกต้อง 1 ตอบผิดได้ 0 ดำเนินการตรวจสอบความตรงเชิงเนื้อหาและความถูกต้องของการกำหนดคุณลักษณะ (Attributes) ลำดับของคุณลักษณะและข้อเสนอแนะใน Profile สารสนเทศย้อนกลับ

โดยผู้เชี่ยวชาญด้านหลักสูตรและการสอนวิชาภาษาอังกฤษจำนวน 5 ท่าน พิจารณาค่าดัชนี Item Objective Congruence: IOC ทำการตรวจสอบคุณภาพแบบสอบวินิจัย (ฉบับร่าง) ตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) พิจารณาค่าความยาก อำนาจจำแนก และความเที่ยง ทำการตรวจสอบคุณภาพของแบบสอบวินิจัย (ฉบับร่าง 2) ด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ด้วยโมเดลการตอบสนองข้อสอบ IRT 3PL ค่าพารามิเตอร์ความยาก (b) ค่าพารามิเตอร์อำนาจจำแนก (a) ค่าสัมประสิทธิ์การเดา (g) ตรวจสอบคุณภาพของแบบสอบวินิจัยด้วยโมเดล G-DINA พิจารณาดัชนีค่าความถูกต้อง (Test Level Accuracy) ค่าระดับและคุณลักษณะ (Attribute Level Accuracy) และ Latent class analysis (LCA) และออกแบบ Profile สารสนเทศจากการวินิจัยแบบย้อนกลับที่ประกอบด้วย 3 ส่วน คือ รูปแบบความรอบรู้ของผู้สอบรายบุคคล คะแนนสอบ ผลการวินิจัยความรอบรู้ สารสนเทศเกี่ยวกับจุดแข็ง และข้อเสนอแนะเพื่อการพัฒนา



กรอบแนวคิดการวิจัย

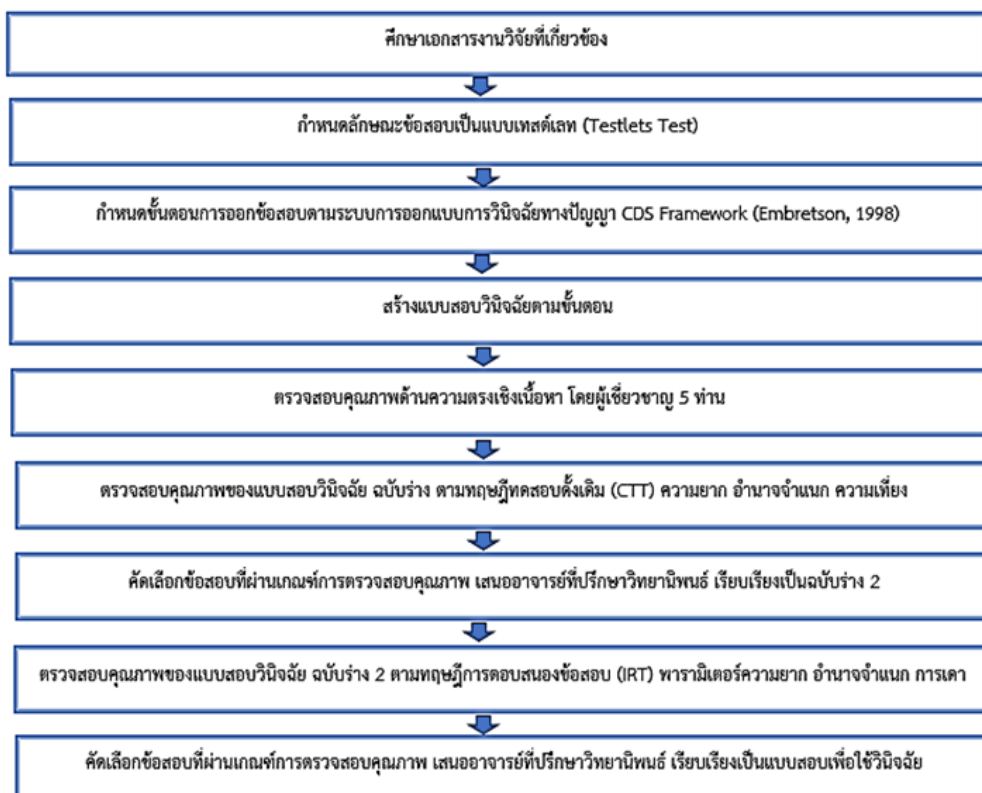


ภาพ 39 กรอบแนวคิดการวิจัย

บทที่ 3 วิธีดำเนินการวิจัย

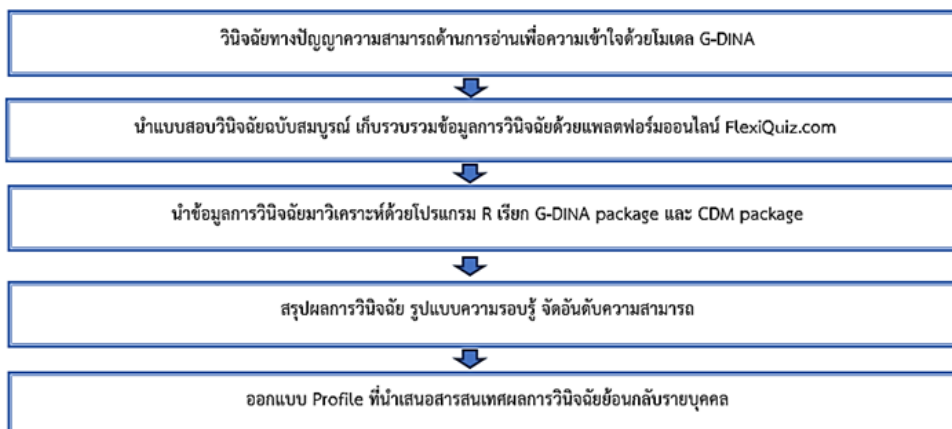
การศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ และทำการวินิจฉัยความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษ และให้สารสนเทศย้อนกลับแก่นักเรียน การดำเนินการวิจัยแบ่งออกเป็น 2 ระยะ ได้แก่ ระยะที่ 1 การสร้างและการตรวจสอบคุณภาพของแบบทดสอบแบบทดสอบเพื่อวินิจฉัยเชิงพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ ระยะที่ 2 การวินิจฉัยและพัฒนา Skill Profile สารสนเทศผลการวินิจฉัยย้อนกลับเกี่ยวกับความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3

ผู้วิจัยออกแบบการดำเนินการวิจัยออกเป็น 6 ส่วน คือ แหล่งข้อมูล ประชากรและกลุ่มตัวอย่าง การออกแบบการประเมินเพื่อวินิจฉัยเชิงพุทธิปัญญา เครื่องมือที่ใช้ในการวิจัย การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูล มีขั้นตอนและวิธีดำเนินการวิจัย ดังภาพ 40



ภาพ 40 แสดงขั้นตอนของการศึกษาในระยะเวลาที่ 1

ระยะที่ 2 การวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ



ภาพ 41 แสดงขั้นตอนของการศึกษาในระยะที่ 2

ระยะที่ 1 การสร้างและการตรวจสอบคุณภาพของแบบสอบแบบทดสอบเพื่อวินิจฉัยเชิงพุทธิปัญญา
ความสามารถการอ่านเพื่อความเข้าใจของนักเรียนชั้นมัธยมศึกษาปีที่ 3

ผู้วิจัยออกแบบการดำเนินการวิจัยออกเป็น 6 ส่วน คือ แหล่งข้อมูล ประชากรและกลุ่มตัวอย่าง ขอบเขตการวิจัย เครื่องมือที่ใช้ในการวิจัย การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูล มีขั้นตอนและวิธีดำเนินการวิจัย ดังนี้

1. แหล่งข้อมูล

แหล่งข้อมูลในขั้นตอนนี้ ได้แก่ ผู้เชี่ยวชาญ ทำการตรวจสอบคุณภาพของเครื่องมือวิจัย ได้แก่ แบบสอบวินิจฉัยด้านความตรงเชิงเนื้อหาของข้อคำถามกับคุณลักษณะที่มุ่งวัด (Q-matrix) จำนวน 5 ท่าน ประกอบด้วย

- ผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษและขอบเขตของเนื้อหาภาษาอังกฤษ ผู้ซึ่งมีคุณสมบัติเป็นครูชำนาญการขึ้นไป และมีวุฒิทางการศึกษาปริญญาตรีหรือปริญญาโทในสาขาวิชาเอกภาษาอังกฤษ จำนวน 3 ท่าน
- ผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษหรืออาจารย์มหาวิทยาลัยจำนวน 1 ท่าน
- ผู้เชี่ยวชาญด้านารวัดผล/การวิจัยประเมินผลทางการศึกษาเป็นอาจารย์มหาวิทยาลัยจำนวน 1 ท่าน

2. ประชากรและกลุ่มตัวอย่าง

2.1 ประชากร

ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนที่กำลังศึกษาอยู่ในระดับชั้นมัธยมศึกษาปีที่ 3 ในปีการศึกษา 2565 จากโรงเรียนมัธยมศึกษาในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเพชรบูรณ์ ทั้งหมด 39 แห่ง จำนวน 4,882 คน

ตาราง 18 แสดงจำนวนประชากรซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำแนกตามขนาดโรงเรียน

ขนาด	จำนวนโรงเรียน	ชื่อโรงเรียน	ประชากร
ใหญ่พิเศษ	6	วิทยานุกูลนารี	506
		เพชรพิทยาคม	420
		หล่มสักพิทยาคม	420
		หล่มเก่าพิทยาคม	408
		นิยมศิลป์อนุสรณ์	286
		หนองไผ่	391
รวมใหญ่พิเศษ			2,431
ใหญ่	3	ดงขุยพิทยาคม	174
		ศรีเทพประชาสรรค์	226
		บึงสามพันพิทยาคม	168
รวมใหญ่			568
กลาง	29	สวนกุหลาบวิทยาลัย เพชรบูรณ์	90
		เนินพิทยาคม	70
		เพชรบูรณ์วิทยา	74
		พัชรพิทยาคม	31
		ชนแดนพิทยาคม	103
		วังโป่งพิทยาคม	53
		วังโป่งศึกษา	52
		ตี่วพิทยาคม	65
		ผาเมืองพิทยาคม	16
		เมืองกลางพิทยาคม	32
		ศรีจันทร์วิทยาคมรัชมังคลาภิเษก	31

ขนาด	จำนวนโรงเรียน	ชื่อโรงเรียน	ประชากร
		กาญจนภิเษกวิทยาลัย เพชรบูรณ์	67
		เมืองрадวิทยาาคม	79
		น้ำหนาววิทยาาคม	56
		แคมป์สนวิทยาาคม	118
		ร่มเกล้าเขาค้อ	49
		พุขามครุฑมณีอุทิศ	25
		น้ำร้อนวิทยาาคม	62
		ชัยสมบูรณัวิทยาาคม	21
		โคกปรังวิทยาาคม	44
		วังใหญ่วิทยาาคม	31
		เมืองศรีเทพ	108
		นาสนุ่นวิทยาาคม	60
		เพชรละครวิทยา	147
		นาเฉลียงพิทยาาคม	115
		ท่าดัวงพิทยาาคม	70
		วังพิกุลพิทยาาคม	51
		ชัยบอนวิทยาาคม	104
		ศรีมงคลวิทยาาคม	36
รวมกลาง			1,860
เล็ก	1	ผาแดงวิทยาาคม	23
รวมเล็ก			23

รวมประชากรทั้งสิ้น 4,882 คน

หมายเหตุ : ขนาดโรงเรียนแบ่งตามเกณฑ์จำนวนนักเรียนที่กำหนดโดยสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (2564) ดังนี้ โรงเรียนขนาดใหญ่พิเศษ มีจำนวนนักเรียนตั้งแต่ 1,680 คนขึ้นไป ขนาดใหญ่ มีจำนวนนักเรียน 720 – 1,679 คน ขนาดกลาง มีจำนวนนักเรียนตั้งแต่ 120 – 719 คน และขนาดเล็กมีจำนวนนักเรียนตั้งแต่ 119 คนลงมา (สำนักพัฒนาระบบบริหารงานบุคคลสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน)

2.2 กลุ่มตัวอย่าง

2.2.1 กลุ่มตัวอย่างใช้ในการตรวจสอบคุณภาพของแบบสอบวินิจฉัยฉบับร่าง

กลุ่มตัวอย่างใช้ในการตรวจสอบคุณภาพของแบบสอบวินิจฉัยฉบับร่าง 1 ทำการตรวจสอบคุณภาพของข้อคำถามตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) พิจารณาค่าความยาก (P) อำนาจจำแนก (r) จำนวน 511 หน่วย ซึ่งได้มาจากการสุ่มอย่างง่าย โดยมีหลักคิดสำหรับการวิเคราะห์โดยใช้ทฤษฎีตอบสนองข้อสอบของ Geoffrey และ Andrej (2012, 10, อ้างถึงใน สุมาลี มีสกุล, 2558, หน้า 124) ที่กำหนดว่ากลุ่มตัวอย่างควรมีขนาดไม่ต่ำกว่า 500 คน มีแบบแผนการสุ่มตามขั้นตอนดังนี้

- 1) ผู้วิจัยทำการสุ่มอย่างง่าย (simple random sampling) ด้วยการจับสลากขนาดของโรงเรียน และการจับสลากได้โรงเรียนขนาดกลาง
- 2) ผู้วิจัยทำการสุ่มอย่างง่าย (simple random sampling) ด้วยการจับสลากชื่อโรงเรียนที่ถูกจัดอยู่ในขนาดกลางตามเกณฑ์ ในตาราง 18 เพื่อให้มีจำนวนตัวอย่างเพียงพอ ผู้วิจัยจึงยุติการจับสลากเมื่อรวมจำนวนกลุ่มตัวอย่างได้มากกว่า 500 หน่วย ดังนั้นจึงมีโรงเรียนที่ถูกสุ่มเพื่อเป็นกลุ่มตัวอย่างจำนวน 8 แห่งในตาราง 19

ตาราง 19 แสดงกลุ่มตัวอย่างที่ใช้ในการตรวจสอบคุณภาพของแบบสอบวินิจฉัยฉบับร่าง

ชื่อโรงเรียนที่สุ่ม	จำนวนนักเรียน	จำนวนตัวอย่าง
เมืองราดวิทยาคม	79	79
ตีวีวิทยาคม	65	65
วังโป่งพิทยาคม	67	67
เมืองกลางวิทยาคม	53	53
กาญจนานิกษะวิทยาลัย	32	32
นาเฉลียงพิทยาคม	115	115
น้ำหนาววิทยาคม	56	56
โคกปรังวิทยาคม	44	44
รวมทั้งสิ้น	511	511

2.2.2 กลุ่มตัวอย่างที่ใช้ในการตรวจสอบคุณภาพของแบบสอบวินิจฉัยฉบับร่าง 2

กลุ่มตัวอย่างจำนวน 1,015 หน่วยนี้ ผู้วิจัยนำมาใช้ใน 2 ขั้นตอน ได้แก่ 1) การตรวจสอบคุณภาพของข้อสอบรายข้อตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ด้วยโมเดลตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (IRT 3PL) ประมาณค่าพารามิเตอร์ความยาก

(b) พารามิเตอร์อำนาจจำแนก (a) พารามิเตอร์การเดา (c) และ 2) ตรวจสอบคุณภาพของแบบสอบฉบับร่างด้วยโมเดลวินิจฉัย G-DINA โดยเรียก CDM package และ G-DINA package โปรแกรมอาร์พีการณาความตรงของ Q-matrix ด้วยวิธี Stepwise Wald Test, ดัชนีความความสอดคล้องของโมเดลเชิงสัมพันธ์ (- 2LL, AIC, BIC, RMSEA) ,ดัชนีค่าความถูกต้อง, ค่าระดับคุณลักษณะ Latent class analysis (ค่าพารามิเตอร์การเดา , พารามิเตอร์ความสะอาด, ค่าดัชนีอำนาจจำแนก (IDI) ค่าความน่าจะเป็นของการตอบรู้แต่ละคุณลักษณะ (Attribute Probabilities)

กลุ่มตัวอย่างที่ใช้ในการตรวจสอบคุณภาพของแบบสอบวินิจฉัยฉบับร่าง 2 เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ประจำปีการศึกษา 2565 จำนวน 1,015 หน่วย ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน (multi-stage sampling โดยมีหลักคิดจากการใช้โมเดลการวินิจฉัยเชิงจำแนกของ รัปและเทมพลิน (Rupp & Templin, 2008: 251, อ้างถึงใน สุมาลี มีสกุล, 2558, หน้า 124) ที่กำหนดว่ากลุ่มตัวอย่างควรมีขนาดไม่ต่ำกว่า 200–300 คน ด้วยการสุ่มตัวอย่างแบบหลายขั้นตอน (Multistage Random Sampling) วิธีการสุ่มนี้เป็นวิธีที่มีข้อดีและเหมาะสมสำหรับนำมาใช้กับโมเดลการวินิจฉัยเชิงจำแนกเนื่องจากทำให้กลุ่มตัวอย่างที่นำมาวิเคราะห์ข้อมูลมีการกระจายตัวในทุกกลุ่มของข้อมูลที่ต้องการจำแนก ซึ่งส่งผลทำให้ค่าสถิติไคสแควร์ที่ใช้ในการทดสอบความถูกต้องเหมาะสมของข้อมูลเชิงประจักษ์กับโมเดลที่ใช้ในการจำแนกกลุ่มมีความน่าเชื่อถือ สอดคล้องกับการศึกษาของ Lei & Li (2016, p.4) ได้ตรวจสอบผลกระทบของขนาดตัวอย่างต่อการกู้คืนพารามิเตอร์ภายในการกำหนดพารามิเตอร์ CDMs ของ Log-linear สรุปว่า ข้อกำหนดขนาดกลุ่มตัวอย่าง (Sample Size) สำหรับความเหมาะสมของโมเดล (Model fit) ที่จะทำงานได้ดีของโมเดลเชิงวินิจฉัย CDMs พบว่าขนาดตัวอย่างประมาณ 500 (Choi et al., 2010, cited in Lei & Li ,2016, p.4) ก็เพียงพอที่จะประมาณค่าดังกล่าว เพื่อให้ได้ขนาดกลุ่มตัวอย่างที่เหมาะสมสำหรับการวิเคราะห์ข้อมูลทางสถิติและลดความเสี่ยงของค่า Missing Data ผู้วิจัยจึงกำหนดขนาดตัวอย่างขั้นต่ำเท่ากับ 600 หน่วย มีแบบแผนการสุ่มตามขั้นตอนดังนี้

1) ผู้วิจัยสุ่มเลือกโรงเรียนที่ใช้เก็บข้อมูลด้วยวิธีสุ่มแบบแบ่งชั้นภูมิ (stratified random sampling) โดยใช้ขนาดโรงเรียนเป็นชั้นภูมิในการสุ่ม พิจารณาจากการกำหนดขนาดสถานศึกษาตามจำนวนนักเรียน เป็น 4 ขนาด ดังนี้ ขนาดเล็ก มีจำนวนนักเรียนตั้งแต่ 119 คนลงมา ขนาดกลาง มีจำนวนนักเรียนตั้งแต่ 120 – 719 คน ขนาดใหญ่ มีจำนวนนักเรียนตั้งแต่ 720 – 1,679 คน ขนาดใหญ่พิเศษ มีจำนวนนักเรียนตั้งแต่ 1,680 คนขึ้นไป (สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2564)

2) สุ่มอย่างง่าย (simple random sampling) ด้วยการจับสลากซึ่งระบุชื่อโรงเรียนและขนาดของโรงเรียนแบบไม่ใส่คืน แต่เนื่องจากโรงเรียนขนาดเล็กในสังกัดสำนักงานเขตพื้นที่มัธยมศึกษาเพชรบูรณ์มีเพียง 1 แห่ง และเพื่อให้มีจำนวนตัวอย่างเพียงพอผู้วิจัยจึงกำหนดขนาดโรงเรียน 3 ขนาด

ได้แก่ ใหญ่พิเศษ ใหญ่ และกลาง ดังนั้นจึงมีโรงเรียนที่ถูกสุ่มเพื่อเป็นกลุ่มตัวอย่างในการตรวจสอบคุณภาพของแบบทดสอบ จำนวน 4 แห่ง ดังนี้ วิทยาลัยนุกุลนารี หล่อมเก่าพิทยาคม บึงสามพันพิทยาคม และชนแดนพิทยาคม

3) ทำการเลือกอย่างเจาะจงนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของแต่ละโรงเรียนที่ถูกสุ่มในข้อที่ 2 โดยพิจารณาตัดห้องเรียนพิเศษออกจำนวน 2 ห้องเรียน สำหรับโรงเรียนขนาดใหญ่พิเศษ เนื่องจากโรงเรียนแจ้งว่าไม่สะดวกให้เก็บข้อมูล ดังนั้นจึงได้กลุ่มตัวอย่างที่สุ่มในขั้นตอนนี้อาจมีจำนวน 1,015 หน่วย ดังแสดงในตาราง 20

ตาราง 20 แสดงผลการสุ่มตัวอย่างที่ใช้ในการตรวจสอบคุณภาพแบบสอบวินิจฉัยฉบับร่าง 2

ขนาด โรงเรียน	ชื่อโรงเรียนที่สุ่ม	จำนวน ห้องเรียน	ห้องที่ สุ่ม	จำนวน นักเรียน	จำนวน ตัวอย่างชั้น ต่ำ	จำนวน ตัวอย่างที่ ใช้
ใหญ่พิเศษ	วิทยาลัยนุกุลนารี	14	12	511		432
	หล่อมเก่าพิทยาคม	10	8	408		320
				919	358	752
ใหญ่	บึงสามพัน พิทยาคม	5	5	160	150	160
กลาง	ชนแดนพิทยาคม	3	3	103	92	103
รวมทั้งสิ้น		32	28	1,182	600	1,015

3. ขอบเขตการวิจัย

3.1 ขอบเขตด้านเนื้อหา

กระทรวงศึกษาธิการใช้ CEFR เป็นกรอบความคิดหลักในการกำหนดเป้าหมายการพัฒนาระดับความสามารถทางภาษาของผู้เรียนในระดับการศึกษาขั้นพื้นฐาน โดยกำหนดให้ ผู้สำเร็จการศึกษาภาคบังคับ (ม.3) มีระดับความสามารถทางภาษาเป็นผู้ใช้ภาษาขั้นพื้นฐาน และมีระดับความสามารถทางภาษาตามกรอบ CEFR ในระดับ A2 ในงานวิจัยนี้ศึกษาเกี่ยวกับความสามารถด้านการอ่าน ผู้วิจัยจึงสร้างแบบสอบวินิจฉัยทางพุทธิปัญญาโดยมีขอบเขตด้านเนื้อหาอิงเป้าหมายของกรอบอ้างอิงความสามารถทางภาษา CEFR ด้านการอ่านระดับ A2 คือ ผู้เรียนสามารถเข้าใจบทความที่สั้น ๆ ง่ายๆ เป็นเรื่องใกล้ตัวและใช้ภาษาที่พบบ่อยในชีวิตประจำวัน ประกอบด้วย

1) การอ่านคำแนะนำ (Reading instruction) สามารถเข้าใจคำแนะนำที่ใช้ภาษาง่ายๆและขั้นตอนการใช้ เช่น วิธีกดเงินจากตู้ ATM หรือวิธีซื้อเครื่องดื่มจากตู้อัตโนมัติ เครื่องจำหน่าย

ตัวสาธิต ข้อมูลเกี่ยวกับความปลอดภัย และการบอกทิศทาง เช่น Passive Voice ที่ใช้ใน โครงสร้างประโยคง่าย ๆ is/are + past participle คำสันธาน (conjunction) เช่น and/ but/ or/ before/ after/ because etc. ตัวเชื่อม (connective words) เช่น First, Second, Third, Fourth, Next, Then, Finally,... etc.

2) การอ่านอย่างมีจุดประสงค์ (Reading for orientation) สามารถค้นหาข้อมูล ที่ต้องการและเข้าใจประเด็นหลัก, ใจความสำคัญ (main idea) ของข้อมูลนั้นๆ เช่น ราคา ขนาด ของ โฆษณา หน้าเว็บ (web page) แผ่นพับ หน้าเว็บ (web page) แคตตาล็อก ตารางเวลา เช่น หัวข้อ เรื่อง ใจความสำคัญ รายละเอียดสนับสนุน คำถามเกี่ยวกับใจความสำคัญของเรื่อง เช่น ใคร ทำอะไร ที่ไหน เมื่อไร อย่างไร ทำไม ใช่หรือไม่ - Yes/No Question - Wh-Question etc.

3) การอ่านเพื่อหาข้อมูลและข้อโต้แย้ง (Reading for information & argument) สามารถเข้าใจประเด็นหลักในบทความสั้น ๆ จากหนังสือพิมพ์ นิตยสาร เกี่ยวกับบุคคลที่มีชื่อเสียงใน สังคม หรือเหตุการณ์ปัจจุบัน

4) การอ่านวรรณกรรม (READ LITERATURE) สามารถเข้าใจประเด็นสำคัญ (main idea) ของงานเขียน เช่น เรื่องสั้น (story) บทกลอน (poem) เกี่ยวกับชีวิตประจำวันซึ่งใช้ภาษาง่ายๆ และมีสื่อหรือรูปภาพประกอบ

5) การอ่านเอกสารโต้ตอบหรือจดหมาย (Reading correspondence) สามารถ เข้าใจข้อความง่ายๆ ในจดหมายอิเล็กทรอนิกส์ (e-mail) การสนทนาทางเว็บหรือแอปพลิเคชัน หรือ ข้อความจากเพื่อนหรือเพื่อนร่วมงานเกี่ยวกับเหตุการณ์ต่างๆในชีวิตประจำวัน เช่น การนัดหมายเวลา ไปรับประทานอาหารกลางวัน /อาหารเย็น We shall meet at six p.m. at BN café. /When shall we meet? เป็นต้น

ขอบเขตด้านเนื้อหาของสิ่งที่มุ่งวินิจฉัยในการวิจัยครั้งนี้ คือความสามารถการอ่านเพื่อความ เข้าใจ ประกอบด้วยคุณลักษณะย่อย 6 คุณลักษณะ ได้แก่ 1) มีความรู้คำศัพท์ 2) มีความสามารถในการระบุข้อมูลที่ต้องการ 3) มีความรู้ทางไวยากรณ์ 4) มีความสามารถในการระบุใจความสำคัญของ เรื่อง 5) มีความสามารถในการอนุมานและ 6) มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน

3.2 ขอบเขตด้านตัวแปร

การสร้างและหาคุณภาพแบบสอบวินิจฉัยทางพุทธิปัญญาในงานวิจัยนี้ มีการศึกษา ตัวแปรคุณภาพของแบบสอบ ดังนี้

1. ความตรงเชิงเนื้อหา (Content validity) ค่าดัชนีความสอดคล้อง (IOC) ระหว่างข้อสอบ (ฉบับร่าง) กับคุณลักษณะการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนชั้น มัธยมศึกษาปีที่ 3 ในแต่ละด้าน ความตรงของ Q-matrix ในเชิงประจักษ์เป็นการตรวจความถูกต้อง และเหมาะสมของ Q-matrix ซึ่งเป็นเมทริกซ์แสดงความสัมพันธ์ระหว่างลำดับที่ของข้อคำถามกับ

คุณลักษณะซึ่งเป็นความสามารถที่ใช้ในการหาคำตอบของข้อคำถามนั้น ตรวจสอบโดยผู้เชี่ยวชาญจำนวน 5 ท่าน และตรวจสอบอีกครั้งด้วย ด้วยวิธี Stepwise Wald Test ใช้โมเดลวินิจัย G-DINA เพื่อให้สอดคล้องกับความเป็นจริงมากที่สุด

2. ความยาก (P) และอำนาจจำแนก (r) ของข้อคำถามรายข้อ ที่ได้จากการคำนวณตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT)

3. ค่าพารามิเตอร์ความยาก (b -parameter) ค่าพารามิเตอร์ความสะเพร่า (a -parameter) พารามิเตอร์โอกาสการเดา (c -parameter) ที่ได้จากการคำนวณตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ด้วยโมเดลตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (IRT 3PL)

4. ดัชนีความสอดคล้องของโมเดลเชิงสัมพัทธ์ ได้แก่ ค่า $-2LL$ (-2Loglikelihood) ค่า AIC (Akaike's information criterion) ค่า BIC (Bayesian information criterion) ค่า RMSEA (Root Mean Square Error of Approximation)

5. ดัชนีค่าความถูกต้อง (Test Level Accuracy)

6. ค่าระดับคุณลักษณะ (Attribute Level Accuracy)

7. Latent class analysis (LCA) ได้แก่ ค่าพารามิเตอร์การเดา (Guessing Parameter) ค่าพารามิเตอร์ความสะเพร่า (Slipping Parameter) ค่าดัชนีอำนาจจำแนก (Generalization Discrimination Index: GDI) ค่าความน่าจะเป็นของการรอบรู้แต่ละคุณลักษณะ (Skill Probabilities)

4. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยแบ่งออกเป็น 2 ประเภท ตามกระบวนการวิจัย ดังนี้

4.1 เครื่องมือที่ใช้ในกระบวนการสร้างและตรวจสอบคุณภาพ

เครื่องมือที่ใช้ในกระบวนการสร้างและตรวจสอบคุณภาพของแบบสอบ สำหรับผู้เชี่ยวชาญ และเก็บข้อมูลจากกลุ่มตัวอย่าง จำนวน 3 ฉบับ ประกอบด้วย

ฉบับที่ 1 แบบประเมินความสัมพันธ์เชิงโครงสร้างของคุณลักษณะย่อย ความสามารถการอ่านเพื่อความเข้าใจ สำหรับผู้เชี่ยวชาญ โดยให้ผู้เชี่ยวชาญจำนวน 5 ท่าน พิจารณาว่าการที่นักเรียนจะมีคุณลักษณะที่พิจารณา นักเรียนต้องมีคุณลักษณะใดมาก่อนดังตัวอย่างในภาพ

(ตัวอย่าง 1) แบบบันทึกการพิจารณาความสัมพันธ์เชิงโครงสร้างของคุณลักษณะย่อยสำหรับผู้เชี่ยวชาญ

คำชี้แจง ให้ผู้เชี่ยวชาญพิจารณาลำดับความสัมพันธ์เชิงโครงสร้างของคุณลักษณะย่อยความสามารถการอ่านเพื่อความเข้าใจ โดยเขียนระบุลำดับความสัมพันธ์ที่ผ่านพิจารณาแล้วว่าการที่นักเรียนจะมีคุณลักษณะที่ A.....นักเรียนต้องมีคุณลักษณะใดมาก่อน โดยลำดับความสัมพันธ์เชิงโครงสร้างของคุณลักษณะที่ได้มาจากการทบทวนวรรณกรรมและสร้างตารางสังเคราะห์ มีดังนี้

- 1) มีความรู้คำศัพท์ (vocabulary knowledge)
- 2) มีความสามารถในการระบุข้อมูลที่ต้องการ (Identifying Explicit Information)
- 3) มีความรู้ทางไวยากรณ์ (Syntactic knowledge)
- 4) มีความสามารถในการระบุใจความสำคัญของเรื่อง (Generalization Main Idea)
- 5) มีความสามารถในการอนุมาน (Getting Inference)
- 6) มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน (Synthesizing information from Text)

ความสัมพันธ์เชิงโครงสร้างของคุณลักษณะโดยผู้เชี่ยวชาญ คนที่.....

ลำดับของคุณลักษณะ	ข้อคิดเห็นที่นำไปใช้ในการพิจารณา
1.....
2.....
3.....
4.....
5.....
6.....

ข้อเสนอแนะเพิ่มเติม.....

ภาพ 42 ตัวอย่างเครื่องมือที่ใช้เก็บรวบรวมข้อมูลในการวิจัยสำหรับผู้เชี่ยวชาญ

ตัวอย่าง 2 รายชื่อคุณลักษณะย่อย และคำอธิบายสำหรับผู้เชี่ยวชาญ

- **มีความรู้คำศัพท์** (vocabulary knowledge) พิจารณาได้จากความเข้าใจเกี่ยวกับความหมายคำศัพท์และวลี (กลุ่มคำ) ที่เกี่ยวข้อง และสามารถระบุคำและวลีที่มีความหมายถูกต้องโดยอาศัยพื้นความรู้เดิม หรืออาศัยการคาดเดาความหมายของคำศัพท์จากบริบท (context clues) ซึ่งช่วยให้ผู้เรียนสามารถทำความเข้าใจความหมายคำศัพท์ที่ยากหรือไม่คุ้นเคย และเข้าใจการใช้คำตามบริบทความหมายที่เป็นความรู้ด้านการเลือกใช้คำในประโยคให้เหมาะสม เช่น

(2) Once I sat on the sofa to watch a cartoon with my 8-year-old son and got completely dumbfounded when I realized a cartoon character was agonizing right in front of us, bleeding as if it was an action adult movie! Everybody knows that the cartoon “The Simpsons” doesn’t spread nice manners or behavior, but what about the “*kid stuff*” they display in the afternoon for young children? They shouldn’t provide some difficult things or danger for them. The next time you take a kid to the movies maybe a cartoon doesn’t make the best choice, unless you take the time to watch it and analyze its content first. Or you take the risk of exposing your child to messages s/he isn’t prepared to

Item 00. What is the meaning of word “*kid stuff*” in line 4 the second paragraph?

- 1) something dangerous for children
- 2) something easy or suitable for children
- 3) something save or comfortable for people
- 4) something valuable or important for people

ในข้อคำถามนี้หากผู้สอบมีคุณลักษณะความรู้คำศัพท์หมายถึงการรู้จักความหมายคำศัพท์ผู้สอบจะสามารถเลือกคำตอบที่ถูกต้องได้ในทันที แต่หากไม่รู้จอาจอาศัยการคาดเดาความหมายของ

คำศัพท์จากบริบทตรงส่วนที่ขีดเส้นใต้มีความหมายว่า พวกเขาไม่ควรนำเสนอการ์ตูนที่มีเนื้อหาเข้าใจยากและอันตรายแก่เด็กเล็กๆจึงมีตรงกับคำตอบในตัวเลือกที่ 2 something easy or suitable for children (บางสิ่งที่เหมาะสมสำหรับเด็ก ๆ)

จากนั้นผู้เชี่ยวชาญจะพิจารณาความถูกต้องเหมาะสมลงในแบบประเมินผู้วิจัยโดยมีเกณฑ์ในการพิจารณา คือ มีคะแนนเฉลี่ยตั้งแต่ 0.5 ขึ้นไป


ฉบับที่ 2 แบบประเมินความสอดคล้องของข้อความและคุณลักษณะที่มุ่งวัด(Q-matrix) สำหรับผู้เชี่ยวชาญ เพื่อประเมินความตรงเชิงเนื้อหา (Index of Item-Objective Congruence: IOC) มีตัวอย่างดังนี้

คำชี้แจง ในแต่ละข้อความ ให้ผู้เชี่ยวชาญพิจารณาประเมินทุกคุณลักษณะ ดังนี้

ประเมิน 1 เมื่อมั่นใจว่าข้อความวัดได้ตรงคุณลักษณะ

ประเมิน 0 เมื่อตัดสินใจไม่ได้ว่าข้อความวัดได้ตรงคุณลักษณะหรือไม่

ประเมิน -1 เมื่อมั่นใจว่าข้อความวัดได้ไม่ตรงคุณลักษณะ

Items	Attributes						ความคิดเห็นผู้เชี่ยวชาญ		
	A1	A2	A3	A4	A5	A6	1	0	-1
Which sentence refers to this picture?  <ol style="list-style-type: none"> 1) Type in the code or press the button for your item. 2) Insert the correct amount of money for the item. 3) Swipe your credit card if it's a newer machine. ✓ 4) Check the coin chute for any change. 	1	1	0	0	1	0			
คำอธิบาย : คุณลักษณะที่ใช้ในการตอบข้อสอบมีดังนี้ เข้าใจความหมายคำศัพท์ กลุ่มคำ และประโยคโดยอาศัยพื้นฐานความรู้เดิม มีความสามารถในการระบุรายละเอียดที่ต้องการซึ่งปรากฏชัดเจนในบทอ่าน และสามารถเชื่อมโยงไปถึงข้อมูลภายนอกซึ่งไม่ได้กล่าวไว้ในบทอ่าน ข้อคิดเห็น/ข้อเสนอแนะเพิ่มเติม.....									

ภาพ 43 ตัวอย่างแบบประเมินความสอดคล้องของข้อความและคุณลักษณะสำหรับผู้เชี่ยวชาญ


ฉบับที่ 3 แบบสอบแบบทดสอบเพื่อวินิจฉัยเชิงพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 (ฉบับร่าง) สำหรับกลุ่มตัวอย่าง จัดทำใน Google form เพื่อเก็บข้อมูลแบบออนไลน์ ดังตัวอย่างภาพ 44

Reading 1: items 1 - 10 * 1 คะแนน

1. Rearrange these sentences according to the correct order of the pictures

(A) Check the coin chute for any change.
 (B) Swipe your credit card if it's a newer machine.
 (C) Insert the correct amount of money for the item.
 (D) Type in the code or press the button for your item.
 (E) Find the price and code underneath the item you wish to buy.

A vending machine is an automated machine that provides items such as snacks, beverages, cigarettes, and lottery tickets to consumers after cash, a credit card, or other forms of payment are inserted into the machine or otherwise made. The first modern vending machines were developed in England in the early 1880s and dispensed postcards. Vending machines exist in many countries and, in more recent times, specialized vending machines that provide common products compared to traditional vending machine items have been created.



Picture 1 Picture 2 Picture 3 Picture 4 Picture 5

C-B-D-A-E
 E-C-B-D-A
 B-D-C-E-A
 E-B-A-C-D

ภาพ 44 ตัวอย่างแบบสอบวินิจฉัยฉบับร่าง

5. การสร้างและการหาคุณภาพเครื่องมือวิจัย

เครื่องมือที่ใช้ในการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจ จำนวน 1 ฉบับ ได้แก่ แบบสอบแบบทดสอบเพื่อวินิจฉัยเชิงพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 เพื่อวินิจฉัยและจัดอันดับความสามารถการอ่านเพื่อความเข้าใจ มีขั้นตอนในการสร้างดังนี้

5.1 ขั้นตอนการสร้างแบบสอบวินิจฉัยทางพุทธิปัญญา

ในการสร้างแบบสอบวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจ ผู้วิจัยอิงกรอบแนวคิดของหลักการออกแบบการทดสอบเชิงวินิจฉัยของเอ็มเบรทสัน หรือ CDS framework ซึ่งมีจุดเด่นที่ให้ความตรงเชิงโครงสร้างของคุณลักษณะที่ทำการวัดสามารถทำความเข้าใจได้ดีในแง่ของการพัฒนาข้อคำถาม อีกทั้งยังสามารถแปลผลคะแนนในเชิงวินิจฉัยได้ง่ายกว่าวิธีการอื่น (สุมาลี มีสกุล, 2558, หน้า 16) เนื่องมาจากข้อคำถามถูกพัฒนาขึ้นด้วยกระบวนการ

ออกแบบเชิงวินิจฉัยโดยตรงสอดคล้องกับ กุลพร พูนสวัสดิ์ (2556) กล่าวว่า การสร้าง Q-matrix ให้สอดคล้องกับข้อสอบที่มีอยู่ และนำข้อมูลมาวินิจฉัย ซึ่งข้อสอบดังกล่าวไม่มีจุดประสงค์เพื่อวินิจฉัย อาจทำให้เกิดปัญหา Flat profile คือ skill profile แบบ 1111 (ผู้สอบชำนาญทุกทักษะ) หรือ 0000 (ผู้สอบไม่ชำนาญทุกทักษะ) และไม่สามารถระบุจุดแข็งหรือจุดอ่อนของผู้สอบได้ ดังนั้นผู้วิจัยจึงสร้างแบบสอบวินิจฉัยโดยอิง CDS framework (Embretson's Cognitive Design System Approach: CDS framework) (Rupp, Templin and Henson, 2010, p. 15-17, อ้างถึงใน วราจคณา ก้อมน้อย, 2563, หน้า 95-97) มีขั้นตอนการออกแบบดังนี้

1) กำหนดจุดมุ่งหมายของการประเมิน

การกำหนดจุดมุ่งหมายของการประเมินให้ชัดเจน มี 2 ขั้นตอนย่อย ดังนี้
 ขั้นตอนที่แรก ผู้วิจัยกำหนดจุดมุ่งหมายของการประเมินให้ชัดเจน พร้อมทั้งระบุถึงโครงสร้างอื่นที่เกี่ยวข้องในเชิงทฤษฎีกับสิ่งที่มุ่งวัด (Nomothetic span) โดยแบบสอบวินิจฉัยทางพุทธิปัญญา ความสามารถในการอ่านเพื่อความเข้าใจ ในงานวิจัยนี้ มีจุดมุ่งหมายเพื่อวินิจฉัยความรอบรู้ (Mastery) หรือไม่รอบรู้ (Non-mastery) ในคุณลักษณะย่อยของความสามารถในการอ่านเพื่อความเข้าใจ ซึ่งจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องพบว่า ความสามารถในการอ่านเพื่อความเข้าใจ เป็นตัวแปรที่สำคัญในการทำนายความสำเร็จในการเรียนภาษาอังกฤษ โดยกลุ่มเป้าหมายในการใช้แบบสอบเชิงวินิจฉัยนี้ คือนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ซึ่งเป็นระดับการศึกษาภาคบังคับ และตามเป้าหมายกรอบอ้างอิงความสามารถทางภาษา CEFR ในระดับ beginner หรือ A2 ซึ่งเป็นช่วงขั้นที่สำคัญในการพัฒนาความสามารถทางภาษา สอดคล้องกับแนวคิดรูปแบบการอ่านภาษาอังกฤษเป็นภาษาที่สองของ Coady (ถิรวัฒน์ ต้นทนิส, 2550, หน้า 128) กล่าวว่า ความเข้าใจในการอ่านจะเกิดขึ้นเนื่องจากการปฏิสัมพันธ์ของปัจจัย 3 ประการคือ ความสามารถในการประมวลข้อมูล (Conceptual Abilities) กระบวนการในการใช้กลยุทธ์ (Process strategies) และพื้นความรู้เดิม (Background knowledge) ปัจจัย 3 ประการนี้มีความสำคัญต่อความเข้าใจในการอ่านเป็นอย่างมาก

2) ระบุคุณลักษณะที่ต้องการวินิจฉัย

ผู้วิจัยระบุคุณลักษณะย่อยของสิ่งที่ต้องการวินิจฉัย โดยมีการดำเนินการดังนี้
 1) ศึกษาเอกสารแนวคิดที่เกี่ยวข้องกับคุณลักษณะของความสามารถในการอ่านเพื่อความเข้าใจ 2) กำหนดคุณลักษณะจากการสังเคราะห์งานวิจัยของนักวิชาการด้านภาษาศาสตร์ โดยมุ่งเน้นการพัฒนาความสามารถในการอ่านภาษาอังกฤษเป็นภาษาที่สอง (Munby's taxonomy ,1978; Heaton's taxonomy, 1991; Jang, 2005; Gao, 2006; DELTA, 2013; Ravand, 2015; Javidanmehr, 2019, โตโพนิกา และรวัน, 2021) คุณลักษณะความสามารถในการอ่านเพื่อความเข้าใจที่ได้จากการทบทวนวรรณกรรมและสร้างตารางสังเคราะห์โดยผู้วิจัย ประกอบด้วย 6 คุณลักษณะ ดังนี้

๔ มีความรู้คำศัพท์ (Vocabulary Knowledge) พิจารณาจากความเข้าใจเกี่ยวกับความหมายคำศัพท์ และกลุ่มคำโดยอาศัยพื้นฐานความรู้เดิม หรือเข้าใจความหมายของคำศัพท์ที่ยากหรือไม่คุ้นเคยโดยอาศัยการคาดเดาความหมายคำศัพท์จากบริบทหรือคำบ่งชี้ที่พบในข้อความ (Context Clue) ประกอบด้วยการนิยามความหมาย การกล่าวซ้ำ การเปรียบเทียบความเหมือนหรือความต่าง และการยกตัวอย่าง

๕ มีความสามารถในการระบุข้อมูลที่ต้องการ (Identifying Explicit Information) พิจารณาได้จากทักษะในการค้นหารายละเอียดหรือข้อมูลเฉพาะที่ต้องการซึ่งปรากฏชัดเจนในบทอ่าน หลังจากที่ได้อ่านแบบผ่าน (Scan) และอ่านแบบพินิจพิเคราะห์ (Skim) ทำให้ทราบว่าข้อมูลอยู่ที่ใดของเรื่อง เช่น จำนวน ราคา วันเวลา สถานที่ ซึ่งข้อมูลดังกล่าวจะปรากฏอย่างชัดเจน ผู้อ่านสามารถเข้าใจได้ โดยมีต้องตีความ

๖ มีความรู้ทางไวยากรณ์ (Syntactic Knowledge) พิจารณาได้จากความเข้าใจเกี่ยวกับวากยสัมพันธ์ซึ่งหมายถึงการศึกษาความสัมพันธ์ของคำในประโยคหรือแบบแผนองค์ประกอบของประโยคในภาษาอังกฤษ และความสัมพันธ์ระหว่างประโยคและสามารถวางตำแหน่งของคำหรือกลุ่มคำประเภทต่าง ๆ ในประโยคได้อย่างถูกต้องตามโครงสร้างทางไวยากรณ์

๗ มีความสามารถในการระบุใจความสำคัญของเรื่อง (Generalization Main Idea) พิจารณาได้จากการทำความเข้าใจเพื่อค้นหาหัวเรื่อง (Topic) นำมารวมกับข้อความที่ถูกควบคุมให้กล่าวถึงสิ่งใดสิ่งหนึ่ง (Controlling Idea) และลงข้อสรุปเป็นใจความสำคัญ หรือแนวคิดหลักที่อยู่ในบทอ่านซึ่งเป็นส่วนที่ครอบคลุมย่อหน้านั้น ๆ หรือเนื้อหาทั้งหมด

๘ มีความสามารถในการอนุมาน (Getting Inference) พิจารณาได้จากความสามารถในการตีความจากเรื่องที่อ่าน โดยใช้ความรู้และประสบการณ์ตรงของผู้อ่าน เพื่อทำความเข้าใจข้อมูลที่ไม่ได้ถูกระบุไว้อย่างชัดเจนซึ่งเป็นจุดมุ่งหมาย ทักษะคิดของผู้เขียน หรือข้อโต้แย้งโดยอาศัยเบาะแสร่องรอยในเนื้อความซึ่งเป็นคำหรือประโยคชี้แนะ หรือทำการตรวจสอบเท็จจริงทั้งหมดของเรื่องที่อ่าน เพื่อพิจารณาหรือตัดสินว่าสิ่งที่อ่านทั้งหมดนั้นหมายความว่าความถึงสิ่งใดและสามารถเชื่อมโยงไปถึงข้อมูลภายนอกซึ่งไม่ได้กล่าวไว้ในบทอ่าน

๙ มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน (Synthesizing Information from Text) พิจารณาได้จากความสามารถในการแยกประโยคแนวคิดหลัก หรือใจความสำคัญออกจากประโยครองหรือข้อความซึ่งเป็นรายละเอียดเสริมจากบทอ่าน นำประเด็นหลักที่รวบรวมได้ มาเชื่อมโยงเชื่อมต่อกัน โดยพิจารณาความเกี่ยวข้อง ความเป็นเหตุเป็นผล ทำการจัดลำดับข้อความ และทำการลงข้อสรุปหรือเรียบเรียงประโยคขึ้นใหม่โดยคงความหมายเดิม

3) พิจารณาความสัมพันธ์เชิงโครงสร้างของคุณลักษณะ

ผู้วิจัยนำคุณลักษณะย่อยของความสามารถการอ่านเพื่อความเข้าใจเสนอต่อผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษและด้านการวัดผล รวมจำนวน 5 ท่าน เพื่อพิจารณาความสัมพันธ์เชิงโครงสร้างของคุณลักษณะ โดยให้ผู้เชี่ยวชาญพิจารณาว่าการที่นักเรียนจะมีคุณลักษณะที่พิจารณา นักเรียนต้องมีคุณลักษณะใดมาก่อน โดยมีเกณฑ์ในการพิจารณา คือ มีคะแนนเฉลี่ยตั้งแต่ 0.5 ขึ้นไป ซึ่งหากพบว่า คุณลักษณะย่อยทั้งหมดคุณลักษณะมีความสัมพันธ์เชิงโครงสร้าง แบบทดสอบที่สร้างขึ้นต้องมีลักษณะเป็นข้อคำถามที่สามารถวัดได้หลายคุณลักษณะ


4) กำหนดเมตริกซ์คิว (Q-matrix) และสร้างข้อคำถาม

ผู้วิจัยกำหนดเมตริกซ์คิว (Q-matrix) ซึ่งเป็นเมตริกซ์ที่แสดงความสัมพันธ์ระหว่างข้อคำถาม (items) กับคุณลักษณะ (Attribute) ซึ่งเป็นความสามารถที่ใช้ในการหาคำตอบของข้อคำถามนั้น โดยกำหนดให้มีค่าคะแนนเท่ากับ 1 เมื่อนักเรียนต้องใช้คุณลักษณะนั้นในการหาคำตอบ และกำหนดให้มีค่าเท่ากับ 0 เมื่อนักเรียนไม่ใช้คุณลักษณะนั้นในการหาคำตอบ

ผู้วิจัยดำเนินการสร้างข้อสอบให้ครอบคลุมตามกรอบการอ่านเพื่อความเข้าใจภาษาอังกฤษ โดยสร้างข้อคำถามมุ่งวัดแต่ละคุณลักษณะด้วยข้อสอบแบบเทสต์เลท (Testlet test) โดยไม่อิงความสัมพันธ์เชิงโครงสร้าง เพื่อเป็นการตรวจสอบยืนยันความสัมพันธ์เชิงโครงสร้างของคุณลักษณะ ในการศึกษาครั้งนี้กำหนดชุดข้อสอบเทสต์เลทใช้สิ่งเร้าร่วมกันเป็นบทอ่าน (passage) จำนวน 5 บทอ่าน ซึ่งจำนวนบทอ่านในแบบสอบอิงเป้าหมายกรอบอ้างอิงความสามารถทางภาษา CEFR ระดับ A2 ใน 5 เรื่อง ได้แก่ 1) การอ่านคำแนะนำ 2) การอ่านข้อมูลเบื้องต้น 3) การอ่านข้อมูลและข้อโต้แย้ง 4) การอ่านวรรณกรรม และ 5) การอ่านเอกสารโต้ตอบหรือจดหมาย ดังนั้นแบบสอบจึงมีจำนวนชุดข้อสอบเทสต์เลท 5 เทสต์เลท แต่ละเทสต์เลทมีจำนวนข้อคำถามไม่เท่ากัน รวมข้อคำถามทั้งสิ้น 46 ข้อ เป็นแบบสอบเลือกตอบ 4 ตัวเลือก ที่มีการตรวจให้คะแนนแบบสองค่า (dichotomous) โดยมุ่งวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจ 6 คุณลักษณะ คือ 1) มีความรู้คำศัพท์ 2) มีความสามารถในการระบุข้อมูลที่ต้องการ 3) มีความรู้ทางไวยากรณ์ 4) มีความสามารถในการระบุใจความสำคัญของเรื่อง 5) มีความสามารถในการอนุมานและ 6) มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่านโดยข้อสอบ 1 ข้อ สามารถวัดความสามารถ ได้ 1 ด้าน แต่วัดได้มากกว่า 1 คุณลักษณะความสามารถด้านนั้น ๆ

5) ตรวจสอบความตรงของเมตริกซ์คิว (Q-matrix)

5.1 ผู้วิจัยนำแบบสอบที่สร้างขึ้นเสนอต่อผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษและด้านการวัดผล รวมจำนวน 5 ท่าน พร้อมแบบประเมินความสอดคล้องของข้อคำถามและคุณลักษณะที่มุ่งวัด (แบบปิด) เพื่อประเมินความตรงเชิงเนื้อหาโดยพิจารณาจากดัชนีความสอดคล้องระหว่างข้อคำถามกับคุณลักษณะที่มุ่งวัด

Items	Attributes						ความคิดเห็นผู้เชี่ยวชาญ		
	A1	A2	A3	A4	A5	A6	1	0	-1
Which sentence refers to this picture? 	1	1	0	0	1	0	✓		
1) Type in the code or press the button for your item. 2) Insert the correct amount of money for the item. 3) Swipe your credit card if it's a newer machine. ✓ 4) Check the coin chute for any change.									
คำอธิบาย : คุณลักษณะที่ใช้ในการตอบข้อสอบมีดังนี้ เข้าใจความหมายคำศัพท์ กลุ่มคำ และประโยคโดยอาศัยพื้นฐานความรู้เดิม มีความสามารถในการระบุละเอียดสิ่งที่ต้องการซึ่งปรากฏชัดเจนในบทอ่าน และสามารถเชื่อมโยงไปถึงข้อมูลภายนอกซึ่งไม่ได้กล่าวไว้ในบทอ่าน ข้อคิดเห็น/ข้อเสนอแนะเพิ่มเติม.....									

ภาพ 45 ตัวอย่างผลการประเมินของผู้เชี่ยวชาญคนที่ ก.

จากภาพ 45 ตัวอย่างผลการประเมิน IOC มีหมายความว่า ผู้เชี่ยวชาญคนที่ ก ประเมินข้อคำถามดังนี้

-มั่นใจว่าข้อคำถามวัดคุณลักษณะที่ 1, 2 และ 5

-มั่นใจว่าข้อคำถามไม่วัดคุณลักษณะที่ 3 และ 6

5.2 ผู้วิจัยนำผลการประเมินของผู้เชี่ยวชาญทั้ง 5 ท่าน มา คำนวณหาค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับคุณลักษณะที่มุ่งวัด (IOC) โดยเกณฑ์การผ่านความตรงเชิงเนื้อหาของข้อคำถาม คือ $IOC \geq 0.5$ คัดเลือกข้อคำถามที่ผ่านเกณฑ์ โดยปรับแก้ไขคำและรูปประโยคตามข้อเสนอแนะของผู้เชี่ยวชาญ จากนั้นเรียบเรียงเป็นแบบสอบฉบับร่าง เพื่อใช้เก็บข้อมูลในขั้นตอนต่อไป

6) ตรวจสอบคุณภาพแบบสอบวินิจฉัย

6.1 นำร่างแบบสอบวินิจฉัยที่ผ่านการคัดเลือก ไปทดลองใช้ (try out) กับกลุ่มตัวอย่าง ซึ่งเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 511 หน่วย ซึ่งได้มาจากการสุ่มอย่างง่าย เพื่อนำไปใช้ตรวจสอบคุณภาพรายข้อของแบบสอบวินิจฉัย โดยการหาค่าความยาก และค่าอำนาจจำแนก ด้วย Excel

6.2 คัดเลือกข้อสอบที่ผ่านเกณฑ์ค่าความยาก ($P = 0.20$ ถึง 0.79) ตามเกณฑ์การแปลความหมาย (ศิริชัย กาญจนวาสี, 2552, อ้างถึงใน เกตรินทร์ สมราช, 2563, หน้า 76-77) ดังนี้

0.80-1.00 ข้อสอบง่าย ควรตัดทิ้งหรือปรับปรุง

0.60-0.79 ข้อสอบค่อนข้างง่าย สามารถเก็บไว้ใช้ได้ต่อไป

0.40-0.59 ข้อสอบยากพอเหมาะ สามารถเก็บไว้ใช้ได้ต่อไป

0.20-0.39 ข้อสอบค่อนข้างยาก สามารถเก็บไว้ใช้ได้ต่อไป

0.00-0.19 ข้อสอบยาก ควรตัดทิ้งหรือปรับปรุง

6.3 คัดเลือกข้อสอบที่ผ่านเกณฑ์ค่าอำนาจจำแนก ($.20 \leq r \leq 1.00$) หากข้อสอบที่มีค่าอำนาจจำแนกเป็นลบหรือเป็นศูนย์ ควรตัดทิ้งเพราะไม่สามารถแยกคนเก่งและคนอ่อนออกจากกันได้ (พิชิต ฤทธิจรุญ, 2550) เกณฑ์พิจารณาอำนาจจำแนก ดังนี้

$.40 \leq r \leq 1.00$ หมายความว่า จำแนกคนเก่งกับคนอ่อนออกจากกันได้ดีมาก

$.30 \leq r \leq .39$ หมายความว่า จำแนกคนเก่งกับคนอ่อนออกจากกันได้

$.20 \leq r \leq .29$ หมายความว่า จำแนกคนเก่งกับคนอ่อนออกจากกันได้ปานกลาง

$-1.00 \leq r \leq .19$ หมายความว่า ไม่สามารถจำแนกคนเก่งกับคนอ่อนออกจากกันได้

นำเสนอต่ออาจารย์ที่ปรึกษา และปรับปรุงตามคำแนะนำของอาจารย์ นำไปเรียบเรียงเป็นแบบสอบวินิจฉัยฉบับร่าง 2 จากนั้นคำนวณหาค่าความเชื่อมั่นทั้งฉบับ ใช้สูตร KR-20 ของ Kuder-Richardson ด้วย Excel โดยแบบทดสอบที่ดีควรมีค่าความเชื่อมั่นมากกว่า 0.70

6.4 ผู้วิจัยนำแบบทดสอบฉบับร่างไปทดลองใช้ (try out) ครั้งที่ 2 กับนักเรียนกลุ่มตัวอย่างจำนวน 1,015 หน่วย ที่ได้มาจากการสุ่มแบบหลายชั้นตอน เพื่อตรวจสอบคุณภาพของแบบทดสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) ด้วยโมเดลตอบสนองข้อสอบโลจิสติกแบบ 3 พารามิเตอร์ ประกอบด้วย พารามิเตอร์ b พารามิเตอร์ a และพารามิเตอร์ c และค่าสารสนเทศของแบบสอบ

b_i คือ ค่าความยากของข้อสอบข้อที่ i (Item Discrimination) มีพิสัยอยู่ระหว่าง $-\infty$ ถึง ∞ แต่ในทางปฏิบัติข้อสอบที่มีค่า b_i อยู่ระหว่าง -3.00 ถึง $+3.00$ ค่า $b = +3.00$ แสดงว่ายากที่สุด ค่า $b = 0$ แสดงว่ายากปานกลาง

a_i คือ ค่าความอำนาจจำแนกของข้อสอบข้อที่ i (Item Discrimination) มีพิสัยอยู่ระหว่าง $-\infty$ ถึง ∞ ค่า a ที่เป็นลบเป็นสิ่งที่ไม่พึงปรารถนา ในทางปฏิบัติข้อสอบจะมีค่า a_i อยู่ระหว่าง 0.5 ถึง 2.5 ค่า a สูง แสดงว่าข้อสอบข้อนั้นมี slope ที่ชัน จึงจำแนกผู้สอบที่มีความสามารถแตกต่างกันได้ดี

c_i คือ ค่าการเดาของข้อสอบข้อที่ i (Pseudo - Chance Level หรือ Guessing Parameter) เป็นความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำมากจะทำข้อสอบข้อที่ i ได้ถูก และค่า c_i ควรมีค่าไม่เกิน 0.30

ผู้วิจัยพิจารณาคัดเลือกข้อสอบตามเกณฑ์ของศิริชัย กาญจนวาสี (2550, หน้า 56, อ้างถึงใน แคนทรียา แสงดาวเทียน, 2560, หน้า 136) ที่กล่าวว่า ค่าอำนาจจำแนกควรมีค่าอยู่ระหว่าง +0.50 ถึง +2.50 ค่าความยากควรมีค่าอยู่ระหว่าง -2.50 ถึง +2.50 และค่าการเดาควรมีค่าไม่เกิน 0.30

6.4 คัดข้อคำถามที่ไม่ผ่านเกณฑ์ออก เสนอต่ออาจารย์ที่ปรึกษาวิทยานิพนธ์ปรับปรุงตามคำแนะนำของอาจารย์ และจัดทำเป็นแบบสอบวินิจฉัยฉบับสมบูรณ์ เพื่อนำไปใช้ในขั้นตอนต่อไป

การพิจารณาอิทธิพลทดสอบ

การสร้างและตรวจสอบคุณภาพแบบสอบในระยะที่ 1 โดยอิงแนวคิดการออกแบบการประเมินเชิงวินิจฉัยทางพุทธิปัญญา แต่เนื่องจากการศึกษาครั้งนี้ ผู้วิจัยพัฒนาแบบสอบแบบทดสอบแบบทดสอบโดยข้อคำถามมีลักษณะทดสอบ ซึ่งมีประเด็นเกี่ยวข้องที่น่าสนใจควรค่าการนำเสนอข้อมูลเพิ่มเติมคือ อิทธิพลของทดสอบ (Testlet Effect) เนื่องจากทัศนะของนักการศึกษาบางท่านกล่าวว่า หากอิทธิพลของทดสอบมีขนาดใหญ่อาจส่งผลกระทบต่อในการประมาณค่าพารามิเตอร์ข้อสอบ และในส่วนของปัญหาด้านความไม่เป็นอิสระกันของข้อคำถาม สุชาติ สกลกิจรุ่งโรจน์ (2558, หน้า 27-28) กล่าวว่า ความเป็นอิสระของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบนั้น เป็นความอิสระของข้อคำถามที่เกิดขึ้นภายหลังการควบคุมพารามิเตอร์บุคคล และพารามิเตอร์ข้อสอบที่มีอิทธิพลโดยตรงต่อการตอบข้อคำถามในโมเดลแล้ว สอดคล้องกับพิสิฐ พิณจสกุล (2563, หน้า 81) กล่าวไว้ว่า หากลดจำนวนข้อสอบในแต่ละทดสอบลงมา ก็จะสามารถลดการขึ้นนำคำตอบซึ่งกันและกันของข้อสอบให้น้อยลงด้วย อีกทั้งข้อสอบหนึ่งข้อต้องใช้ความสามารถหลายอย่างในการตอบทำให้ข้อสอบมีความยากสูงและซับซ้อน ซึ่งจะลดการขึ้นนำคำตอบระหว่างข้อได้อีกระดับหนึ่ง

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในระยะที่ 1 มีขั้นตอนดังนี้

1) ผู้วิจัยนำแบบสอบฉบับร่าง และเครื่องมือที่ใช้ในการวิจัย เสนอผู้เชี่ยวชาญจำนวน 5 ท่านเพื่อเก็บข้อมูล จำนวน 3 ฉบับ ประกอบด้วย

1.1 แบบประเมินความเหมาะสมของการกำหนดความสัมพันธ์เชิงโครงสร้างของคุณลักษณะย่อย

1.2 แบบประเมินความสอดคล้องของข้อคำถามและคุณลักษณะที่มุ่งวัดแบบปิด

2) ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลกับกลุ่มตัวอย่างจำนวน 1,015 คน โดยทำการติดต่อกับครูผู้ประสานงานของโรงเรียนของกลุ่มตัวอย่างทางโทรศัพท์เนื่องจากขณะทำการศึกษายู่ภายใต้

สถานการณ์โรคระบาดโควิด-19 เพื่อชี้แจงจุดมุ่งหมายของการเก็บข้อมูล ลักษณะตัวแปรที่มุ่งวัด รูปแบบของแบบทดสอบ และการบริหารการสอบ และจัดส่งคู่มือการบริหารการสอบที่ผู้วิจัยสร้างขึ้น ส่งให้ครูผู้ประสานงานทางอีเมลล์ โดยแบบสอบวินิจฉัยฉบับร่างจัดพิมพ์ด้วยกูเกิลฟอร์ม (Google form)

3) ผู้วิจัยตรวจสอบความครบถ้วนของข้อมูลตามจำนวนกลุ่มตัวอย่างและความสมบูรณ์ของการตอบแบบสอบ

4) ผู้วิจัยดำเนินการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล

ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลโดยมีขั้นตอนและรายละเอียด ดังนี้

1) ผู้วิจัยทำการวิเคราะห์เชิงพรรณนา (Descriptive Analytics) เป็นการวิเคราะห์ขั้นพื้นฐาน ที่ทำให้เห็นภาพรวมของข้อมูล และความสัมพันธ์ระหว่างข้อมูลจากแบบประเมิน ความสัมพันธ์เชิงโครงสร้างของคุณลักษณะ (Attributes) ของความสามารถอ่านเพื่อความเข้าใจที่ได้รับการพิจารณาโดยผู้เชี่ยวชาญ 5 ท่าน

2) ผู้วิจัยวิเคราะห์ข้อมูลการตรวจสอบคุณภาพของข้อคำถามรายข้อ ตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) วิเคราะห์ค่าความยาก (P) ค่าอำนาจจำแนกของข้อสอบ (r)

3) ผู้วิจัยวิเคราะห์ข้อมูลการตรวจสอบคุณภาพของข้อคำถามรายข้อ ตามทฤษฎีการตอบสนองข้อสอบ (IRT) แบบ 3 พารามิเตอร์ ได้แก่ วิเคราะห์ค่าพารามิเตอร์ความยาก $/b/$ ค่าพารามิเตอร์อำนาจจำแนก $/a/$ และค่าสัมประสิทธิ์การเดา $/g/$

4) ผู้วิจัยวิเคราะห์ข้อมูลด้วยโปรแกรมอาร์ (R) โดยเรียกใช้ GDINA package และ CDM package

สำหรับข้อมูลที่ถูกรวบรวมในการทดลองใช้แบบสอบฉบับร่าง 2

4.1 การวิเคราะห์ประสิทธิภาพความสอดคล้องเชิงสัมพันธ์ของโมเดลการวินิจฉัยประเภท Generalization DINA (Test Level Fit Statistics) ได้แก่ ค่า-2Loglikelihood (-2LL) ค่า Akaike's Information Criterion (AIC) ค่าBayesian Information Criterion (BIC) และค่า Root Mean Square Error of Approximation (RMSEA) และคุณลักษณะย่อย (Attribute) ใช้เกณฑ์การเปรียบเทียบระหว่าง 2 โมเดล คือ CDM package กับ G-DINA Package โดยค่าที่ต่ำกว่าแสดงให้เห็นถึงประสิทธิภาพในการประมาณค่าของโมเดลที่ดีกว่า

4.2 การวิเคราะห์ความตรงของ Q-matrix ด้วยวิธี Stepwise Wald Test ด้วยโมเดล G-DINA เรียก G-DINA package โปรแกรมอาร์

4.3 การวิเคราะห์คุณภาพแบบสอบบินิจฉัย ได้แก่ ค่าพารามิเตอร์การเดา (Guessing Parameter) ใช้เกณฑ์ค่าพารามิเตอร์การเดาไม่เกินกว่า 0.50 ค่าพารามิเตอร์ความสะเพร่า (Slipping Parameter) ใช้เกณฑ์ค่าพารามิเตอร์ความสะเพร่าไม่เกินกว่า 0.50 ค่าดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ ควรมีค่าน้อยกว่า 0.05 $RMSEA < 0.05$ หมายถึง ข้อสอบข้อนั้นมี ความคลาดเคลื่อนต่ำมากและมีความเหมาะสม เป็นอย่างมากที่จะนำไปใช้วินิจฉัยผู้เรียน และค่าดัชนีอำนาจจำแนก (IDI) มีเกณฑ์คือควรมีค่ามากกว่า 0.50 ($IDI > 0.50$) (วรางคณา ก้อมน้อย, 2563)

ระยะที่ 2 การทดลองและวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียน ชั้นมัธยมศึกษาปีที่ 3 และการพัฒนา Skill profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ

ผู้วิจัยนำแบบสอบที่ผ่านการหาคุณภาพของเครื่องมือในระยะที่ 1 เก็บข้อมูลกับกลุ่มตัวอย่าง

ประชากร ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนที่กำลังศึกษาอยู่ในระดับชั้นมัธยมศึกษาปีที่ 3 ใน ปีการศึกษา 2565 จากโรงเรียนมัธยมศึกษาในสังกัดสำนักงานเขตพื้นที่ การศึกษามัธยมศึกษาเขต 40 จังหวัดเพชรบูรณ์ ทั้งหมด 39 แห่ง จำนวน 4,882 คน

กลุ่มตัวอย่าง กลุ่มตัวอย่างที่ใช้ในการวินิจฉัยและจัดอันดับความสามารถการอ่านเพื่อความเข้าใจ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ที่กำลังศึกษาอยู่ในปีการศึกษา 2565 สังกัดสำนักงานเขตพื้นที่ การศึกษามัธยมศึกษาเพชรบูรณ์ได้มาโดยการสุ่มแบบหลายขั้นตอน (multi-stage sampling) จำนวน 396 หน่วย โดยมีหลักคิดจากการใช้โมเดลการวินิจฉัยเชิงจำแนกของ รัปและ เทมพลิน (Rupp & Templin, 2008: 251, อ้างถึงใน สุมาลี มีสกุล, 2558, หน้า 124) ที่กำหนดว่า กลุ่มตัวอย่างควรมีขนาดไม่ต่ำกว่า 200–300 คน การวิจัยในระยะนี้มีแผนการสุ่ม ดังนี้

1) สุ่มเลือกโรงเรียนที่ใช้เก็บข้อมูลด้วยวิธีสุ่มแบบแบ่งชั้นภูมิ (stratified random sampling) โดยใช้ขนาดโรงเรียนเป็นชั้นภูมิในการสุ่ม พิจารณาจากการกำหนดขนาดสถานศึกษาตามจำนวนนักเรียน เป็น 4 ขนาด ดังนี้ ขนาดเล็ก มีจำนวนนักเรียนตั้งแต่ 119 คนลงมา ขนาดกลาง มีจำนวนนักเรียนตั้งแต่ 120 – 719 คน ขนาดใหญ่ มีจำนวนนักเรียนตั้งแต่ 720 – 1,679 คน ขนาดใหญ่พิเศษ มีจำนวนนักเรียนตั้งแต่ 1,680 คนขึ้นไป (สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2564)

2) เลือกอย่างเจาะจงโรงเรียนขนาดเล็กในสังกัดสำนักงานเขตพื้นที่ มัธยมศึกษา เพชรบูรณ์ ซึ่งมีเพียง 1 โรงเรียนได้แก่ โรงเรียนผาแดงวิทยาคม

3) สุ่มอย่างง่าย (simple random sampling) ด้วยการจับสลากซึ่งระบุชื่อโรงเรียนและขนาดของโรงเรียนแบบไม่ใส่คืนขนาดละ 1 แห่ง ผลการจับสลากสุ่มได้โรงเรียนขนาดใหญ่พิเศษ ได้แก่

หล่มสักวิทยาาคม ขนาดใหญ่ ได้แก่ ศรีเทพประชาสรรค์ ขนาดกลาง ได้แก่ แคมป์สนวิทยาาคม ขนาดเล็ก ได้แก่ ผาแดงวิทยาาคม

ผลการสุ่มแบบหลายขั้นตอนนี้ทำให้ได้ตัวแทนเพื่อเป็นกลุ่มตัวอย่างในการวินิจฉัยความสามารถเป็นจำนวนรวมทั้งสิ้น 4 แห่ง ได้แก่ หล่มสักวิทยาาคม ศรีเทพประชาสรรค์ แคมป์สนวิทยาาคม และผาแดงวิทยาาคม โดยมีรายละเอียดในตาราง 21 ดังนี้

ตาราง 21 แสดงจำนวนตัวอย่างที่ใช้ในการวินิจฉัยและจัดอันดับความสามารถการอ่านเพื่อความเข้าใจ

ขนาดโรงเรียน	ชื่อโรงเรียนที่สุ่ม	จำนวนห้องเรียน	จำนวนห้องที่สุ่ม	จำนวนนักเรียนทั้งหมด	จำนวนตัวอย่าง
ใหญ่พิเศษ	หล่มสักวิทยาาคม	11	6	420	228
ใหญ่	ศรีเทพประชาสรรค์	6	3	225	110
กลาง	แคมป์สนวิทยาาคม	3	1	107	35
เล็ก	ผาแดงวิทยาาคม	1	1	23	23
รวมทั้งสิ้น		24	11	775	396

ขอบเขตด้านตัวแปร

ขอบเขตด้านตัวแปรในระยะวินิจฉัยและจัดอันดับความสามารถการอ่านเพื่อความเข้าใจของนักเรียน

ระดับชั้นมัธยมศึกษาปีที่ 3 ในจังหวัดเพชรบูรณ์ คือ รูปแบบการรอบรู้ของนักเรียน (Mastery Pattern)

ขั้นตอนการดำเนินการวิจัยระยะที่ 2

- 1) นำแบบสอบถามแบบสุ่มสุ่มทำการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจกับผู้สอบจำนวน 396 คน ด้วยการเก็บรวบรวมข้อมูลแบบออนไลน์บนแพลตฟอร์ม FlexiQuiz.com
- 2) ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูล
- 3) ผู้วิจัยดำเนินการวิเคราะห์ข้อมูล
- 4) ศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับการให้ข้อมูลย้อนกลับ

5) ดำเนินการพัฒนา Skill Profile รายงานผลการวินิจฉัยย้อนกลับ ซึ่งเป็นข้อมูลสารสนเทศที่ได้จากการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจของนักเรียนชั้นมัธยมศึกษาปีที่ 3 รายบุคคล ดังนี้

5.1 ผู้วิจัยศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับการพัฒนา Skill Profile

5.2 ผู้วิจัยทำการออกแบบ Skill Profile ที่สามารถให้สารสนเทศผลการวินิจฉัยย้อนกลับ จากการทดสอบ (เกศรินทร์ สมราช, 2563, หน้า 120-121) ประกอบด้วย 4 ส่วน ดังนี้

5.2.1 ข้อมูลสารสนเทศที่ได้จากการวินิจฉัยความสามารถประกอบด้วย คะแนนดิบจากการทดสอบแบบข้อต่อข้อ (ตอบถูกได้ 1 ผิดได้ 0) แบบแผนของการมีคุณลักษณะในความสามารถ ด้านนั้น ๆ เช่น แบบแผนการตอบ “111111” หมายถึง ผู้สอบมีคุณลักษณะในความสามารถ ด้านนั้นครบทุกคุณลักษณะ โดยสามารถบอกค่าความน่าจะเป็น (Skill Probabilities) ในการรอบรู้ คุณลักษณะนั้น

5.2.2 สารสนเทศที่เป็นความสามารถที่ผู้สอบผู้นั้นสามารถปฏิบัติได้

5.2.3 สารสนเทศที่เป็นความบกพร่องของผู้สอบผู้นั้นที่ไม่สามารถปฏิบัติได้

5.2.4 สารสนเทศที่เป็นข้อเสนอแนะในการพัฒนาทักษะ/คุณลักษณะความสามารถที่ผู้สอบไม่รอบรู้ ซึ่งเป็นข้อเสนอแนะที่ผ่านการประเมินความถูกต้องเหมาะสมจากผู้เชี่ยวชาญด้านหลักสูตรและการสอนภาษาอังกฤษระดับมัธยมศึกษา จำนวน 3 ท่าน

5.3 การออกแบบการให้สารสนเทศจาก Profile เพื่อให้ข้อมูลสารสนเทศตามผลการวิเคราะห์ด้วยโมเดลการวินิจฉัยประเภท Generalization DINA (G-DINA MODEL) ซึ่ง สามารถให้สารสนเทศเกี่ยวกับการตรวจสอบความสอดคล้องระหว่างโมเดลกับข้อมูลเชิง ประจักษ์ (Model Fit Information) ความถูกต้องของการวินิจฉัย (Test Level Accuracy) และมี ค่าพารามิเตอร์การเดา (Guessing Parameter) มีค่าดัชนีความคลาดเคลื่อนในการประมาณ ค่าพารามิเตอร์ (RMSEA) ซึ่งใช้ตรวจสอบความเหมาะสมของข้อสอบเป็นรายข้อ และค่าดัชนี อำนาจจำแนก (IDI) ซึ่งเป็นดัชนีที่บ่งชี้ว่า ข้อสอบข้อนั้นสามารถจำแนกผู้สอบที่มีความสามารถ กับผู้ที่ไม่มีความสามารถได้ดีมากน้อยเพียงใด สารสนเทศจาก Profile จะต้องมีทั้งข้อความเชิงบรรยาย ตาราง และแผนภาพ เพื่อช่วยต่อการสร้างความเข้าใจให้กับผู้สอบ

6) ลงมือสร้าง Profile และเสนอต่ออาจารย์ที่ปรึกษาเพื่อขอความคิดเห็นและข้อเสนอแนะในการปรับปรุงแก้ไขการสร้าง Profile

7) ปรับปรุงแก้ไขตามข้อเสนอแนะของอาจารย์ที่ปรึกษา

8) เสนอรายงานผลการวินิจฉัยย้อนกลับ Skill Profile ต่อผู้มีส่วนเกี่ยวข้องเพื่อนำไปใช้ประโยชน์

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยวัดความสามารถการอ่านเพื่อความเข้าใจ จำนวน 1 ฉบับ ได้แก่ แบบสอบแบบทดสอบเพื่อวินิจฉัยเชิงพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 เพื่อวินิจฉัยและจัดอันดับความสามารถการอ่านเพื่อความเข้าใจประกอบด้วยข้อสอบที่ครอบคลุมเนื้อหาอิงเป้าหมายของกรอบอ้างอิงความสามารถทางภาษา CEFR ด้านการอ่านระดับ A2 คือ ผู้เรียนสามารถเข้าใจบทอ่านที่สั้น ๆ ง่าย ๆ เป็นเรื่องใกล้ตัวและใช้ภาษาที่พบบ่อยในชีวิตประจำวันประกอบด้วย 5 หัวข้อ ได้แก่ 1) การอ่านคำแนะนำ (Reading Instruction) 2) การอ่านข้อมูลเบื้องต้น (Reading for Orientation) 3) การอ่านข้อมูลและข้อโต้แย้ง (Reading for Information & Argument) 4) การอ่านวรรณกรรม (Read Literature) และ 5) การอ่านเอกสารโต้ตอบหรือจดหมาย (Reading Correspondence) แต่ละหัวข้อมีข้อสอบสถานการณ์ จำนวน 1 สถานการณ์ รวมเป็น 5 สถานการณ์ เป็นข้อสอบปรนัยแบบเลือกตอบ 4 ตัวเลือก กำหนดรูปแบบการให้คะแนนแบบ 2 ค่า คือ ให้ 1 คะแนน สำหรับข้อที่ผู้เรียนตอบถูก และให้ 0 คะแนน สำหรับข้อที่ผู้เรียนตอบผิด

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในขั้นตอนนี้ เป็นการเก็บรวบรวมข้อมูลแบบผสมผสานกล่าวคือ บริหารการสอบโดยจัดสอบในห้องเรียนที่มีกรรมการคุมสอบอย่างน้อยห้องสอบละ 1 คน ทำข้อสอบผ่านแพลตฟอร์มออนไลน์ FlexiQuiz.com มีขั้นตอนดังนี้

1) ผู้วิจัยโทรศัพท์ติดต่อเขตพื้นที่การศึกษา เพื่อขอความอนุเคราะห์ในการส่งหนังสือถึงโรงเรียนเพื่อขอความร่วมมือในการเก็บข้อมูล

2) ผู้วิจัยจัดทำคู่มือการใช้งานแบบสอบวินิจฉัยที่ผู้วิจัยพัฒนาขึ้นทางออนไลน์บนแพลตฟอร์ม FlexiQuiz.com (นำเสนอใน ภาคผนวก) -ประกอบด้วยคำชี้แจง จุดมุ่งหมายของการเก็บข้อมูล ลักษณะตัวแปรที่มุ่งวัด รูปแบบของแบบทดสอบ และวิธีการจัดสอบ สารสนเทศ Skill Profile รายบุคคล และภาพรวม และแนวทางในการนำสารสนเทศไปใช้ประโยชน์ในการพัฒนาผู้เรียน ผู้วิจัยติดต่อกับครูผู้ประสานงานของโรงเรียนกลุ่มตัวอย่างทางโทรศัพท์เนื่องจากขณะทำการศึกษาอยู่ภายใต้สถานการณ์โรคระบาดโควิด-19

3) ผู้วิจัยทำการนัดหมายวันเวลา และสถานที่ในการเก็บข้อมูลกับครูผู้ประสานงานของแต่ละโรงเรียน จากนั้นดำเนินการเก็บรวบรวมข้อมูลตามวันเวลาดังกล่าวแบบผสมผสานกล่าวคือ บริหารการสอบโดยจัดสอบในห้องเรียนที่มีกรรมการคุมสอบอย่างน้อยห้องสอบละ 1 คน ทำข้อสอบผ่านแบบออนไลน์ โดยได้จัดทำแบบทดสอบบนแพลตฟอร์ม FlexiQuiz.com ซึ่งเป็น Freeware ที่สามารถทำการทดสอบบนเครื่องมือสื่อสารแบบไร้สายที่เชื่อมต่อกับ Internet ได้ หลากหลายเหมาะ

กับอุปกรณ์ทุกรุ่น เพื่อให้มีสติกลุ่มตัวอย่างที่มีโทรศัพท์มือถือแบบ Smart Phone ทุกรุ่น สามารถกดเข้าลิงค์ หรือสแกน QR Code เพื่อทำแบบสอบถามวิจัยโดยระบบจะใช้เวลาในการทำข้อสอบจำกัดที่ 40 นาที ผู้ทดสอบจะสามารถทราบผลคะแนนที่ทำได้ทันที

4) ผู้วิจัยดำเนินการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล

1) เนื่องจากวัตถุประสงค์ของการวิจัยระยะที่ 2 เพื่อวินิจฉัยความสามารถ และเพื่อได้รูปแบบความรอบรู้ ดังนั้นผู้วิจัยวิเคราะห์ข้อมูลด้วยโปรแกรมอาร์ (R) โดยเรียกใช้ GDINA package และ CDM package ในการวิเคราะห์ประสิทธิภาพความสอดคล้องเชิงสัมพันธ์ของโมเดลการวินิจฉัยประเภท Generalization DINA (Test Level Fit Statistics) ได้แก่ค่า -2Loglikelihood ($-2LL$) ค่า Akaike's Information Criterion (AIC) ค่า Bayesian Information Criterion (BIC) และค่า Root Mean Square Error of Approximation (RMSEA) จำแนกความสามารถออกเป็น 6 คุณลักษณะย่อย (Attributes) ใช้เกณฑ์การเปรียบเทียบระหว่าง 2 โมเดล คือ CDM package กับ G-DINA Package โดยค่าที่ต่ำกว่าแสดงให้เห็นถึงประสิทธิภาพในการประมาณค่าของโมเดลที่ดีกว่า

2) การวิเคราะห์คุณภาพแบบสอบวินิจฉัยแบบทดสอบความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ ได้แก่ ค่าพารามิเตอร์การเดา (Guessing Parameter) ใช้เกณฑ์ค่าพารามิเตอร์การเดาไม่เกินกว่า 0.50 ค่าพารามิเตอร์ความสะเพร่า (Slipping Parameter) ใช้เกณฑ์ค่าพารามิเตอร์ความสะเพร่าไม่เกินกว่า 0.50 ค่าดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ (RMSEA) ใช้เกณฑ์ของ Maydeu – Olivares & Joe (2014) $RMSEA < 0.89$ หมายถึง ข้อสอบข้อนั้นมีความเหมาะสมเพียงพอที่จะสามารถนำไปใช้ได้ และ $RMSEA < 0.05$ หมายถึง ข้อสอบข้อนั้นมีความคลาดเคลื่อนต่ำมากและมีความเหมาะสม เป็นอย่างมากที่จะนำไปใช้วินิจฉัยผู้เรียน และค่าดัชนีอำนาจจำแนก (GDI) เป็นดัชนีที่บ่งชี้ว่า ข้อสอบข้อ นั้นสามารถจำแนกผู้สอบที่มีความสามารถกับผู้ที่ไม่มีความสามารถได้ดีมากน้อยเพียงใด โดยมีเกณฑ์ในการพิจารณา คือ GDI ควรค่ามากกว่า 0.50

สถิติสำคัญที่ใช้ในงานวิจัย

การศึกษาครั้งนี้ได้ใช้สถิติทั้งหมด 2 ส่วน สถิติที่ใช้ในการวิเคราะห์ข้อมูลประกอบด้วย 1. สถิติที่ใช้ในการหาคุณภาพของเครื่องมือ และ 2. สถิติที่ใช้ในการวิเคราะห์ข้อมูล มีรายละเอียดดังนี้

1. สถิติที่ใช้ในการหาคุณภาพของเครื่องมือ

ในส่วนของสถิติที่ใช้ในการหาคุณภาพ ของเครื่องมือ มีทั้งหมด 2 กลุ่ม ตามทฤษฎีที่ใช้ในการหาคุณภาพของแบบทดสอบ ดังนี้

1.1 วิเคราะห์หาความเที่ยงตรงเชิงเนื้อหา (Content validity) ของแบบทดสอบ โดยการพิจารณาจากค่าดัชนีความสอดคล้อง (IOC) (แคทรียา แสงดาวเทียน, 2560, หน้า 124)

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

เมื่อ IOC แทน ดัชนีความสอดคล้องความเห็นของผู้เชี่ยวชาญ

$\sum R$ แทน ผลรวมของการพิจารณาของผู้เชี่ยวชาญ

N แทน จำนวนผู้เชี่ยวชาญ

1.2 สถิติที่ใช้ในการหาคุณภาพของเครื่องมือ ตามทฤษฎีการตอบข้อคำถาม

1.2.1 วิเคราะห์หาค่าความยาก (b_i) ค่าอำนาจจำแนก (a_i) และค่าการเดา (c_i) ของข้อคำถาม ด้วยการใช้โมเดลโลจิสติกแบบ 3 พารามิเตอร์ ซึ่งมีสมการในการประมาณค่า(แคทรียา แสงดาวเทียน, 2560, หน้า 124) ดังนี้

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

เมื่อ P_i θ แทน ความน่าจะเป็นของการตอบข้อคำถามที่ i ถูกต้องสำหรับ ผู้สอบที่มี ความสามารถ θ

θ แทน ระดับความสามารถของผู้สอบ

D แทน ค่าคงที่ปรับสเกลมีค่าเท่ากับ 1.7

b_i แทน ค่าความยากของข้อคำถามที่ i

a_i แทน ค่าอำนาจจำแนกของข้อคำถามที่ i

c_i แทน ค่าการเดาของข้อคำถามที่ i

1.2.2 วิเคราะห์หาสารสนเทศของแบบทดสอบ (แคทรียา แสงดาวเทียน, 2560, หน้า 125) คำนวณจากสูตร

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

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

เมื่อ $P_i(\theta)$ แทน สมการที่ใช้ในการประมาณค่าความน่าจะเป็นของการตอบ ข้อ
คำถามที่ i

ถูกต้องสำหรับผู้สอบที่มีความสามารถ θ

$Q_i(\theta)$ แทน สมการที่ใช้ในการประมาณค่าความน่าจะเป็นของการตอบ ข้อ
คำถามที่ i ไม่ถูกต้องสำหรับผู้สอบที่มีความสามารถ θ

1.3 วิเคราะห์หาความเชื่อมั่นของแบบทดสอบ โดยใช้สูตร KR-20 ของคูเดอร์-ริชาร์ดสัน
(Kuder-Richardson) (แคทรียา แสงดาวเทียน, 2560, หน้า 126)

$$r_{tt} = \frac{k}{k-1} \left[1 - \frac{\sum pq}{s_x^2} \right]$$

เมื่อ r_{tt} แทน ค่าความเชื่อมั่นของเครื่องมือ

K แทน จำนวนข้อของแบบทดสอบ

p แทน สัดส่วนของผู้ที่ตอบถูกหรือความยากของแต่ละข้อ

q แทน สัดส่วนของผู้ที่ตอบผิด ซึ่งเท่ากับ $1 - p$

s_x^2 แทน คะแนนผลรวมของข้ออื่น ๆ ที่เหลือทุกข้อ

2. สถิติที่ใช้ในการวิเคราะห์ข้อมูล

2.1 สถิติที่ใช้ในการประมาณค่าในโมเดลการวินิจฉัยเชิงพุทธิปัญญา G-DINA model (Generalized-deterministic-input, noisy-and-gate model) มีสมการที่ใช้ในการประมาณค่า คือ สมการความน่าจะเป็นที่นักเรียนคนที่ i จะตอบข้อสอบข้อที่ j ใน G-DINA model กรณีที่ ข้อนั้น ๆ ต้องการคุณลักษณะ (attributes) หรือทักษะ (skills) α จำนวน K คุณลักษณะ คือ (George A.C., Robitzsch A, Kiefer T, GroB J, Ünlü A, 2016, อ้างถึงใน วรางคณา ก้อมน้อย, 2563, หน้า 68)

$$P(\alpha_{ij}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{ik} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^*-1} \delta_{jkk'} \alpha_{ik} \alpha_{ik'} + \dots + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{ik},$$

เมื่อ δ_{j0} คือ ความน่าจะเป็นในการตอบข้อนั้น ๆ ถูก กรณี ผู้ตอบข้อสอบไม่มีคุณลักษณะ
ที่ต้องการในข้อนั้น ๆ

δ_k คือ อิทธิพลหลัก (Main effect) ของการบรรลุคุณลักษณะเดียว

$\delta_{kk'}$ คือ ปฏิสัมพันธ์ของการบรรลุคุณลักษณะสองอย่าง

$\delta_{j1,2,\dots,K_j}$ คือ ปฏิสัมพันธ์ที่ซับซ้อนที่เกิดจากการบรรลุคุณลักษณะทุกอย่าง

โดยยกตัวอย่างเพื่อประกอบการอธิบายตัวอย่างสมการที่ 1 (เกศรินทร์ สมราช, 2563, หน้า 84) ดังนี้ สมการความน่าจะเป็นที่ผู้เรียนคนที่ i จะตอบข้อสอบข้อที่ j ใน G-DINA MODEL กรณีที่ข้อนั้น ๆ ต้องการคุณลักษณะ α_1 และ α_2 สองคุณลักษณะ คือ

$$P(X_{ij} = 1 | \alpha_1, \alpha_2) = \delta_{j0} + \delta_{j1}\alpha_1 + \delta_{j2}\alpha_2 + \delta_{j12}\alpha_1\alpha_2$$

โดยที่ พารามิเตอร์ δ_{j0} คือ ความน่าจะเป็นในการตอบข้อนั้น ๆ ถูก กรณี ผู้ตอบข้อสอบไม่มีคุณลักษณะที่ต้องการในข้อนั้น ๆ พารามิเตอร์ δ_{j1} and δ_{j2} คือ ค่าอิทธิพลหลักของคุณลักษณะ α_1 และ α_2 ตามลำดับ และค่าพารามิเตอร์ δ_{j12} ค่าอิทธิพลปฏิสัมพันธ์ของคุณลักษณะ α_1 และ α_2 สรุป G-DINA MODEL ผู้ตอบไม่จำเป็นต้องมีคุณลักษณะ α_1 และ α_2 ครอบพร้อมกัน ถึงจะมีโอกาสในการทำข้อสอบที่ j ถูก เนื่องจากความน่าจะเป็นในการตอบถูกนั้นขึ้นอยู่กับ 1. ความน่าจะเป็นในการตอบข้อ j ถูก กรณีผู้ตอบข้อสอบไม่มี คุณลักษณะที่ต้องการ 2. ค่าอิทธิพลหลักของคุณลักษณะที่ต้องการ 3. ค่าอิทธิพลปฏิสัมพันธ์

ตัวอย่างที่ 2 สมการความน่าจะเป็นที่ผู้เรียนคนที่ i จะตอบข้อสอบข้อที่ j ใน G-DINA MODEL กรณีที่ข้อนั้น ๆ ต้องการคุณลักษณะ α_1 , α_2 และ α_3 สามคุณลักษณะ (เกศรินทร์ สมราช, 2563, หน้า 85) คือ

$$P(X_{ij} = 1 | \alpha_1, \dots, \alpha_k) = \delta_{j0} + \delta_{j1}\alpha_1 + \delta_{j2}\alpha_2 + \delta_{j3}\alpha_3 + \delta_{j12}\alpha_1\alpha_2 + \delta_{j13}\alpha_1\alpha_3 + \delta_{j23}\alpha_2\alpha_3 + \delta_{j123}\alpha_1\alpha_2\alpha_3$$

โดยพารามิเตอร์ δ_{j0} คือ ความน่าจะเป็นในการตอบข้อนั้น ๆ ถูก กรณี ผู้ตอบข้อสอบไม่มีคุณลักษณะที่ต้องการในข้อนั้น ๆ พารามิเตอร์ δ_{j1} , δ_{j2} และ δ_{j3} คือ ค่าอิทธิพลหลักของคุณลักษณะ α_1 , α_2 และ α_3 พารามิเตอร์ δ_{j12} , δ_{j13} และ δ_{j23} คือ ค่าอิทธิพลปฏิสัมพันธ์ของคุณลักษณะ α_1 , α_2 และ α_3 ที่ละคู่ พารามิเตอร์ δ_{j123} คือ ค่าอิทธิพลปฏิสัมพันธ์ของคุณลักษณะทั้ง α_1 , α_2 และ α_3 สรุป จะเห็นได้ว่า G-DINA MODEL ผู้ตอบข้อสอบไม่จำเป็นต้องมีคุณลักษณะ α_1 , α_2 และ α_3 ครอบพร้อมกัน จึงจะมีโอกาสทำข้อสอบที่ j ถูก เนื่องจากความน่าจะเป็นในการตอบถูกนั้นขึ้นอยู่กับความน่าจะเป็นในการตอบข้อ j ถูก ถ้าผู้ตอบแบบทดสอบไม่มีคุณลักษณะที่ต้องการ ค่าอิทธิพลหลักของคุณลักษณะที่ต้องการค่าอิทธิพลปฏิสัมพันธ์ที่ละคู่และ ค่าอิทธิพลปฏิสัมพันธ์ของคุณลักษณะที่จำเป็นต้องใช้ในข้อนั้น ๆ สรุปจะเห็นได้ว่า G-DINA MODEL ผู้ตอบข้อสอบไม่จำเป็นต้องมี คุณลักษณะ α_1 , α_2 และ α_3 ครอบพร้อมกัน จึงจะมีโอกาสทำข้อสอบที่ j ถูก เนื่องจากความน่าจะเป็นในการตอบถูกนั้นขึ้นอยู่กับความน่าจะเป็นในการตอบข้อ j ถูก

ถ้าผู้ตอบแบบทดสอบไม่มี คุณลักษณะที่ต้องการ ค่าอิทธิพลหลักของคุณลักษณะที่ต้องการค่าอิทธิพล ปฏิสัมพันธ์ที่ละคู่และ ค่าอิทธิพลปฏิสัมพันธ์ของคุณลักษณะที่จำเป็นต้องใช้ในข้อนั้น ๆ (เกศรินทร์ สมราช, 2563, หน้า 85)

2.2 ความตรงเชิงจำแนกรายคุณลักษณะ (Wang, et al., 2015, p. 461)

$$\hat{\tau}_k = \frac{\sum_i \sum_c \left(\mathbf{P}_{N \times 2}^{(k)} * \mathbf{W}_{N \times 2}^{(k)} \right)}{N},$$

2.3 ความตรงเชิงจำแนกรายรูปแบบ (Iaconangelo, 2017, p. 13)

$$\hat{\tau}_l = P(\alpha_s | \alpha_l, \mathbf{X}) I[s = l]$$

2.4 ความตรงเชิงจำแนกรายฉบับ (Iaconangelo, 2017, p. 14)

$$\tau = \frac{\sum_{i=1}^N I[\alpha_i = \hat{\alpha}_i]}{N},$$

2.5 ดัชนีอำนาจจำแนกรายข้อ (de la Torre, 2008, pp. 343 – 362 and de la Torre, Rossi and van der Ark, 2018, pp. 281 – 296)

$$IDI_j = \max_{\alpha_1, \alpha_2, h} \left| P(X_j = h | \alpha_1) - P(X_j = h | \alpha_2) \right|$$

2.6 ความเที่ยงเชิงจำแนกรายคุณลักษณะและรายฉบับ (Johnson and Sinharay, 2018, pp. 635 – 664)

$$\hat{r}_{ij} = \sum_{a \in \Omega} \frac{\left(\sum_{n=1}^N P(A=a | X=x_n) I\{\hat{a}_{nd}=i\} \right) \left(\sum_{n=1}^N P(A=a | X=x_n) I\{\hat{a}_{nd}=j\} \right)}{N^2 P(A=a)}.$$

2.7 การประมาณค่ารูปแบบการรอบรู้ (attribute profiles) แบบ Maximum a Posteriori: MAP (Huebner & Wang, 2011, pp. 407-419)

$$\hat{\alpha}_{MAP} = \underset{i}{\operatorname{argmax}} \left\{ P(\alpha_i | X_i) \right\}$$

บทที่ 4

ผลการวิจัย

การวิจัยเรื่องการพัฒนาแบบสอบแบบทดสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ มีวัตถุประสงค์การวิจัย 2 ข้อ ได้แก่ 1) เพื่อสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ 2) เพื่อวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และพัฒนา Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ เพื่อตอบวัตถุประสงค์การวิจัย จึงแบ่งการวิเคราะห์ออกเป็น 2 ตอน ดังนี้

ตอนที่ 1 ผลการสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3

ตอนที่ 2 ผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และผลการออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ

การนำเสนอผลการวิเคราะห์ข้อมูล

ตอนที่ 1 ผลการสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3

ผลการสร้างและตรวจสอบคุณภาพแบบสอบวินิจฉัย ผู้วิจัยนำเสนอผลการวิเคราะห์ข้อมูลออกเป็น 5 ข้อ ได้แก่ 1) ผลการตรวจสอบความถูกต้องเหมาะสมของการกำหนดคุณลักษณะและความสัมพันธ์เชิงโครงสร้างของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ 2) ผลการสร้างแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ 3) ผลการตรวจสอบความสอดคล้องระหว่างข้อคำถามและคุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ 4) ผลการตรวจสอบคุณภาพของแบบสอบ และ 5) ผลการคัดเลือกข้อสอบและการสร้าง Q-matrix ของแบบสอบวินิจฉัย โดยมีรายละเอียดดังต่อไปนี้

1.1 ผลการกำหนดคุณลักษณะและผลการตรวจสอบความถูกต้องเหมาะสมของคุณลักษณะความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ

ผู้วิจัยศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง ทำการสังเคราะห์คุณลักษณะเพื่อดำเนินการกำหนดคุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ คำอธิบายคุณลักษณะ ทำการประเมินความถูกต้องเหมาะสมโดยให้ผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษและการวัดผล

ประเมินผลจำนวน 5 ท่าน โดยพิจารณาว่าคุณลักษณะใดเป็นคุณลักษณะที่นักเรียนควรมีเป็นอันดับแรก และลำดับถัดไป จนครบทั้ง 6 คุณลักษณะ ผลปรากฏว่ามีคุณลักษณะความสามารถที่นักเรียนใช้จำนวน 6 คุณลักษณะ ดังนี้

A1 มีความรู้คำศัพท์ (Vocabulary Knowledge) พิจารณาจากความเข้าใจเกี่ยวกับความหมายคำศัพท์ และกลุ่มคำโดยอาศัยพื้นฐานความรู้เดิม หรือเข้าใจความหมายของคำศัพท์ที่ยากหรือไม่คุ้นเคยโดยอาศัยการคาดเดาความหมายคำศัพท์จากบริบทหรือคำบ่งชี้ที่พบในข้อความ (Context Clue) ประกอบด้วยการนิยามความหมาย การกล่าวซ้ำ การเปรียบเทียบความเหมือนหรือความต่าง และการยกตัวอย่าง

A2 มีความสามารถในการระบุข้อมูลที่ต้องการ (Identifying Explicit Information) พิจารณาได้จากทักษะในการค้นหารายละเอียดหรือข้อมูลเฉพาะที่ต้องการซึ่งปรากฏชัดเจนในบทอ่าน หลังจากที่ได้อ่านแบบผ่าน (Scan) และอ่านแบบพินิจพิเคราะห์ (Skim) ทำให้ทราบว่าข้อมูลอยู่ที่ใดของเรื่อง เช่น จำนวน ราคา วันเวลา สถานที่ ซึ่งข้อมูลดังกล่าวจะปรากฏอย่างชัดเจน ผู้อ่านสามารถเข้าใจได้โดยมีสติความ

A3 มีความรู้ทางไวยากรณ์ (Syntactic Knowledge) พิจารณาได้จากความเข้าใจเกี่ยวกับวากยสัมพันธ์ซึ่งหมายถึงการศึกษาความสัมพันธ์ของคำในประโยคหรือแบบแผนองค์ประกอบของประโยคในภาษาอังกฤษ และความสัมพันธ์ระหว่างประโยคและสามารถวางตำแหน่งของคำหรือกลุ่มคำประเภทต่าง ๆ ในประโยคได้อย่างถูกต้องตามโครงสร้างทางไวยากรณ์

A4 มีความสามารถในการระบุใจความสำคัญของเรื่อง (Generalization Main Idea) พิจารณาได้จากการทำความเข้าใจเพื่อค้นหาหัวเรื่อง (Topic) นำมารวมกับข้อความที่ถูกควบคุมให้กล่าวถึงสิ่งใดสิ่งหนึ่ง (Controlling Idea) และลงข้อสรุปเป็นใจความสำคัญ หรือแนวคิดหลักที่อยู่ในบทอ่านซึ่งเป็นส่วนที่ครอบคลุมย่อหน้านั้น ๆ หรือเนื้อหาทั้งหมด

A5 มีความสามารถในการอนุมาน (Getting Inference) พิจารณาได้จากความสามารถในการคาดคะเนตามหลักเหตุผล หรือการตีความจากเรื่องที่อ่าน โดยใช้ความรู้และประสบการณ์ตรงของผู้อ่าน เพื่อทำความเข้าใจข้อมูลที่ไม่ได้ถูกระบุไว้อย่างชัดเจนซึ่งเป็นจุดมุ่งหมายที่สคริปต์ของผู้เขียน หรือข้อโต้แย้ง โดยอาศัยเบาะแสร่องรอยในเนื้อความซึ่งเป็นคำหรือประโยคชี้แนะ หรือทำการตรวจสอบเท็จจริงทั้งหมดของเรื่องที่อ่าน เพื่อพิจารณาหรือตัดสินใจว่าสิ่งที่อ่านทั้งหมดนั้นหมายความว่าถึงสิ่งใดและสามารถเชื่อมโยงไปถึงข้อมูลภายนอกซึ่งไม่ได้กล่าวไว้ในบทอ่าน

A6 มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน (Synthesizing Information from Text) พิจารณาได้จากความสามารถในการแยกประโยคแนวคิดหลัก หรือใจความสำคัญออกจากประโยครองหรือข้อความ ซึ่งเป็นรายละเอียดเสริมจากบทอ่าน นำประเด็น

หลักที่รวบรวมได้มาเชื่อมโยงเชื่อมต่อกัน โดยพิจารณาความเกี่ยวข้องความเป็นเหตุเป็นผล ทำการจัดลำดับข้อความและทำการลงข้อสรุปหรือเรียบเรียงประโยคขึ้นใหม่

1.2 ผลการสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ

ผู้วิจัยนำเสนอผลการสร้างและแบบสอบวินิจฉัย แบ่งออกเป็น 3 ระยะ ดังนี้ 1) ผลการสร้างและตรวจสอบคุณภาพด้านความตรงเชิงเนื้อหาของแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจ จำนวน 42 ข้อ โดยผู้เชี่ยวชาญ 2) ผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยทฤษฎีการทดสอบดั้งเดิม (CTT) และ 3) ผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยทฤษฎีการตอบสนองข้อสอบ (IRT 3PL) ดังนี้

1.2.1 ผลการสร้างและตรวจสอบคุณภาพด้านความตรงเชิงเนื้อหาของแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจ จำนวน 42 ข้อ โดยผู้เชี่ยวชาญจำนวน 5 ท่าน ผู้วิจัยดำเนินการสร้างแบบสอบวินิจฉัยแบบ General diagnosis กำหนดเนื้อหาอิงกรอบมาตรฐาน CEFR ด้านการอ่านระดับ A2 คือ ผู้เรียนสามารถเข้าใจบทความที่สั้นและง่ายเป็นเรื่องใกล้ตัวและใช้ภาษาที่พบบ่อยในชีวิตประจำวัน ทำการวินิจฉัยอย่างเป็นทางการ (formal diagnosis) เพื่อค้นหาข้อบกพร่องของนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยมีวิธีการสร้างแบบสอบวินิจฉัยอย่างเป็นทางการเป็นระบบอิงแนวคิดกรอบกระบวนการ CDS ของเอ็มเบรทสัน (Embretson, 1994) เริ่มต้นจากการกำหนดโครงสร้างแบบสอบที่มีจำนวนเทสต์เลทเท่ากับจำนวนเนื้อหาอิงกรอบมาตรฐาน CEFR ระดับ A2 จากนั้นกำหนดบทความที่สอดคล้องกับกรอบ CEFR แต่ละบทความมีข้อคำถามที่เกี่ยวข้องกับบทความเป็นเซตเดียวกัน อาจมีจำนวนข้อคำถาม 3-5 ข้อ หรือบางครั้ง 6-12 ข้อ ซึ่งบทความควรจะมีความยาวประมาณ 300 คำหรือมากกว่า (Wainer, Bradlow, & Wang, 2007, p. 60, อ้างถึงใน พิสิฐ พิณจสกุล, 2563, หน้า 77; Wheadon, 2010, p. 106; อริสสา เตหลิ้ม, 2559, หน้า 2; จงกล บัวแก้ว, 2560, หน้า 4) ให้ครอบคลุมความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษจำแนกตามคุณลักษณะ ประกอบด้วย 5 หัวข้อ (สถาบันภาษาอังกฤษ สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2560) ดังนี้

1) การอ่านคำแนะนำ (Reading Instruction) คำแนะนำที่ใช้ภาษาง่าย ๆ เช่น ตู้ขายสินค้าอัตโนมัติ โทรศัพท์สาธารณะ เครื่องจำหน่ายตั๋วสาธารณะ ข้อมูลเกี่ยวกับความปลอดภัย การบอกทิศทาง สามารถปฏิบัติตามคำสั่ง คำชี้แจง และขั้นตอนง่าย ๆ เช่น ขั้นตอนในการเล่นเกม การใช้อุปกรณ์ที่คุ้นเคย หรือการทำอาหาร เป็นต้น

2) การอ่านอย่างมีจุดประสงค์ (Reading for Orientation) สามารถหาข้อมูลที่สำคัญในบทโฆษณาแผ่นพับ ใบปลิว หนาเว็บ (web page) แคตตาล็อก ตารางเวลา ฯลฯ

3) การอ่านเพื่อหาข้อมูลและข้อโต้แย้ง (Reading for Information & Argument) คำบรรยาย

ข่าวสั้น ๆ บทความสั้น ๆ จากหนังสือพิมพ์ และนิตยสารเกี่ยวกับเหตุการณ์ ปัจจุบัน และหัวข้อที่คุ้นเคย ถ้าผู้อ่านมีความรู้เรื่องเกี่ยวกับเรื่องนั้น ๆ บางอยู่แล้ว เช่น ข่าวเกี่ยวกับกีฬา และบุคคลที่มีชื่อเสียง

4) การอ่านวรรณกรรม (Read Literature) เรื่องในชีวิตประจำวันที่ย้าย ๆ และสั้น ๆ โดยเฉพาะอย่างยิ่งเมื่อมีสื่อประกอบ

5) การอ่านเอกสารโต้ตอบหรือจดหมาย (Reading Correspondence) โต้ตอบจดหมายส่วนตัวเกี่ยวกับเหตุการณ์ต่าง ๆ

ตาราง 22 โครงสร้างแบบสอบวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ

Testlets	เนื้อหาอิงกรอบ CEFR ระดับ A2	Topic	จำนวนข้อคำถาม
Testlets 1	Reading Instruction	The instruction of vending machine	10
Testlets 2	Reading for Orientation and Reading Correspondence	The advertisement of healthy workshop	10
Testlets 3	Reading for Info & Argument	The news about COVID-19	12
Testlets 4	Reading Literature	The story about Thai culture	10
รวมทั้งสิ้น (ข้อ)			42

ผลการตรวจสอบความตรงเชิงเนื้อหาด้วยดัชนีความสอดคล้องระหว่างข้อคำถามและคุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษมีบทอ่าน (Passages) จำนวน 4 เรื่อง ประกอบด้วยข้อสอบแบบทดสอบ (Testlets) จำนวนไม่เท่ากัน โดยเป็นข้อสอบปรนัยแบบเลือกตอบ 4 ตัวเลือก กำหนดรูปแบบการให้คะแนนแบบ 2 ค่า คือ ให้ 1 คะแนน สำหรับข้อที่ผู้เรียนตอบถูก และให้ 0 คะแนน สำหรับข้อที่ผู้เรียนตอบผิด ประกอบด้วยข้อคำถามจำนวน 42 ข้อ พิจารณาความตรงเชิงเนื้อหาด้วยดัชนีวัดความสอดคล้องระหว่างข้อคำถามและคุณลักษณะของความสามารถที่ใช้ในการตอบ (Index of Item-Objective Congruence: IOC) โดยผู้เชี่ยวชาญด้าน

การสอนภาษาอังกฤษ การวิจัยและประเมินผลทางการศึกษาจำนวน 5 ท่าน เพื่อตรวจสอบความสอดคล้องของข้อคำถามและคุณลักษณะที่มุ่งวัด (Q-matrix) ผลปรากฏว่า ข้อคำถาม 42 ข้อ มีค่าดัชนี IOC ตั้งแต่ 0.6 ถึง 1.0 ค่าดัชนี IOC ควรีค่า ≥ 0.5 ดังนั้นข้อคำถามทุกข้อมีความตรงเชิงเนื้อหาและมีความสอดคล้องระหว่างข้อคำถามและคุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ แสดงในตาราง 23 ดังนี้

ตาราง 23 ผลการตรวจสอบความสอดคล้องระหว่างข้อคำถามและคุณลักษณะของความสามารถการอ่าน

เทสต์เลข ที่	ข้อ เดิม	สิ่งที่ พิจารณา	จำนวนผู้เชี่ยวชาญที่			IO C	สรุปผล	ข้อ ใหม่
			พิจารณาให้คะแนน (คน)					
			+1	0	-1			
Testlet 1	1	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	1
		คุณลักษณะ	5	-	-	1.0		
	2	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	2
		คุณลักษณะ	5	-	-	1.0		
	3	เนื้อหา	4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้ ได้	3
		คุณลักษณะ	5	-	-	1.0		
	4	เนื้อหา	4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้ ได้	4
		คุณลักษณะ	5	-	-	1.0		
	5	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	5
		คุณลักษณะ	5	-	-	1.0		
	6	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	6
		คุณลักษณะ	5	-	-	1.0		
	7	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	7
		คุณลักษณะ	5	-	-	1.0		
	8	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	8
		คุณลักษณะ	5	-	-	1.0		
	9	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	9
		คุณลักษณะ	5	-	-	1.0		
	10	เนื้อหา	3	2	-	0.6	ข้อสอบมีความตรง นำไปใช้ ได้	10
		คุณลักษณะ	3	2	-	0.6		

เทสต์เลขที่	ข้อ เดิม	สิ่งที่ พิจารณา	จำนวนผู้เชี่ยวชาญที่ พิจารณาให้คะแนน (คน)			IO C	สรุปผล	ข้อ ใหม่
Testlet 2	11	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้	11
		คุณลักษณะ	5	-	-	1.0	ได้	
	12	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้	12
		คุณลักษณะ	5	-	-	1.0	ได้	
	13	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้	13
		คุณลักษณะ	5	-	-	1.0	ได้	
	14	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้	14
		คุณลักษณะ	5	-	-	1.0	ได้	
	15	เนื้อหา	4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้	15
		คุณลักษณะ	4	1	-	0.8	ได้	
	16	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้	16
		คุณลักษณะ	5	-	-	1.0	ได้	
	17	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้	17
		คุณลักษณะ	5	-	-	1.0	ได้	
	18	เนื้อหา	4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้	18
		คุณลักษณะ	5	-	-	1.0	ได้	
	19	เนื้อหา	4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้	19
		คุณลักษณะ	5	-	-	1.0	ได้	
	20	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้	20
		คุณลักษณะ	5	-	-	1.0	ได้	
Testlet 3	21	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้	21
		คุณลักษณะ	5	-	-	1.0	ได้	
	22	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้	22
		คุณลักษณะ	5	-	-	1.0	ได้	
	23	เนื้อหา	5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้	23
		คุณลักษณะ	5	-	-	1.0	ได้	
	24	เนื้อหา	4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้	24
		คุณลักษณะ	5	-	-	1.0	ได้	
	25	เนื้อหา	4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้	25
		คุณลักษณะ	5	-	-	1.0	ได้	

เทสต์เลข ที่	ข้อ เดิม	สิ่งที่ พิจารณา	จำนวนผู้เชี่ยวชาญที่ พิจารณาให้คะแนน (คน)			IO C	สรุปผล	ข้อ ใหม่
26	เนื้อหา		4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้ ได้	26
	คุณลักษณะ		5	-	-	1.0		
27	เนื้อหา		3	2	-	0.6	ข้อสอบมีความตรง นำไปใช้ ได้	27
	คุณลักษณะ		3	2	-	0.6		
28	เนื้อหา		3	2	-	0.6	ข้อสอบมีความตรง นำไปใช้ ได้	28
	คุณลักษณะ		3	2	-	0.6		
29	เนื้อหา		5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	29
	คุณลักษณะ		5	-	-	1.0		
30	เนื้อหา		5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	30
	คุณลักษณะ		5	-	-	1.0		
31	เนื้อหา		5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	31
	คุณลักษณะ		5	-	-	1.0		
32	เนื้อหา		5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	32
	คุณลักษณะ		5	-	-	1.0		
Testlet 4	เนื้อหา		4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้ ได้	33
	คุณลักษณะ		5	-	-	1.0		
34	เนื้อหา		5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	34
	คุณลักษณะ		5	-	-	1.0		
35	เนื้อหา		5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	35
	คุณลักษณะ		5	-	-	1.0		
36	เนื้อหา		4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้ ได้	36
	คุณลักษณะ		4	1	-	0.8		
37	เนื้อหา		5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	37
	คุณลักษณะ		5	-	-	1.0		
38	เนื้อหา		5	-	-	1.0	ข้อสอบมีความตรง นำไปใช้ ได้	38
	คุณลักษณะ		5	-	-	1.0		
39	เนื้อหา		4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้ ได้	39
	คุณลักษณะ		4	1	-	0.8		
40	เนื้อหา		3	2	-	0.6	ข้อสอบมีความตรง นำไปใช้ ได้	40
	คุณลักษณะ		3	2	-	0.6		

เทสต์เลข ที่	ข้อ เดิม	สิ่งที่ พิจารณา	จำนวนผู้เชี่ยวชาญที่ พิจารณาให้คะแนน (คน)			IO C	สรุปผล	ข้อ ใหม่
41	เนื้อหา		4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้ ได้	41
	คุณลักษณะ		5	-	-	1.0		
42	เนื้อหา		4	1	-	0.8	ข้อสอบมีความตรง นำไปใช้ ได้	42
	คุณลักษณะ		5	-	-	1.0		

1.2.2 ผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยทฤษฎีการทดสอบ ดั้งเดิม (Classical Test Theory: CTT)

ผู้วิจัยนำแบบสอบวินิจฉัย (ฉบับร่าง) ซึ่งมีข้อคำถามจำนวน 42 ข้อ ไปทดลองใช้กับกลุ่มตัวอย่างชุดที่ 1 ซึ่งเป็นนักเรียนที่กำลังศึกษาในระดับชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2565 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเพชรบูรณ์ จำนวน 511 หน่วย ได้มาโดยการสุ่มอย่างง่าย ผลปรากฏว่า ข้อคำถามที่ผ่านเกณฑ์คุณภาพด้านความยากและอำนาจจำแนกตามทฤษฎีทดสอบแบบดั้งเดิมมีจำนวน 33 ข้อ มีค่าความยากในภาพรวมอยู่ระหว่าง 0.25 ถึง 0.70 และค่าอำนาจจำแนกอยู่ระหว่าง 0.26 ถึง 0.87 แบ่งออกเป็นแต่ละเทสต์เลข ดังนี้ เทสต์เลขที่ 1 ข้อคำถามที่ 1 ถึง 10 มีค่าความยากอยู่ระหว่าง 0.51 ถึง 0.66 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.57 ถึง 0.87 ข้อคำถามที่ผ่านเกณฑ์คุณภาพด้านความยากและอำนาจจำแนกมีจำนวน 8 ข้อ ข้อที่ถูกคัดออกได้แก่ ข้อ 2 และข้อ 9 เทสต์เลขที่ 2 ข้อคำถามที่ 11 ถึง 21 มีค่าความยากอยู่ระหว่าง 0.25 ถึง 0.67 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.26 ถึง 0.70 ข้อคำถามที่ผ่านเกณฑ์คุณภาพด้านความยากและอำนาจจำแนกมีจำนวน 8 ข้อ ข้อที่ถูกคัดออกได้แก่ ข้อ 13 และข้อ 16 เทสต์เลขที่ 3 ข้อคำถามที่ 21 ถึง 33 มีค่าความยากอยู่ระหว่าง 0.39 ถึง 0.79 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.22 ถึง 0.81 ข้อคำถามที่ผ่านเกณฑ์คุณภาพด้านความยากและอำนาจจำแนกมีจำนวน 9 ข้อ ข้อที่ถูกคัดออกได้แก่ ข้อ 26 29 30 และข้อ 33 เทสต์เลขที่ 4 ข้อคำถามที่ 34 ถึง 42 มีค่าความยากอยู่ระหว่าง 0.34 ถึง 0.63 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.24 ถึง 0.76 ข้อคำถามที่ผ่านเกณฑ์คุณภาพด้านความยากและอำนาจจำแนกมีจำนวน 8 ข้อ ข้อที่ถูกคัดออกได้แก่ข้อ 34 ผู้วิจัยนำมาเรียบเรียงเป็นแบบสอบวินิจฉัย ฉบับร่าง 1 เสนอต่ออาจารย์ที่ปรึกษา ปรับปรุงตามคำแนะนำเพื่อเตรียมเข้าสู่กระบวนการต่อไป ดังตาราง 24

ตาราง 24 แสดงผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยทฤษฎีการทดสอบดั้งเดิม

Testlet	ข้อที่ผ่าน	ความยาก (P)	อำนาจจำแนก (r)	ข้อใหม่
---------	------------	----------------	-----------------	---------

Testlet	ข้อที่ผ่าน	ความยาก (P)	อำนาจจำแนก (r)	ข้อใหม่
T1: Reading instruction	1	0.61	0.57	1
	3	0.61	0.69	2
	4	0.66	0.64	3
	5	0.51	0.75	4
	6	0.62	0.74	5
	7	0.60	0.68	6
	8	0.63	0.72	7
	10	0.53	0.87	8
T2: Reading for Orientation and Reading Correspondence	11	0.51	0.57	9
	12	0.51	0.54	10
	14	0.44	0.46	11
	15	0.67	0.65	12
	17	0.41	0.26	13
	18	0.48	0.64	14
	19	0.47	0.38	15
	20	0.61	0.70	16
T3: Reading for Info & Argument	21	0.70	0.36	17
	22	0.60	0.66	18
	23	0.60	0.75	19
	24	0.48	0.54	20
	25	0.56	0.81	21
	27	0.39	0.22	22
	28	0.57	0.54	23
	31	0.61	0.64	24
T4: Reading Literature	32	0.54	0.70	25
	35	0.56	0.58	26
	36	0.49	0.48	27
	37	0.57	0.51	28

Testlet	ข้อที่ผ่าน	ความยาก (P)	อำนาจจำแนก (r)	ข้อใหม่
	38	0.57	0.70	29
	39	0.34	0.24	30
	40	0.58	0.68	31
	41	0.60	0.76	32
	42	0.63	0.70	33

1.2.3 ผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT 3PL)

ผู้วิจัยนำแบบสอบวินิจฉัยฉบับร่าง 1 ซึ่งมีข้อคำถามจำนวน 33 ข้อ ไปทดลองใช้กับกลุ่มตัวอย่างชุดที่ 2 ซึ่งเป็นนักเรียนที่กำลังศึกษาในระดับชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2565 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเพชรบูรณ์ จำนวน 1,015 หน่วย ได้มาโดยการสุ่มแบบหลายขั้นตอน พบว่าจากข้อคำถามทั้งหมด 33 ข้อ มีข้อที่ผ่านเกณฑ์จำนวน 18 ข้อ คัดออกจำนวน 15 ข้อ ได้แก่ ข้อ 2, 7, 8, 9, 10, 13, 19, 22, 23, 24, 25, 28, 32 และ 33 เนื่องจาก ค่าพารามิเตอร์ความยากควรมีค่าอยู่ระหว่าง -2.50 ถึง +2.50 ค่าพารามิเตอร์อำนาจจำแนกควรมีค่าอยู่ระหว่าง +0.50 ถึง +2.50 และค่าสัมประสิทธิ์การเดาควรมีค่าไม่เกิน 0.30 (ศิริชัย กาญจนวาสี, 2550, หน้า 56, อ้างถึงใน แคทรียา แสงดาวเทียน, 2560, หน้า 136) แบบสอบมีความเที่ยง Marginal Reliability for Response Pattern Scores เท่ากับ 0.83 ซึ่งมีค่าใกล้เคียง 1.0 หมายความว่า แบบสอบฉบับนี้มีความเที่ยง และแบบสอบให้ค่าสารสนเทศจากการสอบสูงสุดเมื่อนำมาใช้กับกลุ่มผู้สอบที่มีระดับความสามารถ (θ) เท่ากับ 1.60

ผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT 3PL) ค่าพารามิเตอร์ความยาก /b/ ค่าพารามิเตอร์อำนาจจำแนก /a/ และสัมประสิทธิ์การเดา /g/ ผลปรากฏว่า ข้อคำถามที่มีค่าพารามิเตอร์ความยาก ค่าพารามิเตอร์อำนาจจำแนก และสัมประสิทธิ์การเดาผ่านเกณฑ์มีจำนวน 18 ข้อ ในภาพรวมข้อคำถามมีค่าพารามิเตอร์ความยากอยู่ระหว่าง -0.29 ถึง 1.19 ค่าพารามิเตอร์อำนาจจำแนกอยู่ระหว่าง 0.82 ถึง 2.50 และสัมประสิทธิ์การเดา อยู่ระหว่าง 0.12 ถึง 0.28

ผลการตรวจสอบรายเทสต์เลท จำนวน 4 เทสต์เลท ดังนี้ Testlet 1: Reading Instruction ข้อที่ผ่านเกณฑ์มีจำนวน 5 ข้อ ได้แก่ 1, 4, 5, 6, 7 มีค่าพารามิเตอร์ความยาก /b/ อยู่ระหว่าง -0.24 ถึง 0.81 ค่าพารามิเตอร์อำนาจจำแนก /a/ อยู่ระหว่าง 1.83 ถึง 2.46 และ

สัมประสิทธิ์การเดา /g/ อยู่ระหว่าง 0.18 ถึง 0.22 Testlet 2: Reading for Orientation มีข้อที่ผ่านเกณฑ์มีจำนวน 4 ข้อ ได้แก่ 14, 18, 19, 20 มีค่าพารามิเตอร์ความยาก /b/ อยู่ระหว่าง -0.05 ถึง 1.13 ค่าพารามิเตอร์อำนาจจำแนก /a/ อยู่ระหว่าง 0.82 ถึง 2.18 และสัมประสิทธิ์การเดา /g/ อยู่ระหว่าง 0.17 ถึง 0.28 Testlet 3: Reading for Info & Argument ข้อที่ผ่านเกณฑ์มีจำนวน 4 ข้อ ได้แก่ 21, 23, 24, 25 มีค่าพารามิเตอร์ความยาก /b/ อยู่ระหว่าง -0.29 ถึง 1.02 ค่าพารามิเตอร์อำนาจจำแนก /a/ อยู่ระหว่าง 1.66 ถึง 2.48 และสัมประสิทธิ์การเดา /g/ อยู่ระหว่าง 0.15 ถึง 0.24 Testlet 4: Reading Literature ข้อที่ผ่านตรวจสอบคุณภาพ จำนวน 5 ข้อ ได้แก่ 35, 36, 38, 39, 40 มีค่าพารามิเตอร์ความยาก /b/ อยู่ระหว่าง -0.11 ถึง 1.19 ค่าพารามิเตอร์อำนาจจำแนก /a/ อยู่ระหว่าง 0.87 ถึง 2.50 และสัมประสิทธิ์การเดา /g/ อยู่ระหว่าง 0.12 ถึง 0.23 ดังตาราง 25

ตาราง 25 แสดงผลการตรวจสอบคุณภาพข้อคำถามรายข้อด้วยทฤษฎีการตอบสนองข้อสอบ (IRT 3PL)

Testlet	ข้อเดิมที่ผ่าน CTT	ข้อใหม่ ที่ผ่าน IRT	/b/	/a/	/g/	ข้อใหม่	จำนวนข้อ
T1: Reading Instruction	1	1	0.81	2.46	0.22	1	5
	4	3	-	1.97	0.19	2	
	5	4	0.43	1.96	0.18	3	
	6	5	0.21	1.83	0.20	4	
	7	6	-	2.15	0.22	5	
T2: Reading for Orientation & Reading Correspondence	14	11	0.46	2.00	0.28	6	4
	18	14	0.89	1.82	0.17	7	
	19	15	1.13	0.82	0.19	8	
	20	16	-	2.18	0.18	9	
T3: Reading for info & Argument	21	17	-	2.48	0.24	10	4
	23	18	-	2.14	0.16	11	
	24	20	0.62	1.66	0.24	12	

Testlet	ข้อเดิมที่ ผ่าน CTT	ข้อใหม่ ที่ ผ่าน IRT	/b/	/a/	/g/	ข้อ ใหม่	จำนวน ข้อ
	25	21	1.02	2.09	0.15	13	
T4: Reading Literature	35	26	0.82	2.50	0.23	14	5
	36	27	0.64	1.59	0.20	15	
	38	29	-	1.47	0.15	16	
			0.11				
	39	30	1.19	0.87	0.14	17	
	40	31	-	1.72	0.12	18	
			0.02				

ในการศึกษาครั้งนี้ ผู้วิจัยดำเนินการสร้างแบบสอบแบบทดสอบแบบทดสอบที่มีลักษณะสำคัญคือ ข้อคำถามในชุดทดสอบเดียวกันมีการใช้บทอ่านหรือสถานการณ์เดียวกันในการตอบข้อสอบ โดยกำหนดจำนวนของชุดทดสอบเท่ากับเนื้อหาอิงตามกรอบมาตรฐาน CEFR ระดับ A2 (Beginner) แต่ละบทอ่านมีข้อคำถามที่เกี่ยวข้องกับบทอ่านเป็นเซตเดียวกัน อาจมีจำนวนข้อคำถาม 3-5 ข้อ หรือบางครั้ง 6-12 ข้อ ซึ่งบทอ่านควรจะมีควมยาวประมาณ 300 คำหรือมากกว่า (Wainer, Bradlow, & Wang, 2007, p. 60, อ้างถึงใน พิสิฐ พิณิจสกุล, 2563, น.77; Wheadon, 2010, p. 106; อริสพา เตหลิ้ม, 2559, น.2; จงกล บัวแก้ว, 2560, หน้า 4) ตาราง 26 แสดงให้เห็นว่า ผลการถูกต้องออกของข้อคำถามจำนวน 9 ข้อ ในการตรวจสอบด้วยทฤษฎีการวัดแบบดั้งเดิม และจำนวน 15 ข้อในการตรวจสอบคุณภาพรายข้อด้วยทฤษฎีการตอบสนองข้อสอบ ไม่ส่งผลกระทบต่อทางด้านเนื้อหาที่ใช้ประเมินความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 มีรายละเอียดในตาราง 26 ดังนี้

ตาราง 26 แสดงจำนวนชุดทดสอบ เนื้อหาตามกรอบ CEFR ชื่อเรื่อง และจำนวนข้อคำถามของแบบสอบ

Testlets	เนื้อหาอิงกรอบ CEFR ระดับ A2	Topic	แบบ สอบ ร่าง 1	แบบ สอบร่าง 2	แบบ สอบ ใช้ จริง
Testlets 1	Reading Instruction	The Instruction of Vending Machine	10	8	5
Testlets	Reading Orientation &	The Advertisement of	10	8	4

2	Reading Correspondence	Healthy Workshop			
Testlets	Reading for Information &	The News about COVID-	12	9	4
3	Argument	19			
Testlets	Read Literature	The Story of Thai Culture	10	8	5
4					
รวมทั้งสิ้น (ข้อ)			42	33	18


จากกระบวนการพัฒนาแบบสอบในข้างต้น ทำให้ได้แบบสอบแบบเทสต์เลทเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ที่วัดความสามารถการอ่านเพื่อความเข้าใจตามเนื้อหาอิงกรอบมาตรฐาน CEFR ระดับ A2 ประกอบด้วย 4 เทสต์เลท มีข้อคำถามจำนวนทั้งหมด 18 ข้อ แต่ละข้อคำถามมุ่งวัด 2 ถึง 4 คุณลักษณะ โดยแต่ละคุณลักษณะจะมีจำนวนข้อคำถามไม่เท่ากัน แต่ครอบคลุมทุกคุณลักษณะ ดังตาราง 27



ตาราง 27 ผลการสร้างแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่าน
เพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3

Topic 1: READ INSTRUCTIONS

Read the following text and these pictures for answering items 1-5

	Items	Attributes					
		A1	A2	A3	A4	A5	A6
							6
1	<p>1. Rearrange these sentences according to the correct order of the pictures</p> <p>(A) Check the coin chute for any change.</p> <p>(B) Swipe your credit card if it's a newer machine.</p> <p>(C) Insert the correct amount of money for the item.</p> <p>(D) Type in the code or press the button for your item.</p> <p>(E) Find the price and code underneath the item you wish to buy.</p> <p>A vending machine is an automated machine that provides items such as snacks, beverages, and lottery tickets to consumers after cash, a credit card, or other forms of payment are inserted into the machine or otherwise made. The first modern vending machines were developed in England in the 1880s and dispensed postcards. Vending machines exist in many countries and, in more recent years, specialized vending machines that provide common products compared to traditional vending machines have been created.</p>  <p>Picture 1 Picture 2 Picture 3 Picture 4 Picture 5</p> <p><input type="radio"/> C-B-D-A-E</p> <p><input type="radio"/> E-C-B-D-A</p> <p><input type="radio"/> B-D-C-E-A</p> <p><input type="radio"/> E-B-A-C-D</p>	1	1	0		1	0
					0		
							6
2	<p>4. What will you do if you have no banknote? *</p> <p><input type="radio"/> cancel some orders</p> <p><input type="radio"/> pay by cheque</p> <p><input type="radio"/> use your credit card</p> <p><input type="radio"/> borrow friend's money</p>	1	1	0	0	1	0

	Items	Attributes					
		A1	A2	A3	A4	A5	A6
3		1	0	0	0	1	0

5. According to the pictures, what should be the first step of using a vending machine? * 1 answer



- Check the coin chute for any change.
- Type in the code or press the button for your item.
- Insert the correct amount of money for the item.
- Find the price and code underneath the item you wish to buy.

	Items	Attributes					
		A1	A2	A3	A4	A5	A6
4	6. Where <u>can't</u> you generally find the vending machine? *	1	0	0	0	1	0

A vending machine is an automated machine that provides items such as snacks, beverages, cigarette and lottery tickets to consumers after cash, a credit card, or other forms of payment are inserted into the machine or otherwise made. The first modern vending machines were developed in England in the early 1880s and dispensed postcards. Vending machines exist in many countries and, in more recent time, specialized vending machines that provide common products compared to traditional vending machines have been created.

- farm
- airport
- bus terminal
- department store

5 7. Where were the first modern vending machines developed? * 1 1 0 0 0 0

A vending machine is an automated machine that provides items such as snacks, beverages, cigarettes and lottery tickets to consumers after cash, a credit card, or other forms of payment are inserted into the machine or otherwise made. The first modern vending machines were developed in England in the early 1880s and dispensed postcards. Vending machines exist in many countries and, in more recent times specialized vending machines that provide common products compared to traditional vending machine items have been created.



- England
- USA.
- Thailand
- Japan

Topic 2: READ FOR ORIENTATION and READ CORRESPONDENCE

Two passages are related, read the advertisement and the e-mail for answering items 6–9

Items

Attributes

A1 A2 A3 A4 A5 A6

6 14. Which is refer to the word "items" ? * 1 0 0 0 1 0

**GENERAL HOSPITAL
UPCOMING WORKSHOPS**

October 8 Keep Your Heart Healthy
October 10 Walk for Fitness
October 12 Delicious Low-Calorie Meals*

Workshops are free unless otherwise noted. All workshops are held from 10:00 A.M. until noon in Conference Room A. We welcome people ages 15 and up. Advance registration is recommended. E-mail workshops@gh.com to register. There is a \$3 fee for same-day registration, which is permitted on a space-available basis.

For the October 10 workshop, wear comfortable shoes. For all the workshops, bring extra money if you would like to purchase any **items** for cooking healthy food, we will offer 20% discount.

*This workshop has a \$10 materials fee to cover the cost of ingredients. Participants will cook and eat a three-course meal.

- Shoes
- Workshop
- Recipe book
- Healthy food

Items

Attributes

A1 A2 A3 A4 A5 A6

7 18. According to the e-mail, which is NOT true? 0 0 1 0 1 1



- The most popular workshop was October 10.
- Sixty-five people attended the October workshop.
- The meals workshop fee was two dollars.
- The e-mail was sent in January at 12 p.m.

Items	Attributes					
	A1	A2	A3	A4	A5	A6

8 19. What dose the word "you" refer to? 0 0 0 1 1 0



- Sunny Mara
- Kenny A.
- Children
- People

Items	Attributes					
	A1	A2	A3	A4	A5	A6

9 20. How many people showed up for Delicious Low-Calorie Meals? * 0 points 1 1 0 0 0 0



- Twenty-two
- Thirty-two
- Twenty-four
- Thirty-four

Topic 3: READ INFORMATIONS & ARGUMENT
Read the following news for answering items 10-13

Items

Attributes

A	A	A	A	A	A
1	2	3	4	5	6

1 0 Read the news 21. What dose the word "it" refer to? 0 points 1 0 1 1 0 0

(1) COVID-19 is an infectious disease that causes illness in the respiratory system in humans. The term Covid-19 is sort of an acronym, derived from "Corona Virus Disease 2019". It is a new virus that is spreading primarily through contact with the person. We need to control the contagion of it. Although the majority of people develop symptoms within 14 days of getting the virus, in some people, it can take longer. The common symptoms of the viral infection are fever, cold, cough, muscle pain, and respiratory problems.

- A cold
- A cancer
- COVID-19
- A chicken pox

Items

Attributes

A	A	A	A	A	A
1	2	3	4	5	6

1	23. According to the passage, which is NOT the symptoms of COVID-19?	1	1	0	1	0	0
---	--	---	---	---	---	---	---

(1) COVID-19 is an infectious disease that causes illness in the respiratory system in humans. The term Covid-19 is sort of an acronym, derived from "Corona Virus Disease 2019". It is a new virus that is spreading primarily through contact with the person. We need to control the contagion of it. Although the majority of people develop symptoms within 14 days of getting the virus, in some people, it can take longer. The possible symptoms of the viral infection are muscle pain, lungs and respiratory problems.

- an ear pain
- a cough
- a sore throat
- difficulty breathing

Items		Attributes					
		A	A	A	A	A	A
		1	2	3	4	5	6

1	24. What is the second paragraph mainly about? *	0	0	0	0	0	0
---	--	---	---	---	---	---	---

(2) Now that COVID-19 vaccines have reached billions of people worldwide, the evidence is overwhelming that no matter which one you take. Developing immunity through vaccination helps you fight the virus if exposed. Getting vaccinated may also protect people around you, because if you are protected from getting infected and from disease, you are less likely to infect someone else. Washing your hands often, social distancing, and wearing a face mask as ways to help protect yourself and prevent the spread of COVID-19.

- Facts about COVID-19
- The cause of COVID-19
- The impact of COVID-19
- The prevention of COVID-19

Items		Attributes					
		A	A	A	A	A	A
		1	2	3	4	5	6

1	25. What is the best way for stopping the epidemic in Thailand? *	1	0	0	1	1	1
---	---	---	---	---	---	---	---

- Wearing a face mask
- Keeping social distancing
- Washing our hands often
- Getting vaccine against COVID-19

Topic 4: READ LITERATURE

Read the following story about culture for answering items 14-18

Items		Attributes					
-------	--	------------	--	--	--	--	--

		A	A	A	A	A	A
		1	2	3	4	5	6

1 35. Which of the following are NOT the activities people have done? * 1 rest 1 1 0 0 1 0

4

The festival of Phi Ta Khon has become one of the most fun celebrations and a more significant part of Bun Pawet, a Buddhist merit-making festival. During the 3-day festival, townspeople dress up in ghost masks - which is also the actual translation of the festival name - and make merry by participating in parties and parades on the Dan Sai district in the Loei province. The colorful masks are a sight for sore eyes, and can even get pretty scary at times. Skillfully **handcrafted** from dried sticky husks, these masks are generally white and have long noses.

- Making a merit
- Making 'trick or treat'
- Making a costume and mask
- Make merry by participating in parades

	Items	Attributes					
		A	A	A	A	A	A
		1	2	3	4	5	6

1 36. What does the word "handcrafted" in paragraph 3 mean? * 1 rest 1 0 1 0 0 1

5

The festival of Phi Ta Khon has become one of the most fun celebrations and a more significant part of Bun Pawet, a Buddhist merit-making festival. During the 3-day festival, townspeople dress up in ghost masks - which is also the actual translation of the festival name - and make merry by participating in parties and parades on the Dan Sai district in the Loei province. The colorful masks are a sight for sore eyes, and can even get pretty scary at times. Skillfully **handcrafted** from dried sticky husks, these masks are generally white and have long noses.

- To made carefully by agent
- To made carefully by expert
- To made skillfully by hand
- To made skillfully by machine

	Items	Attributes					
		A	A	A	A	A	A
		1	2	3	4	5	6

1 38. What does the word "catch the eye of" in paragraph 4 mean? * 1 0 1 0 1 1

6

The primary attention of every festival-goer would fall on the great monstrous mask, with terrifying teeth and evil eyes. Legend says that it is meant bad luck to keep your mask after the festival gets over, but as times have changed, people now sell their masks and make good money out of it. The masks **catch the eye of** tourists, and thus, the prices can go as high as thousands of Baht. The local population of Loei is always friendly and welcoming. They will be happy to share their culture with you and make sure that you are an integral part of their celebrations. You will also be excited to pose for photos so keep that camera ready to catch the excitement as this is a memory you wouldn't want to forget ever.

Link : <https://www.holiday.com/pages/phi-ta-khon-1215.html>

- Attract the attention of someone
- Concentrate on something
- Look at someone
- Close the eyes

Items	Attributes					
	A	A	A	A	A	A
	1	2	3	4	5	6

1 39. What is the writer's purpose in this paragraph? * 0 1 0 0 1 1 1

7

The primary attention of every festival-goer would fall on the great monstrous mask, with terrifying teeth and evil eyes. Legend says that it is meant bad luck to keep your mask after the festival gets over, but as times have changed, people now sell their masks and make good money out of it. The masks **catch the eye of** tourists, and thus, the prices can go as high as thousands of Baht. The local population of Loei is always friendly and welcoming. They will be happy to share their culture with you and make sure that you are an integral part of their celebrations. You will also be excited to pose for photos so keep that camera ready to catch the excitement as this is a memory you wouldn't want to forget ever.

Link : <https://www.holiday.com/pages/phi-ta-khon-1215.html>

- To describe the story of ghost
- To persuade people for visiting
- To tell how important of festival
- To report the news about festival

Items	Attributes					
	A	A	A	A	A	A
	1	2	3	4	5	6

1 40. Which is the best title for this passage? * 1 0 0 1 0 0

8

- Mask Festival
- Phi Ta Khon in Loei
- Skillfully handcrafted
- The story of Halloween

ตอนที่ 2 ผลการทดลองและวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และผลการออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ

ผู้วิจัยนำเสนอผลการวิเคราะห์ข้อมูลในตอนต้นที่ 2 แบ่งออกเป็น 3 ส่วน คือ 1) ผลการทดลองใช้และตรวจสอบคุณภาพของแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ ด้วยโมเดลการวินิจฉัย Generalization DINA 2) ผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ 3) ผลการออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ ดังนี้

ส่วนที่ 1 ผลการทดลองใช้และตรวจสอบคุณภาพของแบบสอบด้วยโมเดลการวินิจฉัย G-DINA

ผลการทดลองใช้และตรวจสอบคุณภาพของแบบสอบวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษด้วยโมเดลการวินิจฉัย G-DINA โดยนำเสนอผลการวิเคราะห์ข้อมูลแยกออกเป็น 5 หัวข้อย่อย ได้แก่ 1) ผลการวิเคราะห์ข้อมูลเชิงบรรยายของการตรวจสอบคุณภาพของแบบสอบวินิจฉัย 2) ผลการตรวจสอบความตรงของเมตริกซ์คิว (Q-matrix validation) 3) ผลการวิเคราะห์ประสิทธิภาพของความสอดคล้องเชิงสัมพัทธ์ของ G-DINA Model 4) ผลการวิเคราะห์ค่าระดับความแม่นยำในการจำแนกของแบบสอบรายฉบับ (Test Level Accuracy) ค่าระดับความแม่นยำในการจำแนกรายคุณลักษณะ (Attribute Level Accuracy) และ 5) ผลการวิเคราะห์คุณภาพของข้อคำถามรายข้อด้วยโมเดลการวินิจฉัย G-DINA ดังนี้

1) ผลการวิเคราะห์ข้อมูลเชิงบรรยายของการตรวจสอบคุณภาพของแบบสอบวินิจฉัย

ผลการวิเคราะห์ข้อมูลเชิงบรรยายของการตรวจสอบคุณภาพของแบบสอบแบบทดสอบเพื่อวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 18 ข้อ คะแนนเต็ม 18 คะแนน ผลการตอบแบบสอบวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจของนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 396 คน พบว่า นักเรียนตอบแบบสอบวินิจฉัยได้คะแนนสูงสุดเท่ากับ 17 คะแนน จำนวน 6 คน คิดเป็นร้อยละ 1.52 นักเรียนตอบแบบสอบวินิจฉัยได้คะแนนต่ำสุด 3 คะแนน จำนวน 10 คน คิดเป็นร้อยละ 2.53 นักเรียนส่วนใหญ่ตอบแบบสอบวินิจฉัยได้ 5 คะแนน คิดเป็นร้อยละ 14.39 แสดงในตาราง 28

ตาราง 28 แสดงผลการตอบแบบสอบถามวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจ (คะแนนเต็ม 18)

คะแนน	ความถี่	ร้อยละ
3	10	2.53
4	32	8.08
5	57	14.39
6	54	13.64
7	34	8.59
8	33	8.33
9	25	6.31
10	22	5.56
11	17	4.29
12	19	4.8
13	20	5.05
14	27	6.82
15	28	7.07
16	12	3.03
17	6	1.52
Mean = 8.76		S.D.= 3.90

2) ผลการตรวจสอบความตรงของเมตริกซ์คิว (Q-matrix validation)

ผู้วิจัยทำการตรวจสอบความตรงของเมตริกซ์คิว (Q-matrix validation) ด้วยหลักฐานเชิงประจักษ์ โดยวิธี Stepwise Wald test (Ma & de la Torre, 2019) ด้วยโมเดลวินิจฉัย G-DINA Model ได้เมตริกซ์คิวแนะนำ (Suggested Q-matrix) ผลปรากฏว่า จากการตรวจสอบความตรงของเมตริกซ์คิว (Q-matrix validation) ด้วยหลักฐานเชิงประจักษ์ โดยวิธี Stepwise Wald test (Ma & de la Torre, 2019) ได้เมตริกซ์คิวแนะนำ (Suggested Q-matrix) ให้ทำการปรับ จำนวน 10 ข้อ ได้แก่ ข้อ 1, 2, 7, 8, 10, 13, 14, 15, 16 และ 17 โดยข้อคำถามที่ถูกแนะนำให้ปรับมีจำนวน 1-3 คุณลักษณะ แสดงในตาราง 29 ดังนี้

ตาราง 29 แสดงเมตริกซ์ควิแนะนำ (Suggested Q-matrix) โดยวิธี Stepwise Wald Test

ข้อ คำถามที่	เมตริกซ์ควิแนะนำ (Suggested Q-matrix)						ความหมาย
	Vocab (A1)	Explicit info (A2)	Syntax (A3)	Main idea (A4)	Inference (A5)	Synthesize (A6)	
1	0*	1	0	0	1	0	ปรับ 1 คุณลักษณะ
2	1	0*	0	0	1	0	ปรับ 1 คุณลักษณะ
3	1	0	0	0	1	0	คงเดิม
4	1	0	0	0	1	0	คงเดิม
5	1	1	0	0	0	0	คงเดิม
6	1	0	0	0	1	0	คงเดิม
7	0	0	1	0	0*	1	ปรับ 1 คุณลักษณะ
8	1	0	1	0*	1	0	ปรับ 1 คุณลักษณะ
9	1	1	0	0	0	0	คงเดิม
10	1	0	1	0*	0	0	ปรับ 1 คุณลักษณะ
11	1	1	0	1	0	0	คงเดิม
12	1	0	0	1	0	1	คงเดิม
13	1	0	0	1	1	0*	ปรับ 1 คุณลักษณะ
14	1	1*	1	0	1	0	ปรับ 1 คุณลักษณะ
15	1	0	0*	0	0	1	ปรับ 1 คุณลักษณะ
16	1	0	1	1*	0*	0	ปรับ 2 คุณลักษณะ
17	1	0	0	0*	0*	0*	ปรับ 3 คุณลักษณะ

ข้อ	เมตริกซ์คิวแนะนำ (Suggested Q-matrix)						ความหมาย
18	1	0	0	1	0	0	คงเดิม

ผู้วิจัยทำการปรับเมตริกซ์คิว (Q-matrix) ของข้อคำถามจำนวน 9 ข้อ ตามเมตริกซ์คิวแนะนำ (Suggested Q-matrix) เพื่อให้ได้เมตริกซ์คิว (Q-matrix) ที่มีความตรง (Q-matrix validation) ไปใช้ในการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ยกเว้นข้อ 7, ข้อ 8, ข้อ 14 และข้อ 17 มีรายละเอียด ดังนี้ ข้อที่ 7 ผู้วิจัยทำการปรับ Q-matrix เพียง 2 ใน 4 โดยปรับจาก 0 เป็น 1 เนื่องจากผู้เชี่ยวชาญ 5 ท่านและอาจารย์ที่ปรึกษาวิทยานิพนธ์มีความเห็นว่าถูกต้องเหมาะสมควรมีการปรับ ข้อที่ 8 ผู้วิจัยทำการปรับ Q-matrix จำนวน 2 ใน 3 โดยปรับจาก 0 เป็น 1 เนื่องจากผู้เชี่ยวชาญ 5 ท่านและอาจารย์ที่ปรึกษาวิทยานิพนธ์มีความเห็นว่าถูกต้องเหมาะสมควรมีการปรับ ข้อ 14 ผู้วิจัยทำการปรับ 1 ใน 3 ที่โปรแกรมแนะนำ เนื่องจากการปรับจาก 0 เป็น 1 เนื่องจากผู้เชี่ยวชาญ 5 ท่านและอาจารย์ที่ปรึกษาวิทยานิพนธ์มีความเห็นว่าถูกต้องเหมาะสมควรมีการปรับ และข้อ 17 ผู้วิจัยทำการปรับเพียง 1 ใน 3 ของคุณลักษณะที่แนะนำให้ปรับ เนื่องจากผู้เชี่ยวชาญจำนวน 5 ท่านมีความคิดเห็นว่าข้อคำถาม 1 ข้อสามารถวัดทักษะ/คุณลักษณะของความสามารถได้มากกว่า 1 คุณลักษณะ สอดคล้องกับ Senarat, S. et al., 2013, อ้างถึงใน พิสิฐ พิณจสกุล, 2563) กล่าวว่า ลักษณะเฉพาะของข้อสอบแบบทดสอบ (Testlet-test) สามารถวัดคุณลักษณะหลายด้านในข้อเดียวกัน ในตาราง 4.9 แสดงให้เห็นการเปรียบเทียบเมตริกซ์คิว (Q-matrix) เดิมจากการประเมินของผู้เชี่ยวชาญ Initial Q-matrix เมตริกซ์คิวแนะนำ (Suggested Q-matrix) โดยวิธี Stepwise Wald Test และสรุปผลการปรับได้เมตริกซ์คิวสุดท้ายที่นำไปใช้จริง (Final Q-matrix)

ตาราง 30 ผลการตรวจสอบความตรงของเมตริกซ์คิวจากการทดลองวินิจฉัยด้วยโมเดล G-DINA MODEL

		Initial Q-matrix						Suggested Q-matrix						Final Q-matrix						
ite	A	A	A	A	A	A	ite	A	A	A	A	A	A	ite	A	A	A	A	A	A
m	1	2	3	4	5	6	m	1	2	3	4	5	6	m	1	2	3	4	5	6
1	1	1	0	0	1	0	1	0*	1	0	0	1	0	1	0*	1	0	0	1	0
2	1	0	0	0	1	0	2	1	0	0	0	1	0	2	1	0	0	0	1	0
3	1	0	0	0	1	0	3	1	0	0	0	1	0	3	1	0	0	0	1	0
4	1	1	0	0	0	0	4	1	0	0	0	1	0	4	1	0	0	0	1	0
5	1	0	0	0	1	0	5	1	1	0	0	0	0	5	1	1	0	0	0	0
6	1	1	0	1	0	1	6	1	0	0	0	1	0	6	1	0	0	0	1	0
7	1	0	1	1	1	0	7	0*	0	1	0*	0*	1*	7	1	0	1	0	0*	1*

8	1	1	0	0	0	0	8	1	0*	1*	0	1*	0	8	1	1	1*	0	1*	0
9	1	0	1	1	0	0	9	1	1	0	0	0	0	9	1	1	0	0	0	0
10	1	0	0	1	0	1	10	1	0	1	0*	0	0	10	1	0	1	0*	0	0
11	1	0	0	1	1	1	11	1	1	0	1	0	0	11	1	1	0	1	0	0
12	1	0	0	1	1	1	12	1	0	0	1	0	1	12	1	0	0	1	0	1
13	1	1	0	0	1	0	13	1	1	0	1*	1	0	13	1	1	0	1*	1	0
14	1	0	1	0	0	1	14	1	1*	1	0	1*	0*	14	1	1*	1	0	0	1
15	1	0	1	0	1	0	15	1	0	0*	0	0	1	15	1	0	0*	0	0	1
16	1	0	0	1	1	1	16	1	0	1	1*	0*	0	16	1	0	1	1*	0*	0
17	1	0	0	0	1	1	17	1	0	0	0*	0*	0*	17	1	0	0	0*	1	1
18	1	1	0	1	1	1	18	1	0	0	1	0	0	18	1	0	0	1	0	0

3) ผลการวิเคราะห์ประสิทธิภาพด้านความสอดคล้องเชิงสัมพัทธ์ของโมเดล

(Test Level Fit Statistics)

การวิเคราะห์ประสิทธิภาพความสอดคล้องเชิงสัมพัทธ์ของโมเดลการวินิจฉัย Generalization DINA (G-DINA MODEL) แสดงด้วยค่าดัชนีความสอดคล้องเชิงสัมพัทธ์ของโมเดล (Test Level Fit Statistics) ค่าสถิติทดสอบความสอดคล้องของข้อสอบจำแนกเป็นรายด้าน ผู้วิจัย ดำเนินการใน 2 ขั้นตอน ดังนี้

3.1 การวิเคราะห์ประสิทธิภาพของโมเดล ใช้เกณฑ์การเปรียบเทียบระหว่าง 2 โมเดล คือ CDM package กับ G-DINA package

ผู้วิจัยใช้เกณฑ์การเปรียบเทียบระหว่าง 2 โมเดล คือ CDM package กับ G-DINA package ได้แก่ -2Loglikelihood (-2LL) ของโมเดล CDM มีค่าเท่ากับ - 4782.67 และโมเดล G-DINA มีค่าเท่ากับ - 4754.90 โดยค่าที่ต่ำกว่าแสดงให้เห็นถึงประสิทธิภาพในการประมาณค่าของโมเดลที่ดีกว่า เมื่อพิจารณาค่าสถิติทดสอบความสอดคล้องอีก 2 ตัว ของ G-DINA MODEL ได้แก่ Akaike's Information Criterion (AIC) มีค่าเท่ากับ 9923.80 และค่า Bayesian Information Criterion (BIC) มีค่าเท่ากับ 10792.04 จึงสรุปได้ว่า G-DINA MODEL เป็นโมเดลที่มีประสิทธิภาพในการประมาณค่าของโมเดลที่ดีกว่า มีประสิทธิภาพความสอดคล้องเชิงสัมพัทธ์กับข้อมูลเชิงประจักษ์ สามารถวินิจฉัยความสามารถของผู้สอบได้ ดังตาราง 31

ตาราง 31 แสดงผลการวิเคราะห์เปรียบเทียบประสิทธิภาพความสอดคล้องเชิงสัมพัทธ์ของโมเดลการวินิจฉัย G-DINA (Test Level Fit Statistics)

ค่าดัชนีบ่งชี้ประสิทธิภาพความสอดคล้องเชิงสัมพัทธ์	โมเดลก่อนปรับ Q-matrix	
	CDM	G-DINA

-2Loglikelihood (-2LL)	- 4782.67	- 4754.90
Akaike's information criterion (AIC)	9897	9923.80
Bayesian information criterion (BIC)	10594	10792.04

3.2 การวิเคราะห์ประสิทธิภาพความสอดคล้องเชิงสัมพันธ์ของโมเดล (Test Level Fit Statistics) หลังจากที่ถูกวิจัยทำการปรับ Q-matrix ด้วยวิธี Step Wise Wald Test

การวิเคราะห์ประสิทธิภาพความสอดคล้องเชิงสัมพันธ์ของโมเดลการวินิจฉัย Generalization DINA (G-DINA MODEL) แสดงด้วยค่าดัชนีความสอดคล้องเชิงสัมพันธ์ของโมเดล (Test Level Fit Statistics) ค่าสถิติทดสอบความสอดคล้องของข้อสอบจำแนกเป็นรายด้าน ใช้เกณฑ์การเปรียบเทียบระหว่าง 2 โมเดล คือ CDM package กับ G-DINA package ได้แก่ -2Loglikelihood (-2LL) CDM มีค่าเท่ากับ -4098.77 และ G-DINA มีค่าเท่ากับ -4048.34 โดยค่าที่ต่ำกว่าแสดงให้เห็นถึงประสิทธิภาพในการประมาณค่าของโมเดลที่ดีกว่า เมื่อพิจารณาค่าสถิติทดสอบความสอดคล้องอีก 2 ตัว ของ G-DINA MODEL ได้แก่ Akaike's Information Criterion (AIC) มีค่าเท่ากับ 8590.69 และค่า Bayesian Information Criterion (BIC) มีค่าเท่ากับ 9574.10 จึงสรุปได้ว่า แม้จะทำการปรับ Q-matrix โมเดล G-DINA MODEL ยังคงเป็นโมเดลที่มีประสิทธิภาพในการประมาณค่าของโมเดลที่ดีกว่า มีประสิทธิภาพความสอดคล้องเชิงสัมพันธ์กับข้อมูลเชิงประจักษ์สามารถวินิจฉัยความสามารถของผู้สอบได้ ดังตาราง 32

ตาราง 32 แสดงผลการวิเคราะห์เปรียบเทียบประสิทธิภาพความสอดคล้องเชิงสัมพันธ์ของโมเดลการวินิจฉัย G-DINA (Test Level Fit Statistics)

ค่าดัชนีประสิทธิภาพความสอดคล้องเชิงสัมพันธ์	โมเดลหลังปรับ Q-matrix	
	CDM	G-DINA
-2Loglikelihood (-2LL)	-4098.77	-4048.34
Akaike's information criterion (AIC)	8610	8590.69
Bayesian information criterion (BIC)	9430	9574.10

4) ผลการวิเคราะห์ค่าระดับความแม่นยำในการจำแนกของแบบสอบรายฉบับ (Test Level Accuracy) ค่าระดับความแม่นยำในการจำแนกรายคุณลักษณะ (Attribute Level Accuracy)

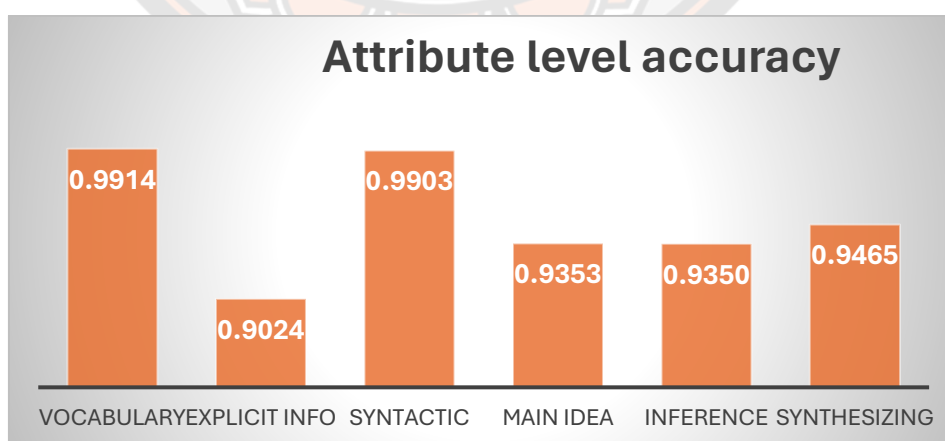
4.1 ผลการวิเคราะห์ค่าระดับความแม่นยำในการจำแนกของแบบสอบรายฉบับ (Test Level Accuracy) ด้วยโมเดลวินิจฉัย Generalization DINA (G-DINA MODEL) พบว่าแบบสอบแบบทดสอบเพื่อวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจที่ผู้วิจัยพัฒนาขึ้น สามารถ

จำแนกผู้สอบเข้าชั้นแฝงตามคุณลักษณะความสามารถรายบุคคลได้อย่างถูกต้อง เท่ากับ 0.82 โดยมีค่ามากกว่า 0.5 และมีค่าเข้าใกล้ 1.0

ค่าระดับความแม่นยำในการจำแนกรายคุณลักษณะ (Attribute Level Accuracy) ด้วยโมเดลวินิจฉัย Generalization DINA (G-DINA MODEL) จำนวน 6 คุณลักษณะ พบว่า แบบสอบมีค่าระดับความแม่นยำในการจำแนกรายคุณลักษณะอยู่ระหว่าง 0.90 – 0.99 โดยที่คุณลักษณะที่ 1 (A1 มีความรู้คำศัพท์) และคุณลักษณะที่ 3 มีความรู้ทางไวยากรณ์ มีค่าความถูกต้องแม่นยำสูงที่สุด เท่ากับ 0.99 รองลงไปคือคุณลักษณะที่ 6 มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน มีค่าเท่ากับ 0.95 คุณลักษณะที่ 4 มีความสามารถในการระบุใจความสำคัญของเรื่อง และคุณลักษณะที่ 5 มีความสามารถในการอนุมานมีค่าเท่ากับ 0.94 คุณลักษณะที่ 2 มีความสามารถในการระบุข้อมูลที่ต้องการ มีค่าเท่ากับ 0.90 ตามลำดับ ซึ่งมีค่าใกล้ 1.0 ทุกคุณลักษณะดังตาราง 33 และภาพ 46

ตาราง 33 ค่าระดับความแม่นยำในการจำแนกของแบบสอบรายฉบับและรายคุณลักษณะ

Reading Comprehension	Test Level Accuracy	Attribute Level Accuracy					
		A1	A2	A3	A4	A5	A6
	0.82	0.99	0.90	0.99	0.94	0.94	0.95



ภาพ 46 แสดงระดับความเที่ยงตรงเชิงจำแนกรายคุณลักษณะที่ใช้ในการวินิจฉัย

4.2 ผลการตรวจสอบคุณภาพด้านความตรงเชิงจำแนกและความเที่ยงเชิงจำแนกรายคุณลักษณะของแบบสอบด้วยโมเดลวินิจฉัย G-DINA (Class Accuracy)

ผลการตรวจสอบคุณภาพด้านความตรงในการจำแนกและความเที่ยงเชิงจำแนกของแบบสอบด้วยโมเดลวินิจฉัย G-DINA (Class Accuracy) ด้วยวิธีการประมาณค่าความน่าจะเป็นสูงสุด (Maximum Likelihood Estimation: MLE) ผลปรากฏว่า แบบสอบมีความตรงในการจำแนกรายคุณลักษณะอยู่ระหว่าง 0.882 ถึง 0.982 ซึ่งมีค่าใกล้เคียง 1.0 หมายความว่า แบบสอบมีคุณภาพด้านความตรงเชิงจำแนก และมีค่าความเที่ยงเชิงจำแนกรายคุณลักษณะ อยู่ระหว่าง 0.809 ถึง 0.965 ซึ่งมีค่าใกล้เคียง 1.0 หมายความว่า แบบสอบมีคุณภาพด้านความเที่ยงเชิงจำแนกรายคุณลักษณะ โดยมีรายละเอียดในตาราง 34

ตาราง 34 แสดงผลการวิเคราะห์ความตรงและความเที่ยงเชิงจำแนกของแบบสอบรายคุณลักษณะ

คุณลักษณะของความสามารถอ่านเพื่อความเข้าใจ	ความตรงเชิงจำแนก	ความเที่ยงเชิงจำแนก
A1 มีความรู้คำศัพท์	0.982	0.965
A2 มีความสามารถในการระบุข้อมูลที่ต้องการ	0.882	0.809
A3 มีความรู้ทางไวยากรณ์	0.894	0.839
A4 มีความสามารถในการระบุใจความสำคัญของเรื่อง	0.962	0.930
A5 มีความสามารถในการอนุมาน	0.925	0.869
A6 มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน	0.899	0.829

5) ผลการวิเคราะห์คุณภาพของข้อคำถามรายข้อด้วยโมเดลการวินิจฉัย G-DINA ในการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยโมเดล G-DINA ในการศึกษาครั้งนี้ ผู้วิจัยนำเสนอผลการวิเคราะห์ค่าสถิติ ได้แก่ ค่าพารามิเตอร์การเดา พารามิเตอร์ความสะเพร่า ค่าดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ของข้อคำถาม (RMSEA) และค่าดัชนีอำนาจจำแนกของข้อสอบรายข้อ (Item Discrimination Index: IDI) มีรายละเอียด ดังนี้

ค่าพารามิเตอร์การเดา (Guessing Parameter) อยู่ระหว่าง 0.001 – 0.4338 ข้อสอบทุกข้อมีค่าพารามิเตอร์การเดาตามเกณฑ์ คือไม่มีข้อใดที่มีค่าพารามิเตอร์การเดาเกินกว่า 0.5 แสดงว่าข้อสอบทั้ง 18 ข้อ เป็นข้อสอบที่นักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีความสามารถอ่านเพื่อความเข้าใจในระดับต่ำ จะไม่สามารถเดาคำตอบได้ถูกต้อง ค่าพารามิเตอร์ความสะเพร่า (Slipping parameter) อยู่ระหว่าง 0.001 - 0.8893 ข้อสอบจำนวน 17 ข้อ มีค่าพารามิเตอร์ความสะเพร่าตามเกณฑ์ คือไม่มีข้อที่มีค่าพารามิเตอร์ความสะเพร่าเกินกว่า 0.5 ยกเว้นข้อ 7 มี

ค่าพารามิเตอร์ความสะเพร่าเท่ากับ 0.8893 หมายความว่า ผู้สอบมักจะตอบข้อคำถามที่ 7 ผิด เพราะความสะเพร่า

ค่าดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ของข้อคำถามรายข้อ (RMSEA) เมื่อตรวจสอบความเหมาะสมของข้อสอบเป็นรายข้อด้วยดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ (RMSEA) มีค่าระหว่าง 0.032 ถึง 0.169 และเมื่อตรวจสอบโดยใช้เกณฑ์ของ Maydeu – Olivares & Joe (2014, อ้างถึงใน เกศรินทร์ สมราช, 2563, น.188) $RMSEA < 0.089$ หมายถึง ข้อสอบข้อนั้นมีความเหมาะสมสามารถนำไปใช้ได้ และ $RMSEA < 0.05$ หมายถึง ข้อสอบข้อนั้นมีความคลาดเคลื่อนต่ำมากและมีความเหมาะสมเป็นอย่างมากที่จะนำไปใช้วินิจฉัยผู้เรียน ผลปรากฏว่า ค่า RMSEA ของข้อสอบจำนวน 11 ข้อ มีค่าอยู่ระหว่าง 0.032 ถึง 0.081 ซึ่งเป็นไปตามเกณฑ์ ยกเว้น ข้อคำถามจำนวน 7 ข้อ ได้แก่ ข้อ 2, 6, 7, 10, 11, 13 และ 16 โดยมีค่า 0.119, 0.148, 0.109, 0.169, 0.112, 0.114 และ 0.139 ตามลำดับ อย่างไรก็ตาม เมื่อพิจารณาค่าเฉลี่ยของ RMSEA ของข้อสอบ 18 ข้อ มีค่าเท่ากับ 0.086 ซึ่งเป็นไปตามเกณฑ์ สามารถนำไปใช้วินิจฉัยความสามารถผู้เรียนได้

ค่าดัชนีอำนาจจำแนกของข้อสอบรายข้อ (Item Discrimination Index: IDI) เป็นดัชนีที่บ่งชี้ว่าข้อสอบข้อนั้นสามารถจำแนกผู้สอบที่มีความสามารถกับผู้ที่ไม่มีความสามารถได้ดีมากน้อยเพียงใดโดยมีเกณฑ์การผ่าน คือ มีค่า $IDI \geq 0.20$ (de la Torre, 2008, pp. 343 – 362 and de la Torre, Rossi and van der Ark, 2018, pp. 281 – 296, อ้างถึงใน วรางคณา ก้อนน้อย, 2563, น.134) ผลปรากฏว่า ข้อคำถามทั้ง 18 ข้อมีค่าดัชนีอำนาจจำแนกผ่านเกณฑ์ โดยมีค่าดัชนี IDI อยู่ระหว่าง 0.2105 ถึง 0.9238 โดยมีรายละเอียดในตาราง 35 ดังนี้

ตาราง 35 แสดงผลการวิเคราะห์คุณภาพของข้อคำถามด้วยโมเดล G-DINA

ITEM	Guessing	Slip	RMSEA	IDI	สรุปผล
Item1	0.1866	0.0001	0.081	0.8132	นำไปใช้ได้
Item2	0.3010	0.0001	0.119	0.6988	นำไปใช้ได้
Item3	0.1780	0.0522	0.051	0.7698	นำไปใช้ได้
Item4	0.3583	0.3377	0.066	0.3040	นำไปใช้ได้
Item5	0.4338	0.0303	0.054	0.5358	นำไปใช้ได้
Item6	0.3917	0.1836	0.148	0.4246	นำไปใช้ได้
Item7	0.0001	0.8893	0.109	0.2105	นำไปใช้ได้
Item8	0.2133	0.1155	0.052	0.6711	นำไปใช้ได้

ITEM	Guessing	Slip	RMSEA	IDI	สรุปผล
Item9	0.0001	0.0761	0.063	0.9238	นำไปใช้ได้
Item10	0.2944	0.0001	0.169	0.7055	นำไปใช้ได้
Item11	0.2542	0.0001	0.112	0.7457	นำไปใช้ได้
Item12	0.1149	0.0844	0.071	0.8007	นำไปใช้ได้
Item13	0.3294	0.1649	0.114	0.5057	นำไปใช้ได้
Item14	0.3880	0.0001	0.032	0.6119	นำไปใช้ได้
Item15	0.0796	0.1017	0.053	0.8187	นำไปใช้ได้
Item16	0.0001	0.4730	0.139	0.4269	นำไปใช้ได้
Item17	0.2706	0.0001	0.049	0.7293	นำไปใช้ได้
Item18	0.4325	0.0001	0.068	0.5673	นำไปใช้ได้

ส่วนที่ 2 ผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียน
ชั้นมัธยมศึกษาปีที่ 3

ผู้วิจัยนำเสนอผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของ
นักเรียนชั้นมัธยมศึกษาปีที่ 3 แบ่งออกเป็น 3 หัวข้อย่อย คือ 2.1 ผลการวินิจฉัยรายคุณลักษณะ 2.2
ผลการวินิจฉัยรูปแบบความรู้ของนักเรียน (Attribute Profile) และ 2.3 ตัวอย่างผลการวินิจฉัย
รายบุคคล ดังนี้

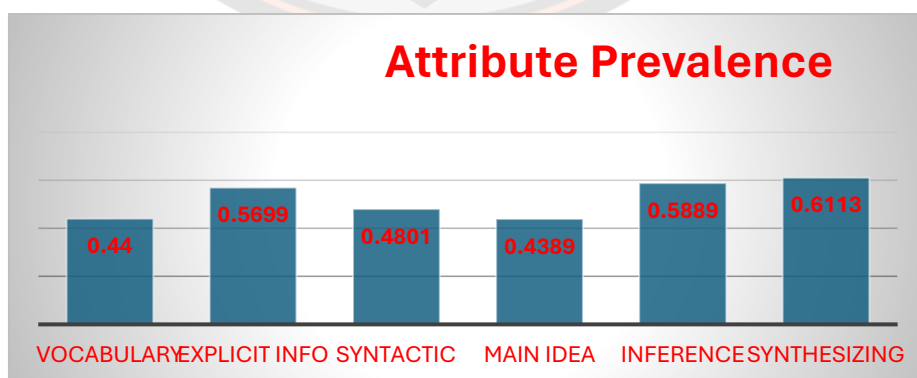
2.1 ผลการวินิจฉัยรายคุณลักษณะ (Attribute Prevalence)

ผลการวินิจฉัยรายคุณลักษณะในภาพรวมของความสามารถการอ่านเพื่อ
ความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ดังนี้ ความรอบรู้อันดับที่ 1 นักเรียน
ระดับชั้นมัธยมศึกษาปีที่ 3 มีความน่าจะเป็นที่จะรอบรู้ (Mastery) ในคุณลักษณะ A6 มี
ความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน ร้อยละ 61.13 อันดับที่ 2 นักเรียนมีความน่าจะเป็น
ที่จะรอบรู้ (Mastery) ในคุณลักษณะ A5 มีความสามารถในการอนุมานร้อยละ 58.89 อันดับที่ 3
นักเรียนมีความน่าจะเป็นที่จะรอบรู้ (Mastery) ในคุณลักษณะ A2 มีความสามารถในการระบุข้อมูลที่
ต้องการ ร้อยละ 56.99 อันดับที่ 4 นักเรียนมีความน่าจะเป็นที่จะรอบรู้ (Mastery) ในคุณลักษณะ A3
มีความรู้ทางไวยากรณ์ ร้อยละ 48.01 รองลงไปนักเรียนมีความน่าจะเป็นที่จะรอบรู้ในคุณลักษณะ A1
มีความรู้คำศัพท์ ร้อยละ 44.00 และอันดับสุดท้าย นักเรียนมีความน่าจะเป็นที่จะรอบรู้ในคุณลักษณะ
A4 มีความสามารถในการระบุใจความสำคัญของเรื่อง ร้อยละ 43.89 ตามลำดับ ในตาราง 36 และ
ภาพ 47 ดังนี้

ตาราง 36 แสดงผลการวินิจฉัยรายคุณลักษณะ (Attribute Prevalence)

คุณลักษณะ	คำอธิบายทักษะความสามารถ	ร้อยละ รอบรู้	อันดับ ความ รอบรู้
A1	มีความรู้คำศัพท์ (Vocabulary Knowledge) พิจารณาจากความเข้าใจเกี่ยวกับความหมายคำศัพท์ และกลุ่มคำโดยอาศัยพื้นความรู้เดิม หรือเข้าใจความหมายของคำศัพท์ที่ยากหรือไม่คุ้นเคยโดยอาศัยการคาดเดาความหมายคำศัพท์จากบริบทหรือคำบ่งชี้ที่พบในข้อความ (Context Clue) ประกอบด้วยการนิยามความหมาย การกล่าวซ้ำ การเปรียบเทียบความเหมือนหรือความต่าง และการยกตัวอย่าง	0.4400	5
A2	มีความสามารถในการระบุข้อมูลที่ต้องการ (Identifying Explicit Information) พิจารณาได้จากทักษะในการค้นหารายละเอียดหรือข้อมูลเฉพาะที่ต้องการซึ่งปรากฏชัดเจนในบทอ่าน หลังจากที่ได้อ่านแบบผ่าน (Scan) และอ่านแบบพินิจพิเคราะห์ (Skim) ทำให้ทราบว่าข้อมูลอยู่ที่ใดของเรื่อง เช่น จำนวน ราคา วันเวลา สถานที่ ซึ่งข้อมูลดังกล่าวจะปรากฏอย่างชัดเจน ผู้อ่านสามารถเข้าใจได้โดยมีสติความ	0.5699	3
A3	มีความรู้ทางไวยากรณ์ (Syntactic Knowledge) พิจารณาได้จากความเข้าใจเกี่ยวกับวากยสัมพันธ์ซึ่งหมายถึงการศึกษาความสัมพันธ์ของคำในประโยคหรือแบบแผนองค์ประกอบของประโยคในภาษาอังกฤษ และความสัมพันธ์ระหว่างประโยคและสามารถวางตำแหน่งของคำหรือกลุ่มคำประเภทต่าง ๆ ในประโยคได้อย่างถูกต้องตามโครงสร้างทางไวยากรณ์	0.4801	4
A4	มีความสามารถในการระบุใจความสำคัญของเรื่อง (Generalization Main Idea) พิจารณาได้จากการทำความเข้าใจเพื่อค้นหาหัวเรื่อง (Topic) นำมารวมกับข้อความที่ถูกควบคุมให้กล่าวถึงสิ่งใดสิ่งหนึ่ง (Controlling Idea) และลงข้อสรุปเป็นใจความสำคัญ หรือแนวคิดหลักที่อยู่ในบทอ่านซึ่งเป็นส่วนที่	0.4389	6

คุณลักษณะ	คำอธิบายทักษะความสามารถ	ร้อยละ รอบรู้	อันดับ ความ รอบรู้
	ครอบคลุมย่อหน้านั้น ๆ หรือเนื้อหาทั้งหมด		
A5	มีความสามารถในการอนุมาน(Getting Inference) พิจารณาได้จากความสามารถในการคาดคะเนตามหลักเหตุผล หรือการตีความจากเรื่องี่อ่าน โดยใช้ความรู้และประสบการณ์ตรงของผู้อ่าน เพื่อทำความเข้าใจข้อมูลที่ไม่ได้ถูกระบุไว้อย่างชัดเจนซึ่งเป็นจุดมุ่งหมาย ที่ศนคติของผู้เขียน หรือข้อโต้แย้ง โดยอาศัยเบาะแส ร่องรอยในเนื้อความซึ่งเป็นคำหรือประโยคชี้แนะ หรือทำการตรวจสอบข้อเท็จจริงทั้งหมดของเรื่องี่อ่าน เพื่อพิจารณาหรือตัดสินว่าสิ่งี่อ่านทั้งหมดนั้นหมายความว่าถึงสิ่งใดและสามารถเชื่อมโยงไปถึงข้อมูลภายนอกซึ่งไม่ได้กล่าวไว้ในบทอ่าน	0.5889	2
A6	มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน (Synthesizing Information from Text) พิจารณาได้จากความสามารถในการแยกประโยคแนวคิดหลัก หรือใจความสำคัญออกจากประโยครองหรือข้อความซึ่งเป็นรายละเอียดเสริมจากบทอ่าน นำประเด็นหลักที่รวบรวมได้ มาเชื่อมโยงเชื่อมต่อกัน โดยพิจารณาความเกี่ยวข้องความเป็นเหตุเป็นผล ทำการจัดลำดับข้อความ และทำการลงข้อสรุปหรือเรียบเรียงประโยคขึ้นใหม่	0.6113	1



ภาพ 47 ผลการวินิจฉัยรายคุณลักษณะ (Attribute Prevalence)

2.2 ผลการวินิจฉัยรูปแบบความรอบรู้ของนักเรียน (Attribute Profile)

การคำนวณชั้นแฝง (Latent Class) จะทำให้ทราบจำนวนโปรไฟล์แบบแผนการตอบที่เป็นไปได้ทั้งหมด (Attribute Profile) ความสามารถการอ่านเพื่อความเข้าใจประกอบด้วย 6 คุณลักษณะ 2^k ($k=6$) ดังนั้น $2^6 = 64$ หมายความว่า ผู้สอบจะถูกจำแนกเข้าชั้นแฝงเป็น 64 ชั้นแฝง (Latent Class) เลข “1” หมายถึง มีความรอบรู้ (Mastery) และเลข “0” หมายถึง ไม่รอบรู้ (Non-mastery) ในการแสดงแบบแผนใช้ตัวเลข 6 ตำแหน่ง ตำแหน่งที่ 1 แทนคุณลักษณะ A1 มีความรู้คำศัพท์ ตำแหน่งที่ 2 แทนคุณลักษณะ A2 มีความสามารถในการระบุข้อมูลที่ต้องการ ตำแหน่งที่ 3 แทนคุณลักษณะ A3 มีความรู้ทางไวยากรณ์ ตำแหน่งที่ 4 แทนคุณลักษณะ A4 มีความสามารถในการระบุใจความสำคัญของเรื่อง ตำแหน่งที่ 5 แทนคุณลักษณะ A5 มีความสามารถในการอนุมานตำแหน่งที่ 6 แทนคุณลักษณะ A6 มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน ผลปรากฏว่า ความน่าจะเป็นที่ผู้สอบจะถูกจัดเข้าชั้นแฝงในแบบแผนการตอบ “000000” หมายความว่า ผู้สอบไม่รอบรู้ในทุกคุณลักษณะคิดเป็นร้อยละ 2.15 ความน่าจะเป็นที่ผู้สอบจะถูกจัดเข้าชั้นแฝงในแบบแผน “111111” หมายความว่า ผู้สอบรอบรู้ในทุกคุณลักษณะคิดเป็นร้อยละ 3.63 และมีความถี่ของผู้สอบที่ถูกจัดเข้าชั้นแฝงในแบบแผนเดียวกันเท่ากับ 14 ส่วนแบบแผน “110111” เป็นแบบแผนที่มีความน่าจะเป็นที่ผู้สอบจะถูกจัดเข้าชั้นแฝงสูงที่สุดร้อยละ 8.59 และมีความถี่ของผู้สอบที่ถูกจัดเข้าชั้นแฝงในแบบแผนเดียวกันสูงสุดเท่ากับ 34 ยกเว้น ชั้นแฝงที่ 2 แบบแผนการตอบ “100000” ชั้นแฝงที่ 11 แบบแผนการตอบ “010100” ชั้นแฝงที่ 15 แบบแผนการตอบ “011100” ชั้นแฝงที่ 18 แบบแผนการตอบ “100010” ชั้นแฝงที่ 31 แบบแผนการตอบ “011110” ชั้นแฝงที่ 34 แบบแผนการตอบ “100001” ชั้นแฝงที่ 47 แบบแผนการตอบ “011101” ชั้นแฝงที่ 50 แบบแผนการตอบ “100011” ซึ่งไม่มีผู้สอบถูกจัดเข้าชั้นแฝงดังกล่าว รายละเอียดในตาราง 37 ดังนี้

ตาราง 37 แสดงผลการวิเคราะห์ผลการจำแนกกลุ่มผู้สอบตามคุณลักษณะความสามารถ 64 แบบแผน

Latent Class	Attribute Profile	Class Probabilities (ความน่าจะเป็นที่ผู้สอบถูกจัดเข้าชั้นแฝง)	Class Expected Frequency (จำนวนผู้สอบที่ถูกจัดเข้าชั้นแฝงเดียวกัน)
1	000000	0.0215	8
2	100000	0.0012	0
3	010000	0.0230	9
4	110000	0.0042	1
5	001000	0.0523	20

Latent Class	Attribute Profile	Class Probabilities (ความน่าจะเป็นที่ผู้สอบถูกจัดเข้าชั้น แผง)	Class Expected Frequency (จำนวนผู้สอบที่ถูกจัดเข้าชั้นแผง เดียวกัน)
6	101000	0.0089	3
7	011000	0.0255	10
8	111000	0.0139	5
9	000100	0.0065	2
10	100100	0.0037	1
11	010100	0.0022	0
12	110100	0.0041	1
13	001100	0.0043	1
14	101100	0.0073	2
15	011100	0.0006	0
16	111100	0.0037	1
17	000010	0.0125	4
18	100010	0.0013	0
19	010010	0.0278	11
20	110010	0.0092	3
21	001010	0.0130	5
22	101010	0.0040	1
23	011010	0.0131	5
24	111010	0.0130	5
25	000110	0.0151	5
26	100110	0.0155	6
27	010110	0.0110	4
28	110110	0.0362	14
29	001110	0.0042	1
30	101110	0.0130	5
31	011110	0.0014	0
32	111110	0.0139	5
33	000001	0.0202	8
34	100001	0.0011	0

Latent Class	Attribute Profile	Class Probabilities (ความน่าจะเป็นที่ผู้สอบถูกจัดเข้าชั้น แฝง)	Class Expected Frequency (จำนวนผู้สอบที่ถูกจัดเข้าชั้นแฝง เดียวกัน)
35	010001	0.0288	11
36	110001	0.0052	2
37	001001	0.0542	21
38	101001	0.0091	3
39	011001	0.0352	13
40	111001	0.0191	7
41	000101	0.0092	3
42	100101	0.0052	2
43	010101	0.0043	1
44	110101	0.0078	3
45	001101	0.0067	2
46	101101	0.0113	4
47	011101	0.0014	0
48	111101	0.0077	3
49	000011	0.0148	5
50	100011	0.0015	0
51	010011	0.0440	17
52	110011	0.0145	5
53	001011	0.0170	6
54	101011	0.0052	2
55	011011	0.0229	9
56	111011	0.0224	8
57	000111	0.0271	10
58	100111	0.0276	10
59	010111	0.0263	10
60	110111	0.0859	34
61	001111	0.0085	3
62	101111	0.0256	10
63	011111	0.0037	1

Latent Class	Attribute Profile	Class Probabilities (ความน่าจะเป็นที่ผู้สอบถูกจัดเข้าชั้น แฝง)	Class Expected Frequency (จำนวนผู้สอบที่ถูกจัดเข้าชั้นแฝง เดียวกัน)
64	111111	0.0363	14

2.3 ตัวอย่างผลการวินิจฉัยรายบุคคล

การศึกษาครั้งนี้ ผู้วิจัยทำการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2565 ของโรงเรียนมัธยมสังกัดสำนักงานเขตพื้นที่การศึกษา มัธยมศึกษาเพชรบูรณ์ จำนวน 396 คน ด้วยแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญา ความสามารถในการอ่านเพื่อความเข้าใจที่พัฒนาขึ้น มีเนื้อหาอิงกรอบมาตรฐาน CEFR ระดับ A2 วิเคราะห์ข้อมูลด้วยโมเดลวินิจฉัยเชิงจำแนก G-DINA ในตาราง 4.14 นำเสนอตัวอย่างผลการวินิจฉัย ความสามารถในการอ่านเพื่อความเข้าใจของนักเรียนในระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 16 ตัวอย่าง (R001 ถึง R016) จากผู้สอบจำนวน 396 คน พบว่า ผู้สอบ R001 มีผลสอบ 17 คะแนน มีแบบแผนการตอบ “111111” ถูกจัดเข้าชั้นแฝงที่ 64 ผู้สอบคนนี้มีโอกาสที่จะรอบรู้ในทักษะหรือคุณลักษณะ A1, A2, A3, A4, A5, A6 ร้อยละ 99.99, 98.36, 63.29, 99.95, 99.99 และ 98.67 ตามลำดับ หมายความว่า R001 เป็นผู้รอบรู้ (Mastery) ในทุกคุณลักษณะความสามารถการอ่านเพื่อความเข้าใจ ในขณะที่ R016 มีผลสอบ 3 คะแนน มีแบบแผนการตอบ “000000” ถูกจัดเข้าชั้นแฝงที่ 1 ผู้สอบคนนี้ ไม่มีโอกาสที่จะรอบรู้ในทักษะหรือคุณลักษณะใดๆ หมายความว่า R016 เป็นผู้ไม่รอบรู้ (Non-mastery) ในทุกคุณลักษณะความสามารถการอ่านเพื่อความเข้าใจ

เมื่อพิจารณาผลการวินิจฉัยของผู้สอบ R009 และ R010 พบว่า มีผลสอบ 8 คะแนนเท่ากัน แต่มีแบบแผนการตอบแตกต่างกัน โดยที่ R009 มีแบบแผนการตอบ “010011” ถูกจัดเข้าชั้นแฝงที่ 51 ผู้สอบคนนี้มีโอกาสที่จะรอบรู้ในทักษะหรือคุณลักษณะ A2, A5, A6 ร้อยละ 80.76, 100.00 และ 60.03 ตามลำดับ หมายความว่า R009 เป็นผู้รอบรู้ในคุณลักษณะการค้นหาข้อมูลที่ปรากฏชัดเจนในบทความมีความสามารถในการอนุมาน และการสังเคราะห์ข้อมูลจากบทอ่าน ในขณะที่ R010 มีแบบแผนการตอบ “101101” ถูกจัดเข้าชั้นแฝงที่ 46 ผู้สอบคนนี้มีโอกาสที่จะรอบรู้ในทักษะหรือคุณลักษณะ A1, A3, A4, A6 ร้อยละ 97.00, 98.07, 100 และ 97.05 ตามลำดับ หมายความว่า R010 เป็นผู้รอบรู้ในคุณลักษณะ ความรู้คำศัพท์ มีความรู้ทางไวยากรณ์ มีความสามารถในการระบุใจความสำคัญของเรื่อง และมีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน เช่นเดียวกับ R011 และ R012 ซึ่งมีผลสอบ 6 คะแนนเท่ากัน แต่มีแบบแผนการตอบ และถูกจำแนกเข้าชั้นแฝงแตกต่างกัน รายละเอียดในตาราง 38 ดังนี้

ตาราง 38 แสดงตัวอย่างผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจของนักเรียนใน
ระดับชั้นมัธยมศึกษาปีที่ 3

Respondent s	score (18)	Pattern	Attribute Profile	Attributes of Reading Comprehension					
				A1	A2	A3	A4	A5	A6
R001	17	1111111111111101	111111	99.99	98.3	63.2	99.95	99.99	98.67
		1			6	9			
R002	16	1111111111110101	110111	99.89	98.6	43.0	99.82	99.97	98.29
		1			2	8			
R003	14	10111111101110011	101011	100.0	12.5	93.1	28.53	99.88	99.55
		1		0	0	0			
R004	12	11111001111101001	100111	99.99	26.4	50.1	96.79	98.55	96.75
		0			6	5			
R005	11	11111001101010101	110110	99.94	99.8	56.0	100.0	99.96	12.99
		0			4	7	0		
R006	10	10101001101111100	111010	99.95	99.2	99.9	30.57	99.78	29.04
		0			1	0			
R007	9	11001001100101101	010111	24.32	96.0	37.3	100.0	100.0	100.0
		0			9	6	0	0	0
R008	7	01110000110010000	001110	19.81	19.5	46.0	99.98	99.98	33.46
		1			2	1			
R009	8	10110010010001110	010011	22.92	80.7	19.1	18.75	100.0	60.03
		0			6	6		0	
R010	8	11111001110000000	101101	97.00	18.2	98.0	100.0	30.02	97.05
		0			5	7	0		
R011	6	00111001000010000	110000	99.98	99.9	13.1	11.66	11.66	34.22
		1			8	9			
R012	6	10001010000011010	000011	21.78	42.8	20.7	25.29	51.69	65.35
		0			7	5			
R013	5	11100000100000001	010111	33.25	65.7	26.6	100.0	99.66	100.0
		0			4	2	0		0
R014	4	00001100000001100	000001	35.20	29.4	58.2	43.02	13.88	99.00
		0			2	1			
R015	3	00010000000000011	000000	11.95	27.4	38.4	27.43	12.01	12.01

Respondent s (18)	score (18)	Pattern	Attribute Profile	Attributes of Reading Comprehension					
				A1	A2	A3	A4	A5	A6
	0			1	8				

ส่วนที่ 3 ผลการออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ

ในการศึกษาครั้งนี้ ผู้วิจัยกำหนดประเภทข้อมูลย้อนกลับตามรูปแบบ (Shute, 2000, p. 1-2) ที่เป็นการให้ข้อมูลย้อนกลับเพื่อความก้าวหน้า (Formative Feedback) โดยให้ข้อเสนอแนะเพื่อส่งเสริมสนับสนุนความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของผู้เรียนมีจุดมุ่งหมายในการส่งเสริมการเรียนรู้หรือความสามารถโดยให้สารสนเทศที่ถูกต้องตรงกับคุณลักษณะหรือทักษะที่ทำการวินิจฉัย และให้สารสนเทศย้อนกลับแบบสรุปผลรวม (Summative Feedback) โดยเป็นการให้ผลการวินิจฉัยหลังจากทำแบบทดสอบแล้ว ผู้วิจัยใช้หลักการให้ข้อมูลย้อนกลับที่มีการอธิบายประกอบเนื้อหาซึ่งมีผลต่อการเรียนรู้ของผู้เรียนมากทำให้ผู้เรียนสามารถทำความเข้าใจได้ดีกว่า และส่งผลต่อการพัฒนาการเรียนรู้ของผู้เรียนได้ดีกว่าการให้ข้อมูลย้อนกลับ ที่เพียงแต่บอกว่าถูกหรือผิด โดยไม่มีการอธิบายใด ๆ และผู้ให้ข้อมูลย้อนกลับนำเสนอแนวทางที่ผู้เรียนสามารถยอมรับและนำไปปฏิบัติได้ (Krumboltz and Bouawitz, 1962, อ้างถึงใน ทิพนิกา อรุณวิภาส, 2552, หน้า 11; De Cecco, 1968, pp. 284-294)

ผู้วิจัยทำการออกแบบ Profile ที่สามารถให้สารสนเทศผลการวินิจฉัยย้อนกลับ จากการทดสอบ (เกสรินทร์ สมราช, 2563, หน้า 120-121) ประกอบด้วย 4 ส่วน ดังนี้

1) ข้อมูลสารสนเทศที่ได้จากการวินิจฉัยความสามารถประกอบด้วย คะแนนดิบจากการทดสอบแบบข้อต่อข้อ (ตอบถูกได้ 1 ผิดได้ 0) แบบแผนของการมีคุณลักษณะใน ความสามารถ ด้านนั้น ๆ เช่น แบบแผนการตอบ “111111” หมายถึง ผู้สอบมีคุณลักษณะในความสามารถ ด้านนั้นครบทุกคุณลักษณะ โดยสามารถบอกค่าความน่าจะเป็น (Skill Probabilities) ในการรอบรู้คุณลักษณะนั้น

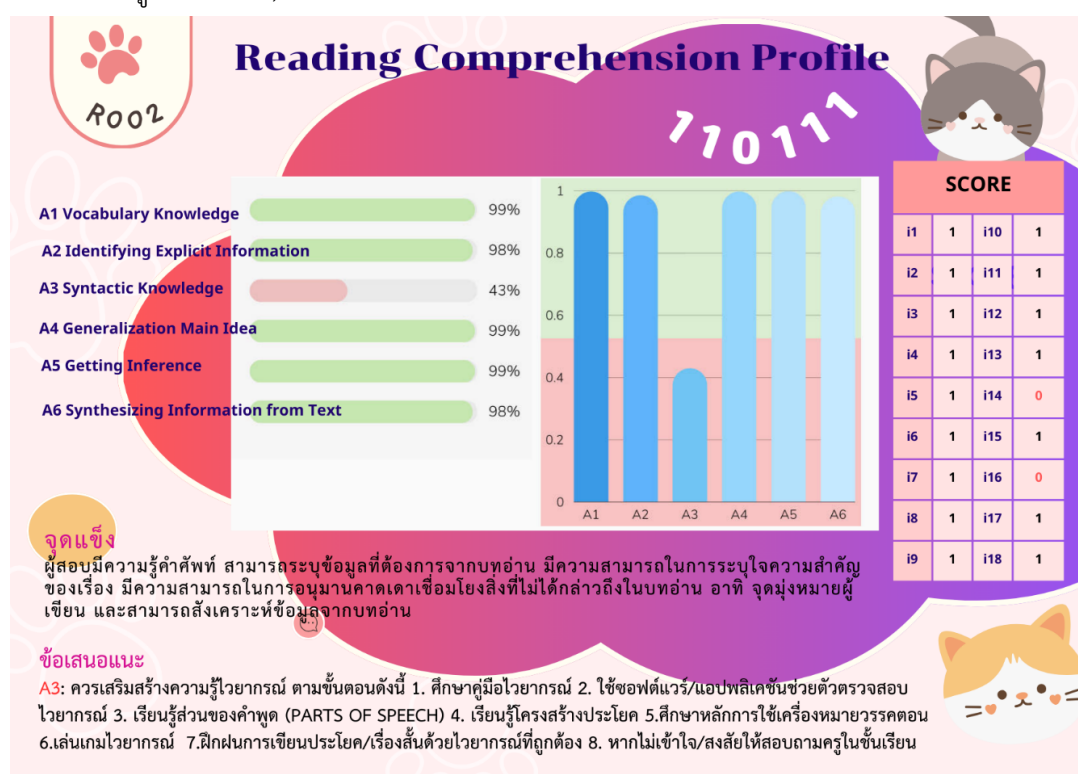
2) สารสนเทศที่เป็นความสามารถที่ผู้สอบผู้นั้นสามารถปฏิบัติได้

3) สารสนเทศที่เป็นความบกพร่องของผู้สอบผู้นั้นที่ไม่สามารถปฏิบัติได้

4) สารสนเทศที่เป็นข้อเสนอแนะในการพัฒนาทักษะ/คุณลักษณะความสามารถที่ผู้สอบไม่รอบรู้ ซึ่งเป็นข้อเสนอแนะที่ผ่านการประเมินความถูกต้องเหมาะสมจากผู้เชี่ยวชาญด้านหลักสูตรและการสอนภาษาอังกฤษระดับมัธยมศึกษา จำนวน 3 ท่าน

ผลการออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ พบว่า Profile สามารถให้ข้อมูล สารสนเทศที่ได้จากการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจ ประกอบด้วย

คะแนนดิบจากการทดสอบแบบตอบถูกได้ 1 ตอบผิดได้ 0 และแบบแผนการมีคุณลักษณะความสามารถด้านนั้น ๆ เช่น แบบแผนการตอบ 111111 หมายถึง ผู้สอบมีความรอบรู้ทุกคุณลักษณะด้านนั้นครบทุกคุณลักษณะ โดยสามารถบอกค่า ความน่าจะเป็น (Skill Probabilities) ในการรอบรู้คุณลักษณะนั้น รวมถึงสารสนเทศความสามารถที่ ผู้สอบผู้นั้นสามารถปฏิบัติได้ อาทิ ผู้สอบมีความรู้คำศัพท์ แต่ควรได้รับการพัฒนาในคุณลักษณะการอนุมาน เป็นต้น ภาพ 48, 49 และ 50 แสดง Profile ของผู้สอบ R002, R007 และ R008

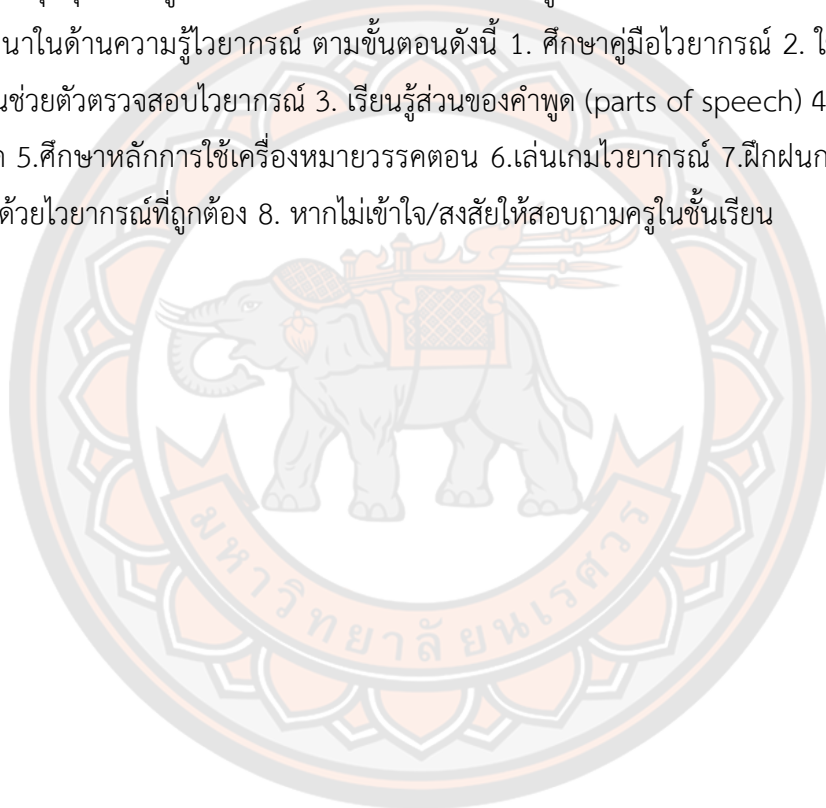


ภาพ 48 แสดง Profile ของผู้สอบ R002

จากภาพ 48 แสดงผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ มีส่วนประกอบ 6 ส่วน ได้แก่ 1) รหัสผู้สอบ 2) ผลการทดสอบด้วยแบบสอบแบบทดสอบเลขเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจที่ผู้วิจัยพัฒนาขึ้น จำนวน 18 ข้อ (ตอบถูกได้ 1 ตอบผิดได้ 0) 3) แบบแผนการตอบข้อสอบ 4) แผนภูมิแสดงระดับความรอบรู้ในแต่ละคุณลักษณะของผู้สอบ 5) สารสนเทศแสดงจุดแข็งของผู้สอบ และ 6) สารสนเทศแสดงข้อเสนอแนะเชิงบวกเพื่อการปรับปรุงจุดอ่อนเกิดประโยชน์ในการพัฒนาตนเองให้มีความรอบรู้ต่อไป

ผู้สอบ R002 มีผลการทดสอบด้วยแบบสอบแบบทดสอบเลขเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจเท่ากับ 16 คะแนน ดังนี้ 1111111111111010111 มีแบบแผนการตอบข้อสอบ “110111” จากแผนภูมิ แปลความหมายได้ว่า ผู้สอบ R002 มีความรอบรู้ใน 5 คุณลักษณะ โดยแท่งสีสูงเกิน แถบสีแดง ซึ่งเป็นสัญลักษณ์ของขอบเขตที่แสดงว่ายังไม่มีความ

รอบรู้ ดังนั้น ผู้สอบคนนี้มี ความรอบรู้ในคุณลักษณะ A1: มีความรู้คำศัพท์ เท่ากับร้อยละ 99.89 มีความรอบรู้ในคุณลักษณะ A2: สามารถระบุข้อมูลที่ต้องการจากบทอ่าน เท่ากับร้อยละ 98.62 มีความรอบรู้ในคุณลักษณะ A4: มีความสามารถในการระบุใจความสำคัญของเรื่อง ร้อยละ 99.82 มีความรอบรู้ในคุณลักษณะ A5: มีความสามารถในการอนุมาน ร้อยละ 99.97 และมีความรอบรู้ในคุณลักษณะ A6: มีความสามารถสังเคราะห์ข้อมูลจากบทอ่านเท่ากับร้อยละ 98.29 R002 มีจุดแข็ง A1: มีความรู้คำศัพท์ A2: สามารถระบุข้อมูลที่ต้องการจากบทอ่าน A4: มีความสามารถในการระบุใจความสำคัญของเรื่อง A5: มีความสามารถในการอนุมานคาดเดาเชื่อมโยงสิ่งที่ไม่ได้กล่าวถึงในบทอ่าน อาทิ จุดมุ่งหมายผู้เขียน A6: สามารถสังเคราะห์ข้อมูลจากบทอ่าน และควนปรับปรุงและได้รับการพัฒนาในด้านความรู้ไวยากรณ์ ตามขั้นตอนดังนี้ 1. ศึกษาคู่มือไวยากรณ์ 2. ใช้ซอฟต์แวร์/แอปพลิเคชันช่วยตรวจสอบไวยากรณ์ 3. เรียนรู้ส่วนของคำพูด (parts of speech) 4. เรียนรู้โครงสร้างประโยค 5. ศึกษาหลักการใช้เครื่องหมายวรรคตอน 6. เล่นเกมไวยากรณ์ 7. ฝึกฝนการเขียนประโยค/เรื่องสั้นด้วยไวยากรณ์ที่ถูกต้อง 8. หากไม่เข้าใจ/สงสัยให้สอบถามครูในชั้นเรียน



บทที่ 5

สรุปผลการวิจัย

การวิจัยเรื่องการพัฒนาแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ มีวัตถุประสงค์การวิจัย 2 ข้อ ได้แก่ 1) เพื่อสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และ 2) เพื่อวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และพัฒนา Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ เพื่อตอบวัตถุประสงค์การวิจัย จึงนำเสนอสรุปผลการวิจัยออกเป็น 2 ตอน ดังนี้

ตอนที่ 1 ผลการสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3

ตอนที่ 2 ผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 และผลการออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ

สรุปผลการวิจัย

ตอนที่ 1 ผลการสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3

1) ผลการตรวจสอบความถูกต้องเหมาะสมของการกำหนดคุณลักษณะ และความสัมพันธ์เชิงโครงสร้างของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ

สรุปผลได้ว่า คุณลักษณะและความสัมพันธ์เชิงโครงสร้างของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษมีความถูกต้องเหมาะสมเป็นไปตามกรอบมาตรฐาน CEFR ระดับ A2 Beginner ประกอบด้วย 6 คุณลักษณะ ได้แก่ A1 มีความรู้คำศัพท์ A2 มีความสามารถในการระบุข้อมูลที่ต้องการ A3 มีความรู้ทางไวยากรณ์ A4 มีความสามารถในการระบุใจความสำคัญของเรื่อง A5 มีความสามารถในการอนุมาน และ A6 มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน

2) ผลการสร้างแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ

สรุปผลได้ว่า แบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจ ที่ผู้วิจัยพัฒนาขึ้น ฉบับร่างมีจำนวน 42 ข้อ โดยมีเนื้อหาอิงกรอบมาตรฐาน CEFR ด้านการอ่านระดับ A2 คือ ผู้เรียนสามารถเข้าใจบทอ่านที่สั้นและง่าย เป็นเรื่องใกล้ตัวและใช้

ภาษาที่พบบ่อยในชีวิตประจำวัน ครอบคลุมความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ จำแนกตามคุณลักษณะ ประกอบด้วย 5 หัวข้อ (สถาบันภาษาอังกฤษ สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2560) ดังนี้ 1) การอ่านคำสั่ง/วิธีทำหรือวิธีใช้บางสิ่งบางอย่าง (Reading Instruction) คำแนะนำที่ใช้ ภาษาง่าย ๆ เช่น โพรศัพทสารณะ เครื่องจำหน่ายตัวสารณะ ข้อมูลเกี่ยวกับความปลอดภัย การบอกทิศทาง สามารถปฏิบัติตามคำสั่ง คำชี้แจง และขั้นตอนง่าย ๆ เช่น ขั้นตอนในการเล่นเกมส์ การใช้อุปกรณ์ชนิดต่าง ๆ ที่คุ้นเคย หรือการทำอาหาร เป็นต้น 2) การอ่านเพื่อหาข้อมูลเบื้องต้น (Reading for Orientation) สามารถหาข้อมูลที่สำคัญในบทโฆษณาแผ่นพับ ใบปลิว หนาเว็บ (web page) แคตตาล็อก ตารางเวลา ฯลฯ 3) การอ่านข้อมูลและข้อโต้แย้ง (Reading for Information & Argument) คำบรรยาย ข่าวสั้น ๆ บทความสั้น ๆ จากหนังสือพิมพ์ และนิตยสารเกี่ยวกับเหตุการณ์ ปัจจุบัน และหัวข้อที่คุณเคย ถ้าผู้อ่านมีความรู้เรื่องเกี่ยวกับเรื่องนั้น ๆ บางอยู่แล้ว เช่น ข่าวเกี่ยวกับกีฬา และบุคคลที่มีชื่อเสียง 4) การอ่านวรรณกรรม (Read Literature) เรื่องในชีวิตประจำวันที่ง่าย ๆ และสั้น ๆ โดยเฉพาะอย่างยิ่งเมื่อมีสื่อประกอบและ 5) การอ่านจดหมายอิเล็กทรอนิกส์ (Reading Correspondence) เกี่ยวกับเหตุการณ์ต่าง ๆ ทำการวินิจฉัยอย่างเป็นทางการ (formal diagnosis) เพื่อค้นหาข้อบกพร่องของนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยมีวิธีการสร้างแบบสอบวินิจฉัยอย่างเป็นทางการเป็นระบบอิงแนวคิดกระบวนการ CDS ของเอ็มเบรทสัน (Embretson, 1994)

3) ผลการตรวจสอบความสอดคล้องระหว่างข้อคำถามและคุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ

สรุปผลได้ว่า ผลการตรวจสอบความตรงเชิงเนื้อหาด้วยดัชนีความสอดคล้องระหว่างข้อคำถามและคุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษมีบทอ่าน (Passages) จำนวน 4 เรื่อง ประกอบด้วยข้อสอบแบบทดสอบ (Testlets) จำนวนไม่เท่ากัน โดยเป็นข้อสอบปรนัยแบบเลือกตอบ 4 ตัวเลือก กำหนดรูปแบบการให้คะแนนแบบ 2 ค่า คือ ให้ 1 คะแนน สำหรับข้อที่ผู้เรียนตอบถูก และให้ 0 คะแนน สำหรับข้อที่ผู้เรียนตอบผิด ประกอบด้วยข้อคำถามจำนวน 42 ข้อ พิจารณาความตรงเชิงเนื้อหาด้วยดัชนีวัดความสอดคล้องระหว่างข้อคำถามและคุณลักษณะของความสามารถที่ใช้ในการตอบ (Index of Item-Objective Congruence: IOC) โดยผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษ การวิจัยและประเมินผลทางการศึกษาจำนวน 5 ท่าน เพื่อตรวจสอบความสอดคล้องของข้อคำถามและคุณลักษณะที่มุ่งวัด (Q-matrix) ผลปรากฏว่า ข้อคำถาม 42 ข้อ มีค่าดัชนี IOC ตั้งแต่ 0.6 ถึง 1.0 ค่าดัชนี IOC ควรมีค่า ≥ 0.5 ดังนั้นข้อคำถามทุกข้อมีความตรงเชิงเนื้อหาและมีความสอดคล้องระหว่างข้อคำถามและคุณลักษณะของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ

4) ผลการตรวจสอบคุณภาพของแบบสอบ

4.1 ผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยทฤษฎีการทดสอบดั้งเดิม (Classical Test Theory: CTT) ผลปรากฏว่า ข้อคำถามที่ผ่านเกณฑ์คุณภาพด้านความยากและอำนาจจำแนกตามทฤษฎีการทดสอบแบบดั้งเดิมมีจำนวน 33 ข้อ มีค่าความยากในภาพรวมอยู่ระหว่าง 0.25 ถึง 0.70 และค่าอำนาจจำแนกอยู่ระหว่าง 0.26 ถึง 0.87 แบ่งออกเป็นแต่ละเทสต์เลท ดังนี้ เทสต์เลทที่ 1 ข้อคำถามที่ 1 ถึง 10 มีค่าความยากอยู่ระหว่าง 0.51 ถึง 0.66 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.57 ถึง 0.87 ข้อคำถามที่ผ่านเกณฑ์คุณภาพด้านความยากและอำนาจจำแนกมีจำนวน 8 ข้อ ข้อที่ถูกคัดออกได้แก่ ข้อ 2 และข้อ 9 เทสต์เลทที่ 2 ข้อคำถามที่ 11 ถึง 21 มีค่าความยากอยู่ระหว่าง 0.25 ถึง 0.67 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.26 ถึง 0.70 ข้อคำถามที่ผ่านเกณฑ์คุณภาพด้านความยากและอำนาจจำแนกมีจำนวน 8 ข้อ ข้อที่ถูกคัดออกได้แก่ ข้อ 13 และข้อ 16 เทสต์เลทที่ 3 ข้อคำถามที่ 21 ถึง 33 มีค่าความยากอยู่ระหว่าง 0.39 ถึง 0.79 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.22 ถึง 0.81 ข้อคำถามที่ผ่านเกณฑ์คุณภาพด้านความยากและอำนาจจำแนกมีจำนวน 9 ข้อ ข้อที่ถูกคัดออกได้แก่ ข้อ 26 29 30 และข้อ 33 เทสต์เลทที่ 4 ข้อคำถามที่ 34 ถึง 42 มีค่าความยากอยู่ระหว่าง 0.34 ถึง 0.63 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.24 ถึง 0.76 ข้อคำถามที่ผ่านเกณฑ์คุณภาพด้านความยากและอำนาจจำแนกมีจำนวน 8 ข้อ ข้อที่ถูกคัดออกได้แก่ ข้อ 34

4.2 ผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT 3PL) ผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT 3PL) ค่าพารามิเตอร์ความยาก $/b/$ ค่าพารามิเตอร์อำนาจจำแนก $/a/$ และสัมประสิทธิ์การเดา $/g/$ ผลปรากฏว่า ข้อคำถามที่มีค่าพารามิเตอร์ความยาก ค่าพารามิเตอร์อำนาจจำแนก และสัมประสิทธิ์การเดาผ่านเกณฑ์มีจำนวน 18 ข้อ ในภาพรวม ข้อคำถามมีค่าพารามิเตอร์ความยากอยู่ระหว่าง -0.29 ถึง 1.19 ค่าพารามิเตอร์อำนาจจำแนกอยู่ระหว่าง 0.82 ถึง 2.50 และสัมประสิทธิ์การเดา อยู่ระหว่าง 0.12 ถึง 0.28

ผลการตรวจสอบรายเทสต์เลท จำนวน 4 เทสต์เลท ดังนี้ Testlet 1: Reading Instruction ข้อที่ผ่านเกณฑ์มีจำนวน 5 ข้อ ได้แก่ 1, 4, 5, 6, 7 มีค่าพารามิเตอร์ความยาก $/b/$ อยู่ระหว่าง -0.24 ถึง 0.81 ค่าพารามิเตอร์อำนาจจำแนก $/a/$ อยู่ระหว่าง 1.83 ถึง 2.46 และสัมประสิทธิ์การเดา $/g/$ อยู่ระหว่าง 0.18 ถึง 0.22 Testlet 2: Reading for Orientation มีข้อที่ผ่านเกณฑ์มีจำนวน 4 ข้อ ได้แก่ 14, 18, 19, 20 มีค่าพารามิเตอร์ความยาก $/b/$ อยู่ระหว่าง -0.05 ถึง 1.13 ค่าพารามิเตอร์อำนาจจำแนก $/a/$ อยู่ระหว่าง 0.82 ถึง 2.18 และสัมประสิทธิ์การเดา $/g/$ อยู่ระหว่าง 0.17 ถึง 0.28 Testlet 3: Reading for Info & Argument ข้อที่ผ่านเกณฑ์มีจำนวน 4 ข้อ ได้แก่ 21, 23, 24, 25 มีค่าพารามิเตอร์ความยาก $/b/$ อยู่ระหว่าง -0.29 ถึง 1.02 ค่าพารามิเตอร์อำนาจจำแนก $/a/$ อยู่ระหว่าง 1.66 ถึง 2.48 และสัมประสิทธิ์การเดา $/g/$ อยู่ระหว่าง 0.15 ถึง 0.24 Testlet 4: Reading Literature ข้อที่ผ่านตรวจสอบคุณภาพ จำนวน 5 ข้อ ได้แก่ 35, 36, 38, 39,

40 มีค่าพารามิเตอร์ความยาก /b/ อยู่ระหว่าง -0.11 ถึง 1.19 ค่าพารามิเตอร์อำนาจจำแนก /a/ อยู่ระหว่าง 0.87 ถึง 2.50 และสัมประสิทธิ์การเดา /g/ อยู่ระหว่าง 0.12 ถึง 0.23

5) ผลการคัดเลือกข้อสอบและการสร้าง Q-matrix ของแบบสอบวินิจฉัย

สรุปได้ว่า ผลการคัดเลือกข้อสอบมาเรียงเรียงเป็นแบบสอบแบบทดสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ที่วัดความสามารถการอ่านเพื่อความเข้าใจตามเนื้อหาอิงกรอบมาตรฐาน CEFR ระดับ A2 ประกอบด้วย 4 ทดสอบ มีข้อคำถามจำนวนทั้งหมด 18 ข้อ แต่ละข้อคำถามมุ่งวัด 2 ถึง 4 คุณลักษณะ โดยแต่ละคุณลักษณะจะมีจำนวนข้อคำถามไม่เท่ากัน แต่ครอบคลุมทุกคุณลักษณะ ได้ Q-matrix ของแบบสอบวินิจฉัยมีจำนวน 1 เมตริกซ์โดยแนวนอนมี 19 คอลัมน์ และแนวตั้งมี 7 คอลัมน์

ตอนที่ 2 ผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้น

มัธยมศึกษาปีที่ 3 และผลการออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ

1. สรุปผลการทดลองใช้และตรวจสอบคุณภาพของแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ ด้วยโมเดลการวินิจฉัย Generalization DINA

1.1 สรุปผลการวิเคราะห์ข้อมูลเชิงบรรยายของการตรวจสอบคุณภาพของแบบสอบแบบทดสอบเพื่อวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 18 ข้อ คะแนนเต็ม 18 คะแนน ผลการตอบแบบสอบวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจของนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 396 คน พบว่า นักเรียนตอบแบบสอบวินิจฉัยได้คะแนนสูงสุดเท่ากับ 17 คะแนน จำนวน 6 คน คิดเป็นร้อยละ 1.52 นักเรียนตอบแบบสอบวินิจฉัยได้คะแนนต่ำสุด 3 คะแนน จำนวน 10 คน คิดเป็นร้อยละ 2.53 นักเรียนส่วนใหญ่ตอบแบบสอบวินิจฉัยได้ 5 คะแนน คิดเป็นร้อยละ 14.39

1.2 สรุปผลการตรวจสอบความตรงของเมตริกซ์คิว (Q-matrix validation) ด้วยหลักฐานเชิงประจักษ์ โดยวิธี Stepwise Wald test (Ma & de la Torre, 2019) ด้วยโมเดลวินิจฉัย G-DINA Model ได้เมตริกซ์คิวแนะนำ (Suggested Q-matrix) ผลปรากฏว่า จากการตรวจสอบความตรงของเมตริกซ์คิว (Q-matrix validation) ด้วยหลักฐานเชิงประจักษ์ โดยวิธี Stepwise Wald test (Ma & de la Torre, 2019) ได้เมตริกซ์คิวแนะนำ (Suggested Q-matrix) ให้ทำการปรับจำนวน 10 ข้อ ได้แก่ ข้อ 1, 2, 7, 8, 10, 13, 14, 15, 16 และ 17 โดยข้อคำถามที่ถูกแนะนำให้ปรับมีจำนวน 1-3 คุณลักษณะ

1.3 สรุปผลการวิเคราะห์ประสิทธิภาพความสอดคล้องเชิงสัมพัทธ์ของโมเดลการวินิจฉัย Generalization DINA (G-DINA MODEL) แสดงด้วยค่าดัชนีความสอดคล้องเชิงสัมพัทธ์ของโมเดล (Test Level Fit Statistics) ค่าสถิติทดสอบความสอดคล้องของข้อสอบ ด้วยเกณฑ์การเปรียบเทียบ ระหว่าง 2 โมเดล คือ CDM package กับ G-DINA package ได้แก่ $-2\text{Loglikelihood} (-2LL)$ ของโมเดล CDM มีค่าเท่ากับ -4782.67 และโมเดล G-DINA มีค่าเท่ากับ -4754.90 โดยค่าที่ต่ำกว่า แสดงให้เห็นถึงประสิทธิภาพในการประมาณค่าของโมเดลที่ดีกว่า เมื่อพิจารณาค่าสถิติทดสอบความสอดคล้องอีก 2 ตัว ของ G-DINA MODEL ได้แก่ Akaike's Information Criterion (AIC) มีค่าเท่ากับ 9923.80 และค่า Bayesian Information Criterion (BIC) มีค่าเท่ากับ 10792.04 จึงสรุปได้ว่า G-DINA MODEL เป็นโมเดลที่มีประสิทธิภาพในการประมาณค่าของโมเดลที่ดีกว่า มีประสิทธิภาพความสอดคล้องเชิงสัมพัทธ์กับข้อมูลเชิงประจักษ์ สามารถวินิจฉัยความสามารถของผู้สอบได้

1.4 สรุปผลการวิเคราะห์ค่าระดับความแม่นยำในการจำแนกของแบบสอบรายฉบับ (Test Level Accuracy) ด้วยโมเดลวินิจฉัย Generalization DINA (G-DINA MODEL) พบว่า แบบสอบแบบทดสอบเพื่อวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจที่ผู้วิจัยพัฒนาขึ้น สามารถจำแนกผู้สอบเข้าชั้นแ่งตามคุณลักษณะความสามารถรายบุคคลได้อย่างถูกต้อง เท่ากับ 0.82 โดยมีค่ามากกว่า 0.5 และมีค่าเข้าใกล้ 1.0 สรุปผลการวิเคราะห์ค่าระดับความแม่นยำในการจำแนกรายคุณลักษณะ (Attribute Level Accuracy) ด้วยโมเดลวินิจฉัย Generalization DINA (G-DINA MODEL) จำนวน 6 คุณลักษณะ พบว่า แบบสอบมีค่าระดับความแม่นยำในการจำแนกรายคุณลักษณะอยู่ระหว่าง $0.90 - 0.99$ โดยที่คุณลักษณะที่ 1 (A1 มีความรู้คำศัพท์) และคุณลักษณะที่ 3 มีความรู้ทางไวยากรณ์ มีค่าความถูกต้องแม่นยำสูงที่สุด เท่ากับ 0.99 รองลงไปคือคุณลักษณะที่ 6 มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน มีค่าเท่ากับ 0.95 คุณลักษณะที่ 4 มีความสามารถในการระบุใจความสำคัญของเรื่อง และคุณลักษณะที่ 5 มีความสามารถในการอนุมานมีค่าเท่ากับ 0.94 คุณลักษณะที่ 2 มีความสามารถในการระบุข้อมูลที่ต้องการ มีค่าเท่ากับ 0.90 ตามลำดับ ซึ่งมีค่าใกล้ 1.0 ทุกคุณลักษณะ

สรุปผลการตรวจสอบคุณภาพด้านความตรงในการจำแนกและความเที่ยงเชิงจำแนกของแบบสอบด้วยโมเดลวินิจฉัย G-DINA (Class Accuracy) ด้วยวิธีการประมาณค่าความน่าจะเป็นสูงสุด (Maximum Likelihood Estimation: MLE) และ ค่าความน่าจะเป็น MAP (Maximum a posterior) ปรากฏว่า แบบสอบมีความตรงในการจำแนกรายคุณลักษณะอยู่ระหว่าง 0.882 ถึง 0.982 ซึ่งมีเข้าใกล้ 1.0 หมายความว่า แบบสอบมีคุณภาพด้านความตรงเชิงจำแนก และมีค่าความเที่ยงเชิงจำแนกรายคุณลักษณะ อยู่ระหว่าง 0.809 ถึง 0.965 ซึ่งมีเข้าใกล้ 1.0 หมายความว่า แบบสอบมีคุณภาพด้านความเที่ยงเชิงจำแนกรายคุณลักษณะ

1.5 สรุปผลการตรวจสอบคุณภาพของข้อคำถามรายข้อด้วยโมเดล G-DINA

1.5.1 ค่าพารามิเตอร์การเดา (Guessing Parameter) อยู่ระหว่าง 0.001 – 0.4338 ข้อสอบทุกข้อมีค่าพารามิเตอร์การเดาตามเกณฑ์ คือไม่มีข้อใดที่มีค่าพารามิเตอร์การเดาเกินกว่า 0.5 แสดงว่าข้อสอบทั้ง 18 ข้อ เป็นข้อสอบที่นักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีความสามารถการอ่านเพื่อความเข้าใจในระดับต่ำ จะไม่สามารถเดาคำตอบได้ถูกต้อง ค่าพารามิเตอร์ความสะเพร่า (Slipping parameter) อยู่ระหว่าง 0.001 - 0.8893 ข้อสอบจำนวน 17 ข้อ มีค่าพารามิเตอร์ความสะเพร่าตามเกณฑ์ คือไม่มีข้อที่มีค่าพารามิเตอร์ความสะเพร่าเกินกว่า 0.5 ยกเว้นข้อ 7 มีค่าพารามิเตอร์ความสะเพร่าเท่ากับ 0.8893 หมายความว่า ผู้สอบมักจะตอบข้อคำถามที่ 7 ผิด เพราะความสะเพร่า

1.5.2 ค่าดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ของข้อคำถามรายข้อ (RMSEA) เมื่อตรวจสอบความเหมาะสมของข้อสอบเป็นรายข้อด้วยดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ (RMSEA) มีค่าระหว่าง 0.032 ถึง 0.169 และเมื่อตรวจสอบโดยใช้เกณฑ์ของ Maydeu – Olivares & Joe (2014, อ้างถึงใน เกศรินทร์ สมราช, 2563, น.188) $RMSEA < 0.089$ หมายถึง ข้อสอบข้อนั้นมีความเหมาะสมสามารถนำไปใช้ได้ และ $RMSEA < 0.05$ หมายถึง ข้อสอบข้อนั้นมีความคลาดเคลื่อนต่ำมากและมีความเหมาะสมเป็นอย่างมากที่จะนำไปใช้วินิจฉัยผู้เรียน ผลปรากฏว่า ค่า RMSEA ของข้อสอบจำนวน 11 ข้อ มีค่าอยู่ระหว่าง 0.032 ถึง 0.081 ซึ่งเป็นไปตามเกณฑ์ ยกเว้น ข้อคำถามจำนวน 7 ข้อ ได้แก่ ข้อ 2, 6, 7, 10, 11, 13 และ 16 โดยมีค่า 0.119, 0.148, 0.109, 0.169, 0.112, 0.114 และ 0.139 ตามลำดับ อย่างไรก็ตาม เมื่อพิจารณาค่าเฉลี่ยของ RMSEA ของข้อสอบ 18 ข้อ มีค่าเท่ากับ 0.086 เป็นไปตามเกณฑ์ สามารถนำไปใช้วินิจฉัยความสามารถผู้เรียนได้

1.5.3 ค่าดัชนีอำนาจจำแนกของข้อสอบรายข้อ (Item Discrimination Index: IDI) เป็นดัชนีที่บ่งชี้ว่าข้อสอบข้อนั้นสามารถจำแนกผู้สอบที่มีความสามารถกับผู้ที่ไม่มีความสามารถได้ดีมากน้อยเพียงใดโดยมีเกณฑ์การผ่าน คือ มีค่า $IDI \geq 0.20$ (de la Torre, 2008, pp. 343 – 362 and de la Torre, Rossi and van der Ark, 2018, pp. 281 – 296, อ้างถึงใน วรางคนา ก้อมน้อย, 2563, น.134) ผลปรากฏว่า ข้อคำถามทั้ง 18 ข้อมีค่าดัชนีอำนาจจำแนกผ่านเกณฑ์ โดยมีค่าดัชนี IDI อยู่ระหว่าง 0.2105 ถึง 0.9238

2. ผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3

2.1 สรุปผลการวินิจฉัยรายคุณลักษณะ พบว่า ผลการวินิจฉัยรายคุณลักษณะในภาพรวมของความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ดังนี้ ความรอบรู้อันดับที่ 1 นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 มีความน่าจะเป็นที่จะรอบรู้ (Mastery) ในคุณลักษณะ A6 มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน ร้อยละ 61.13 อันดับที่ 2 นักเรียนมีความน่าจะเป็นที่จะรอบรู้ (Mastery) ในคุณลักษณะ A5 มีความสามารถในการ

อนุমানร้อยละ 58.89 อันดับที่ 3 นักเรียนมีความน่าจะเป็นที่จะรอบรู้ (Mastery) ในคุณลักษณะ A2 มีความสามารถในการระบุข้อมูลที่ต้องการ ร้อยละ 56.99 อันดับที่ 4 นักเรียนมีความน่าจะเป็นที่จะรอบรู้ (Mastery) ในคุณลักษณะ A3 มีความรู้ทางไวยากรณ์ ร้อยละ 48.01 รองลงไปนักเรียนมีความน่าจะเป็นที่จะรอบรู้ในคุณลักษณะ A1 มีความรู้คำศัพท์ ร้อยละ 44.00 และอันดับสุดท้าย นักเรียนมีความน่าจะเป็นที่จะรอบรู้ในคุณลักษณะ A4 มีความสามารถในการระบุใจความสำคัญของเรื่อง ร้อยละ 43.89 ตามลำดับ

2.2 สรุปผลการวินิจฉัยรูปแบบความรู้ของนักเรียน (Attribute Profile) พบว่า ผู้สอบจะถูกจำแนกเข้าชั้นแฝงเป็น 64 ชั้นแฝง (Latent Class) เลข “1” หมายถึง มีความรอบรู้ (Mastery) และเลข “0” หมายถึง ไม่รอบรู้ (Non-mastery) ในการแสดงแบบแผนใช้ตัวเลข 6 ตำแหน่ง ตำแหน่งที่ 1 แทนคุณลักษณะ A1 มีความรู้คำศัพท์ ตำแหน่งที่ 2 แทนคุณลักษณะ A2 มีความสามารถในการระบุข้อมูลที่ต้องการ ตำแหน่งที่ 3 แทนคุณลักษณะ A3 มีความรู้ทางไวยากรณ์ ตำแหน่งที่ 4 แทนคุณลักษณะ A4 มีความสามารถในการระบุใจความสำคัญของเรื่อง ตำแหน่งที่ 5 แทนคุณลักษณะ A5 มีความสามารถในการอนุมานตำแหน่งที่ 6 แทนคุณลักษณะ A6 มีความสามารถในการสังเคราะห์ข้อมูลจากบทอ่าน ผลปรากฏว่า ความน่าจะเป็นที่ผู้สอบจะถูกจัดเข้าชั้นแฝงในแบบแผนการตอบ “000000” หมายความว่า ผู้สอบไม่รอบรู้ในทุกคุณลักษณะคิดเป็นร้อยละ 2.15 ความน่าจะเป็นที่ผู้สอบจะถูกจัดเข้าชั้นแฝงในแบบแผน “111111” หมายความว่า ผู้สอบรอบรู้ในทุกคุณลักษณะคิดเป็นร้อยละ 3.63 และมีความถี่ของผู้สอบที่ถูกจัดเข้าชั้นแฝงในแบบแผนเดียวกันเท่ากับ 14 ส่วนแบบแผน “110111” เป็นแบบแผนที่มีความน่าจะเป็นที่ผู้สอบจะถูกจัดเข้าชั้นแฝงสูงที่สุดร้อยละ 8.59 และมีความถี่ของผู้สอบที่ถูกจัดเข้าชั้นแฝงในแบบแผนเดียวกันสูงสุดเท่ากับ 34 ยกเว้น ชั้นแฝงที่ 2 แบบแผนการตอบ “100000” ชั้นแฝงที่ 11 แบบแผนการตอบ “010100” ชั้นแฝงที่ 15 แบบแผนการตอบ “011100” ชั้นแฝงที่ 18 แบบแผนการตอบ “100010” ชั้นแฝงที่ 31 แบบแผนการตอบ “011110” ชั้นแฝงที่ 34 แบบแผนการตอบ “100001” ชั้นแฝงที่ 47 แบบแผนการตอบ “011101” ชั้นแฝงที่ 50 แบบแผนการตอบ “100011” ซึ่งไม่มีผู้สอบถูกจัดเข้าชั้นแฝงดังกล่าว

3. ผลการออกแบบ Profile สารสนเทศจากการวินิจฉัยแบบย้อนกลับ สรุปได้ว่า Profile ของผู้สอบรายบุคคล ที่ผู้วิจัยพัฒนาขึ้นเป็นประเภทข้อมูลย้อนกลับตามรูปแบบ (Shute, 2008: 1-2) ที่เป็นการให้ข้อมูลย้อนกลับเพื่อความก้าวหน้า (Formative Feedback) โดยให้ข้อเสนอแนะเพื่อส่งเสริมสนับสนุนความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของผู้เรียนมีจุดมุ่งหมายในการส่งเสริมการเรียนรู้หรือความสามารถโดยให้สารสนเทศที่ถูกต้องตรงกับคุณลักษณะหรือทักษะที่ทำการวินิจฉัย และให้สารสนเทศย้อนกลับแบบสรุปผลรวม (Summative Feedback) โดยเป็นการให้ผลการวินิจฉัยหลังจากทำแบบทดสอบแล้ว (ทิพนิกา อรุณวิภาส, 2552, หน้า 11) มี

ส่วนประกอบ 6 ส่วน ได้แก่ 1) รหัสผู้สอบ 2) ผลการทดสอบด้วยแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจที่ผู้วิจัยพัฒนาขึ้น จำนวน 18 ข้อ (ตอบถูกได้ 1 ตอบผิดได้ 0) 3) แบบแผนการตอบข้อสอบ 4) แผนภูมิแสดงระดับความรอบรู้ในแต่ละคุณลักษณะของผู้สอบ 5) สารสนเทศแสดงจุดแข็งของผู้สอบ และ 6) สารสนเทศแสดงข้อเสนอแนะเชิงบวกเพื่อการปรับปรุงจุดอ่อนเกิดประโยชน์ในการพัฒนาตนเองให้มีความรอบรู้ต่อไป

อภิปรายผลการวิจัย

ผลการวิจัยในครั้งนี้พบประเด็นและข้อสังเกตที่น่าสนใจสามารถนำมาสรุปและอภิปรายผลการวิจัย โดยมีรายละเอียดดังนี้

1. การสร้างและตรวจสอบคุณภาพแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญา ด้านความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากการพัฒนาแบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาด้านความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ผู้วิจัย สร้างขึ้นครอบคลุม 6 คุณลักษณะ มีความถูกต้องเหมาะสม เนื่องจากจากผู้วิจัยทำการสังเคราะห์ทักษะ/คุณลักษณะการอ่านเพื่อความเข้าใจภาษาอังกฤษ (Reading Comprehension) ของนักวิชาการทางภาษาศาสตร์ (Munby's taxonomy, 1978; Heaton's taxonomy, 1991; Gao, 2006; Jang, 2009; Li et. al., 2015; Thip, 2017; Javidanmehr, 2019; Tonekaboni, 2021)

และผ่านการประเมินความถูกต้องจากผู้เชี่ยวชาญทางด้านหลักสูตรและการสอนภาษาอังกฤษ จำนวน 18 ข้อ ซึ่งเป็นข้อสอบที่ผ่านการตรวจสอบคุณภาพ ด้านความตรงเชิงเนื้อหา (IOC = 0.6 ถึง 1.0) ทุกข้อคำถาม (จากข้อสอบทั้งหมด 42 ข้อ) และผ่านการตรวจสอบคุณภาพรายข้อ ค่าความยาก ($P = 0.25$ ถึง 0.70) ค่าอำนาจจำแนก ($r = 0.26$ ถึง 0.87) ผ่านเกณฑ์จำนวน 33 ข้อ จากข้อสอบทั้งหมด 42 ข้อ การตรวจสอบคุณภาพรายข้อ ด้วยทฤษฎีการตอบสนองข้อสอบ (IRT 3 PL) ข้อคำถามมีค่าพารามิเตอร์ความยากอยู่ระหว่าง -0.29 ถึง 1.19 ค่าพารามิเตอร์อำนาจจำแนกอยู่ระหว่าง 0.82 ถึง 2.50 และสัมประสิทธิ์การเดา อยู่ระหว่าง 0.12 ถึง 0.28 มีข้อคำถามผ่านเกณฑ์ในขั้นตอนนี้จำนวน 18 ข้อ ซึ่งข้อคำถามที่ผ่านเกณฑ์ทั้ง 18 ข้อ ยังคงวัดครอบคลุมคุณลักษณะ 6 คุณลักษณะ และมีจำนวนเทสต์เลทที่วัดเนื้อหา 5 เรื่องตามกรอบมาตรฐาน CEFR A2 Beginner อย่างครอบคลุมครบถ้วน เนื่องจากจากผู้วิจัยได้สร้างข้อคำถามที่มีจำนวนมากกว่าที่จำเป็นต้องใช้ และทำการตรวจสอบคุณภาพค่าความยาก ค่าอำนาจจำแนก โดยนำข้อมูลที่ได้มาวิเคราะห์เป็นรายเทสต์เลท สอดคล้องกับ DeMars, C. (2006). กล่าวว่ ข้อคำถามแบบมัตรวมเป็นชุดที่ใช้วัดสถานการณ์เดียวกันถูกนำมาวิเคราะห์คุณภาพพร้อมกันเป็นชุด และยังสอดคล้องกับ ญัฐกานต์ ประจันบาน (2563) กล่าวว่ ควรสร้างข้อคำถามเพิ่มเติมไว้ล่วงหน้าสอดคล้องกับผังการสร้างข้อสอบ (Test Blueprint) และยัง

สอดคล้องกับ พิสิฐ พิณิจสกุล (2563, หน้า 77; อริสสา เตหลิ้ม, 2559, หน้า 2; จงกล บัวแก้ว, 2560, หน้า 4) กล่าวว่า แต่ละบทอ่านมีข้อความที่เกี่ยวกับบทอ่านเป็นเซตเดียวกัน อาจมีจำนวนข้อความ 3-5 ข้อ หรือบางครั้ง 6-12 ข้อ ซึ่งบทอ่านควรมีความยาวประมาณ 300 คำก็เพียงพอ ทั้งนี้แบบสอบวินิจฉัยฉบับนี้ ยังสอดคล้องกับมาตรฐานและตัวชี้วัดรายวิชาภาษาอังกฤษตามหลักสูตรแกนกลางกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ วิชาภาษาอังกฤษ พุทธศักราช 2551

การตรวจสอบขนาดอิทธิพลของเทสต์เลท (Testlet effect) ในประเด็นนี้ ผู้วิจัยได้ทำการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง และผลการตรวจสอบคุณภาพของแบบสอบวินิจฉัยฉบับร่างตามทฤษฎีทดสอบแบบดั้งเดิม (CTT) พิจารณาค่าความยาก ค่าอำนาจจำแนก ของข้อสอบเทสต์เลทรายข้อ และความเที่ยงของแบบสอบทั้งฉบับ พบว่ามีค่าเป็นไปตามเกณฑ์ จำนวน 33 ข้อ ต่อมาเมื่อตรวจสอบคุณภาพของแบบสอบด้วยทฤษฎีการตอบสนองข้อสอบ (IRT) พบว่า มีค่าเป็นไปตามเกณฑ์ผ่านการคัดเลือกจำนวน 18 ข้อ โดยขั้นตอนทั้งหมดไม่กระทบต่อจำนวนข้อสอบในชุดเทสต์เลทเดียวกัน ยังคงมีข้อสอบจำนวน 2-10 ข้อ ตามหลักการของ Testlet Theory จากนั้นผู้วิจัยทำการวิเคราะห์ด้วยโปรแกรม R พบว่า อิทธิพลมีขนาดเล็ก และไม่มีนัยสำคัญทางสถิติหมายความว่าไม่มีผลกระทบต่อการวินิจฉัยด้วยแบบสอบวินิจฉัยทางปัญญาถึงแม้ว่าจะใช้รูปแบบของเทสต์เลทก็ตามสามารถใช้แบบสอบในการวินิจฉัยความสามารถด้านการอ่านเพื่อความเข้าใจได้

การตรวจสอบความตรงของ Q-matrix ในการศึกษาครั้งนี้ดำเนินการใน 2 ขั้นตอน กล่าวคือ ตรวจสอบความตรงโดยผู้เชี่ยวชาญด้านหลักสูตรและการสอนวิชาภาษาอังกฤษ จำนวน 5 ท่าน ผลการตรวจสอบได้ Q-matrix ในขั้นต้นที่เรียกว่า Initial Q-matrix จากนั้นใช้โมเดล G-DINA ในโปรแกรม R โดยเรียก package G-DINA เลือกใช้วิธี Stepwise Wald Test ทำการตรวจสอบความตรงด้วยโมเดลทางสถิติ ผลการตรวจสอบ พบว่า โปรแกรมแนะนำให้ปรับจำนวน 9 ข้อ โดยมีจำนวน 1 ถึง 4 คุณลักษณะ อย่างไรก็ตามการตรวจสอบความตรงโดยผู้เชี่ยวชาญ 5 ท่านมีพื้นฐานการพิจารณาจากองค์ความรู้และหลักวิชาการ และมีความคิดเห็นสอดคล้องกับอาจารย์ที่ปรึกษาวิทยานิพนธ์ ดังนั้น ผู้วิจัยจึงปรับ Q-matrix โดยยึดผลการตรวจสอบของผู้เชี่ยวชาญเป็นหลัก และพิจารณาปรับจาก 0 เป็น 1 ตามคำแนะนำของโปรแกรมเพื่อความถูกต้องเหมาะสมของ Q-matrix ที่นำไปใช้ในขั้นตอนของการวินิจฉัยต่อไป

การพิจารณาเลือกใช้โปรแกรมในการวิเคราะห์ที่มีส่วนสำคัญ ผู้วิจัยควรศึกษาข้อดีและข้อจำกัดของโปรแกรมการวิเคราะห์ทางสถิติที่เหมาะสมสอดคล้องกับบริบทของงานวิจัย การศึกษาครั้งนี้เลือกใช้โปรแกรม R เนื่องจากเป็น Freeware ไม่เสียค่าใช้จ่าย สามารถดาวน์โหลดและทำการติดตั้งได้ทั้งระบบปฏิบัติการ window และ Mac ใช้งานได้ง่ายโดยการเรียกใช้ package มีคู่มือการใช้งานให้ศึกษาเพื่อเขียนคำสั่งในการวิเคราะห์ โดยเรียกการทำงานคู่กับ R Studio มีหน้าต่าง pop up เป็นคำแนะนำในการเขียนคำสั่งวิเคราะห์ที่สะดวกและง่ายต่อการใช้งาน

2. การวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากผลการวินิจฉัย ด้วยโมเดลวินิจฉัยเชิงจำแนก G-DINA model ได้ผลการวินิจฉัยทางพุทธิปัญญาที่มีความแม่นยำของการวินิจฉัย (Test Level Accuracy) ผ่านเกณฑ์ทุกด้าน และมีความแม่นยำของการวินิจฉัย รายคุณลักษณะ (Attribute Level Accuracy) ผ่านเกณฑ์ทุกคุณลักษณะ มีค่าพารามิเตอร์การเดา และค่าพารามิเตอร์ความสะเพร่าผ่านเกณฑ์ทุกข้อ ค่าดัชนีความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ (RMSEA) ส่วนใหญ่อยู่ในระดับต่ำ และผลการวิเคราะห์สามารถแบ่งกลุ่มแฝงของผู้สอบที่มีความสามารถต่างกันได้ แสดงให้เห็นว่า ข้อสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจที่ผู้วิจัยสร้างขึ้น เมื่อวิเคราะห์ด้วยโมเดลการวินิจฉัยประเภท Generalization DINA (G-DINA MODEL) จะสามารถวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 จำแนกตามทักษะ/คุณลักษณะที่จำเป็นต้องใช้เพื่อเป็นผู้รอบรู้/ไม่รอบรู้ในทักษะ/คุณลักษณะนั้น ๆ ได้อย่างน่าเชื่อถือ เนื่องจาก G-DINA Model เป็นโมเดลวินิจฉัยเชิงจำแนกที่มีข้อตกลง (Assumptions) ไม่เข้มงวด มีความอ่อนปรนเป็นโมเดลที่ชดเชยได้ กล่าวคือยินยอมให้ทักษะ/คุณลักษณะที่ต่ำกว่าหรือเป็นพื้นฐานหรือง่ายกว่า ถูกทดแทนได้ด้วยทักษะ/คุณลักษณะที่สูงกว่า สอดคล้องกับ ญัฐกานต์ ประจันบาน (2563, หน้า 94) กล่าวว่า ข้อตกลงเบื้องต้นเกี่ยวกับความน่าจะเป็นในการทำแบบสอบ มีความยืดหยุ่นสูง และไม่เคร่งครัดจนเกินไป เหมือนกับโมเดลการวินิจฉัยอื่นๆ ซึ่งทำให้ได้ผลการวิเคราะห์ ทางสถิติที่แม่นยำตรง มีประสิทธิภาพสูง มีความน่าเชื่อถือสูง (De la Torre., 2011; Rupp et al. ,2010; Basokcu et al. ,2013; Li et al. ,2016; Ma and De la Torre, 2019, อ้างถึงใน ญัฐกานต์ ประจันบาน, 2563, หน้า 94) ที่อธิบายว่า โมเดล G-DINA มีลักษณะชดเชยความสามารถได้ (Compensatory Model) และเป็นโมเดลที่ประหยัด (Parsimonious) สามารถแปลผลได้จาก 2 พารามิเตอร์ในแต่ละข้อ คือ พารามิเตอร์ความสะเพร่า และพารามิเตอร์การเดาข้อสอบถูก การประมาณค่าพารามิเตอร์ทำได้ง่าย ใช้เวลาน้อย โมเดลมีความ เหมาะสมสอดคล้องกับข้อมูล (Good Model Fit) ดีกว่าโมเดลอื่นๆ สามารถประมาณค่าพารามิเตอร์ได้ อย่างถูกต้อง และมีข้อจำกัดน้อยกว่า สามารถประเมินประสิทธิภาพของโมเดลโดยพิจารณาจากค่าสถิติ ความสอดคล้องเชิงสัมพัทธ์ เพราะฉะนั้นนักวิจัยและนักการศึกษาในสาขาวิชาการวัดและประเมินผล จึงสนับสนุนให้ผู้สนใจเลือกนำโมเดลการวินิจฉัยนี้ไปวิเคราะห์ข้อมูลเพื่อให้ได้ผลสรุปแบบสอบวินิจฉัย เป็นที่น่าพอใจ (Somrat, 2023, อ้างถึงใน ญัฐกานต์ ประจันบาน, 2563, หน้า 94) สอดคล้องกับ เกศรินทร์ สมราช (2563, หน้า 218) กล่าวว่า นักวิชาการด้านวัดและประเมินผล จึงสนับสนุนให้ผู้สนใจเลือกนำโมเดลการวินิจฉัยนี้ไปวิเคราะห์ข้อมูลซึ่งได้ผลสรุปแบบสอบวินิจฉัยเป็นที่น่าพอใจ ปัจจุบันมีการนำไปใช้ในการวินิจฉัยทางการศึกษา

กันอย่างหลากหลาย เช่น การวินิจฉัยความรู้ทาง คณิตศาสตร์ วิทยาศาสตร์ การอ่านจับใจความ ภาษาไทย และภาษาอังกฤษ

ในอีกประเด็นที่น่าสนใจ พบว่า มีผู้สอบจำนวน 34 คน ถูกจัดเข้าชั้นแ่งที่ 60 มีแบบแผนการตอบข้อสอบ “110111” หมายถึง ผู้สอบ 34 คน มีความรอบรู้คำศัพท์ สามารถค้นหาข้อมูลที่ระบุอย่างชัดเจน สามารถค้นหาสาระสำคัญของเรื่อง/ย่อหน้า สามารถตีความหรืออนุมานสิ่งที่ไม่ได้กล่าวถึงชัดเจนในบทอ่าน และสามารถสังเคราะห์ข้อมูลได้ โดยไม่ต้องรอบรู้ในเรื่องไวยากรณ์ นอกจากนี้ ชั้นแ่งที่ 57 แบบแผนการตอบ “000111” มีผู้สอบจำนวน 10 คน ถูกจัดเข้าชั้นแ่งดังกล่าว หมายถึง ผู้สอบสามารถค้นหาสาระสำคัญของเรื่อง สามารถตีความหรืออนุมานสิ่งที่ไม่ได้กล่าวถึงชัดเจนในบทอ่าน และสามารถสังเคราะห์ข้อมูลได้ โดยไม่ต้องรอบรู้คำศัพท์ ไวยากรณ์ และทักษะการค้นหาข้อมูลที่ระบุอย่างชัดเจนในบทอ่าน ทั้งนี้เนื่องมาจากโมเดล G-DINA เป็นโมเดลแบบชดเชยได้และยินยอมให้ทักษะที่แกร่งชดเชยทักษะที่อ่อน

ข้อเสนอแนะจากการวิจัย

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. แบบสอบแบบทดสอบเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษสำหรับฉบับนี้ มีแบบสอบจำนวน 18 ข้อ ครอบคลุม 6 คุณลักษณะ โดยที่ข้อคำถาม 1 ข้อสามารถวัดคุณลักษณะ ความสามารถในการอ่านเพื่อความเข้าใจ ได้มากกว่า 1 คุณลักษณะ และให้สารสนเทศย้อนกลับ (Skill Profile) เกี่ยวกับคุณลักษณะที่นักเรียนมีความรอบรู้ (Mastery) และคุณลักษณะที่นักเรียนไม่มีความรอบรู้ (Mastery) เป็นคำร้อยละ นักเรียนสามารถเข้าใจได้ง่าย นำไปสู่การปรับปรุงพัฒนาความสามารถในทักษะที่ยังไม่รอบรู้ รวมทั้งเพิ่มพูนทักษะความสามารถที่รอบรู้แล้วอีกด้วย รวมทั้งครูผู้สอนสามารถนำผลการวินิจฉัยและสารสนเทศย้อนกลับ ไปใช้ออกแบบการจัดการเรียนรู้ที่ตอบสนองทักษะความสามารถของผู้เรียนรายบุคคล

2. การวินิจฉัยทางพุทธิปัญญาเป็นการเสริมแรงเชิงบวกในกระบวนการเรียนรู้ กล่าวคือ การวัดประเมินผลการเรียนรู้ในห้องเรียนส่วนใหญ่ เป็นการประเมินเพื่อตัดสินผลการเรียน อาทิ “ผ่าน” หรือ “ไม่ผ่าน” “สอบได้” “สอบตก” หรือการทดสอบเพื่อตัดเกรด หรือการจัดลำดับคะแนน ซึ่งเกณฑ์การผ่านมักจะอยู่ที่ร้อยละ 50 ดังนั้น นักเรียนที่มีผลการทดสอบต่ำกว่าร้อยละ 50 จึงถูกตัดสินให้ “สอบตก” หรือ “ไม่ผ่าน” และวิธีการสอบแก้ตัวคือทำแบบสอบชุดเดิมซ้ำ ๆ จนกว่าจะได้คะแนนผ่านเกณฑ์ การกระทำเช่นนี้ นอกจากไม่เกิดประโยชน์ยังก่อให้เกิดผลเสียที่ร้ายแรงต่อนักเรียน ทำให้เกิดความเครียดและเบื่อหน่ายต่อการเรียน ปัญหาไม่ถูกแก้ไขและสะสมให้มากขึ้นได้ อย่างไรก็ตาม ปัญหาดังกล่าวอาจหมดไป หากครูผู้สอนนำการประเมินเพื่อวินิจฉัยจุดแข็งและจุดอ่อนของผู้เรียนมาใช้ในการประเมินผลระหว่างเรียน (Formative evaluation) เพื่อสำรวจความก้าวหน้า

ทางการเรียนของผู้เรียนเป็นรายบุคคล และแก้ไขทันที่ ตัวอย่างเช่น จากการวิจัยครั้งนี้ มีผู้สอบ R011 ได้คะแนนสอบ 6 คะแนน (คะแนนเต็ม 18) มีแบบแผนการตอบ “110000” หมายถึง มีความรอบรู้ ในเรื่องคำศัพท์ และทักษะการระบุข้อมูลที่ต้องการจากบทอ่าน ดังนั้นครูผู้สอนควรออกแบบ กิจกรรมการเรียนรู้ที่ให้ R011 นำความรู้คำศัพท์ และการค้นหาข้อมูลไปใช้เป็นพื้นฐานการเรียนรู้ ไวยากรณ์ ทักษะการหาสาระสำคัญ การอนุมานตีความ และสังเคราะห์ข้อมูลจากเรื่องที่อ่าน เป็นต้น

3. ครูผู้สอน หรือนักวิจัยที่เกี่ยวข้องทางการศึกษา สามารถนำโมเดลวินิจฉัยเชิงจำแนก G-DINA ซึ่งเป็นโมเดลแบบยืดหยุ่นให้ชัดเจนได้และมีความยืดหยุ่นสูง สามารถวิเคราะห์ด้วยโปรแกรม ทางสถิติที่ทำได้ง่าย และให้ประสิทธิภาพการวิเคราะห์ข้อมูลเชิงการวินิจฉัยดีกว่าโมเดลอื่น เพื่อ วินิจฉัยทักษะความสามารถในวิชาต่าง ๆ ได้

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. แบบสอบวินิจฉัยในการศึกษาครั้งนี้มีจำนวนเพียง 18 ข้อ ซึ่งหากผู้สอบทำแบบสอบ หลายครั้ง จะสามารถจดจำคำตอบได้ อาจส่งผลต่อผลการวินิจฉัย ดังนั้น นักวิจัยจึงควรพัฒนาเป็น คลังข้อสอบวินิจฉัย (Items Bank) สำหรับความสามารถการอ่านเพื่อความเข้าใจ ที่มีจำนวนหลาย ฉบับในรูปแบบแพลตฟอร์มออนไลน์ และเป็นแอปพลิเคชันที่สามารถใช้งานได้ด้วยโทรศัพท์มือถือ

2. การให้สารสนเทศผลการวินิจฉัยย้อนกลับไปยังผู้สอบ เพื่อที่จะทำให้ผู้สอบได้ทราบจุด แข็งและจุดที่ควรพัฒนาของตนเองในแต่ละทักษะในทันที ด้วยโปรแกรมทางคอมพิวเตอร์ free ware หรือแพลตฟอร์มออนไลน์ หรือแอปพลิเคชัน ออกแบบ Profile ที่มีความสวยงามดึงดูดความสนใจ และมีสารสนเทศที่มีประโยชน์ทางการวินิจฉัยถูกต้องครบถ้วนเพื่อนำไปใช้พัฒนาผู้เรียนต่อไป

บรรณานุกรม



บรรณานุกรม

- กระทรวงศึกษาธิการ. (2551). *หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551*. กรุงเทพฯ.
- กุลพร พูลสวัสดิ์. (2556). *การพัฒนาวิธีวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสาร สำหรับนักเรียนมัธยมศึกษาปีที่ 3*. (วิทยานิพนธ์ปริญญาคุชฎีบัณฑิต). กรุงเทพฯ: คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- เกศรินทร์ สมราช. (2563). *การพัฒนาแบบสอบวินิจฉัยความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารของนิสิตครู โดยประยุกต์ใช้ G-DINA MODEL*. (วิทยานิพนธ์ปริญญาคุชฎีบัณฑิต). พิษณุโลก: คณะศึกษาศาสตร์ มหาวิทยาลัยนเรศวร.
- แคทรียา แสงดาวเทียน. (2562). *การเปรียบเทียบประสิทธิภาพการประเมินวินิจฉัยเชิงพุทธิปัญญา เรื่องสมการเชิงเส้นตัวแปรเดียว ระหว่างโมเดลดีไอเอเอ็นไอ (DINO model) โมเดลดีไอเอเอ็นเอ (DINA model) และโมเดลจีดีไอเอเอ็นเอ (G-DINA model)*. (คุชฎีนิพนธ์). หลักสูตรปรัชญาคุชฎีบัณฑิต, ชลบุรี: มหาวิทยาลัยบูรพา.
- จกกล บัวแก้ว. (2560). *ประสิทธิผลของการเชื่อมโยงคะแนนสำหรับแบบสอบทดสอบทดสอบตามทฤษฎีการตอบสนองข้อสอบ*. (ปรัชญาคุชฎีบัณฑิต). ชลบุรี: มหาวิทยาลัยบูรพา.
- จกกล บัวแก้ว, ไพรัตน์ วงษ์นาม และสมพงษ์ ปั่นหุ่น. (2561). การประมาณค่าพารามิเตอร์ของแบบสอบทดสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองทดสอบ. *วารสารราชภัฏสุราษฎร์ธานี*, 5(1), 139-155.
- จันทร์ฉาย เศษแสงศรี. (2562). *ปัจจัยเชิงสาเหตุที่ส่งผลต่อความสามารถในการอ่านภาษาอังกฤษอย่างมีวิจารณญาณของนักเรียนชั้นมัธยมศึกษาตอนปลาย สังกัดสำนักงานเขตพื้นที่การมัธยมศึกษา เขต 27*
- จุกา ธรรมชาติ. (2554). *บทบาทของโปรแกรมช่วยวิเคราะห์เชิงคุณภาพ : การศึกษาโปรแกรม ATLAS.ti*. (วารสารศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี), 22(1).
- ฉลอง รัตนพงษ์. (2557). *การพัฒนารูปแบบการสอนการอ่านเพื่อความเข้าใจ โดยใช้ยุทธวิธีการสอนแบบแลกเปลี่ยนบทบาทและแผนภูมิความหมายเพื่อส่งเสริมความสามารถในการอ่านเพื่อความเข้าใจของนักศึกษาระดับปริญญาตรี*. (วิทยานิพนธ์)
- ฉัตรประวิณ. (2555). *การสังเคราะห์งานวิจัยที่เกี่ยวข้องกับการอ่านภาษาอังกฤษระดับชั้นมัธยมศึกษา*. (ปริญญาศึกษาศาสตรมหาบัณฑิต). มหาวิทยาลัยศิลปากร: กรุงเทพฯ.

- เฉลิมลาภ ทองอาจ. (13 ธันวาคม 2555). *คิดออกเสียง: การวัดประเมินความสามารถในการอ่านเพื่อความเข้าใจตามสภาพจริง*. สืบค้นเมื่อ 4 มกราคม 2565. จากเว็บไซต์ <https://www.gotoknow.org/posts/508264>
- ทิพย์ ขำอยู่, โสฬส สุขานนท์สวัสดิ์, และสุตาพรรณ อาจกล้า. (2557). ผลการวิจัยย้ยักษะการอ่านภาษาอังกฤษของนักศึกษาระดับปริญญาตรีโดยวิธีลำดับชั้นคุณลักษณะและการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์. (*วารสารการวัดผลการศึกษา*), 20(1), 103- 117.
- ชญารัตน์ บุญพุมิกร.(2561). *การพัฒนามาตรวัดเจตคติต่อวิชาชีพรูตามโมเดลการตอบสนองข้อสอบแบบพหุมิติ: การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์*. (หลักสูตรปรัชญาดุษฎีบัณฑิตสาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา). มหาวิทยาลัยบูรพา
- โชติกา ภาชีผล. (2554). *การสร้างและพัฒนาเครื่องมือในการวัดและประเมินผลการศึกษา*. (พิมพ์ครั้งที่ 2). กรุงเทพฯ: คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
- ณฐมน วงศ์ทาทอง (2560). *การพัฒนาความสามารถในการอ่านภาษาอังกฤษเพื่อความเข้าใจ ด้วยวิธีสอนอ่านแบบ SQ3R โดยใช้ข้อมูลอาเซียน สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6*. (วิทยานิพนธ์ศึกษาศาสตรมหาบัณฑิต สาขาวิชาหลักสูตรและการนิเทศ). มหาวิทยาลัยศิลปากร.
- ถิรวัดน์ ต้นทนิส. (2550). กระบวนการอ่านเพื่อความเข้าใจภาษาอังกฤษกับทฤษฎีพื้นฐานความรู้เดิม. (*วารสารอักษรศาสตร์ มหาวิทยาลัยศิลปากร*), 29(ฉบับพิเศษ), 120-133.
- นฤมล อุดรประจักษ์. (2555). *การสร้างแบบทดสอบวินิจฉัยข้อบกพร่องทางการเรียนวิชาคณิตศาสตร์ เรื่องระบบสมการเชิงเส้น ชั้นมัธยมศึกษาปีที่ 3*. (วิทยานิพนธ์ครุศาสตรมหาบัณฑิต, สาขาวิชาวิจัยและประเมินผลการศึกษา) คณะครุศาสตร์, มหาวิทยาลัยราชภัฏมหาสารคาม.
- พิสิฐ พิณจสกุล. (2563). ทฤษฎีการตอบสนองข้อสอบกับข้อสอบแบบเทสต์เลท. (*วารสารวิจัยและประเมินผลอุบลราชธานี*), 9(2), 74-84.
- พิศิษฐ ตัฒทวนิช. (2558). *การพัฒนาแบบสอบเพื่อการวินิจฉัยโดยแนวคิดด้านพุทธิปัญญา*. (*วารสารสถาบันวิจัยและพัฒนา*), มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา, 2(2), 1-13.
- ไพฑูรย์ เอื้อเอื้อบุญประดิษฐ์. (2563). *การพัฒนาแบบสอบวินิจฉัยการรู้เรื่องคณิตศาสตร์ที่มีการให้ข้อมูลย้อนกลับสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้วิธีลำดับชั้นของคุณลักษณะ*. สาขาวิชาการศึกษาและประเมินผลการศึกษา, คณะครุศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย.
- วรางคณา ก้อมน้อย. (2563). *การพัฒนาแบบทดสอบเชิงวินิจฉัยทางพุทธิปัญญา เรื่อง ความรู้สึกเชิงจำนวน สำหรับนักเรียนชั้นประถมศึกษาตอนต้น: การประยุกต์ใช้โมเดลจี ดีไอเอ็นเอ*. (วิทยานิพนธ์การศึกษาดุษฎีบัณฑิต), คณะศึกษาศาสตร์,มหาวิทยาลัยนเรศวร.

- สถาบันภาษาอังกฤษ สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน. (2560). *คู่มือการจัดการเรียนการสอนภาษาอังกฤษแนวใหม่ตามกรอบมาตรฐานความสามารถทาง ภาษาอังกฤษที่เป็นสากล*. กรุงเทพฯ: องค์การสงเคราะห์ทหารผ่านศึก.
- สุปราณี บุระ. (2557). *การพัฒนาแบบสอบเชิงวินิจฉัยทางพุทธิปัญญาเรื่องการดำเนินการเลขคณิตพื้นฐานโดยใช้คอมพิวเตอร์: การประยุกต์โมเดลไอเอ็นเอ*. กรุงเทพฯ: คณะครุศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย.
- สุมาลี มีสกุล. (2558). *การประเมินเพื่อวินิจฉัยความสามารถในการแก้ปัญหาเรื่องสมการเชิงเส้น ตัวแปรเดียวของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 โดยการประยุกต์ใช้โมเดล การวินิจฉัยเชิงจำแนก*. (วิทยานิพนธ์ปรัชญาดุษฎีบัณฑิต, สาขาวิชาการวิจัยและพัฒนาศักยภาพมนุษย์), มหาวิทยาลัยศรีนครินทรวิโรฒ.
- สังวรรณ ังตกระโทก. (2558). *การปฏิรูปการศึกษาวิทยาศาสตร์ที่ขับเคลื่อนด้วยข้อมูลเพื่อยกระดับคุณภาพการเรียนการสอนวิชา วิทยาศาสตร์ระดับมัธยมศึกษา* (รายงานผลการวิจัย). นนทบุรี: มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- ศิริชัย กาญจนวาสี. (2552). *ทฤษฎีการทดสอบแนวใหม่* (พิมพ์ครั้งที่3). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแนวใหม่* (พิมพ์ครั้งที่ 4). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศุภามณ จันท์สกุล. (2561). *การเปรียบเทียบผลการวินิจฉัยพุทธิปัญญาโดยใช้แผนที่ตัวแปรเชิงทฤษฎีและแบบทดสอบวินิจฉัยสามระดับในกระบวนการพยาบาลพื้นฐาน*. (สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์), จุฬาลงกรณ์มหาวิทยาลัย.
- อรพินทร์ ชูชม. (2555). *แรงจูงใจในการทำงาน: ทฤษฎีและการประยุกต์*. (วารสารมหาวิทยาลัยเกษมบัณฑิต), 2(2), 52–61.
- อรอนงค์ อุ่นแก้ว, ภัฐธีร์ตา วัฒนประดิษฐ์, และสมพงษ์ ปั่นหุ่น. (2562). *การสร้างแบบทดสอบวินิจฉัย*
 ความสามารถในการอ่านเพื่อความเข้าใจภาษาอังกฤษ สำหรับนักเรียนชั้น มัธยมศึกษาปีที่ 1 โรงเรียนในสังกัดกรมส่งเสริมการปกครองท้องถิ่น จังหวัดสมุทรสาคร. (*e-Journal*), 1(1), 54–67.
- Chen, J. (2014). *Model selection for IRT equating of Testlet-based tests in the random groups design* (Doctoral dissertation). University of Iowa, United States of America.

- Cohen, J. (1960). *A coefficient of agreement for nominal scales*. *Educational and Psychological Measurement*, 20, 37-46
- DeMars, C. (2010). *Item response theory: Understanding statistics measurement*. Oxford University Press: Oxford.
- de la Torre, J. (2011). *The generalized DINA model framework*. *Psychometrika*, 76(2), 179-199.
- Eckes, T., & Baghaei P. (2015). *Using testlet response theory to examine local dependence in C-Tests*. *Applied Measurement in Education*, 28(2), 85-98.
- Embretson, S. E. (1983). *Construct validity: Construct representation versus nomothetic span*. *Psychological Bulletin*, 93, 179-197.
- Embretson, S. E. (1994). *Applications of cognitive design systems to test development*. In C. R. Reynolds (Ed.), *Cognitive assessment: A multidisciplinary perspective* (pp. 107-135). New York: Plenum Press.
- Embretson SE, Yang X. (2013). *Multicomponent Latent Trait Model for Diagnosis*. *Psychometrika*. Jan; 78(1), 14-36. doi: 10.1007/s11336-012-9296-y.
- Jang, E. E. (2009). *Cognitive diagnostic assessment of L2 reading comprehension ability: Validity arguments for Fusion Model application to LanguEdge assessment*. *Language Testing*, 26(1), 031-073.
- Johnson M.S., & Sinharay, S. (2018). *Measurement of Agreement to Assess Attribute – Level Classification Accuracy and Consistency for Cognitive Diagnostic Assessments*. (*Journal of Education Measurement*), 54(4), 635 – 664. doi: 10.1111/jedm.12196.
- Muthén, L. K., Muthén, B. O. (2013). *Obtaining Diagnostic Classification Model Estimates Using M-plus*. *Educational Measurement Issues and Practice*, 32(2), 146 – 162.
- Primary National Strategy. (2005). *Understanding reading comprehension: 1. Primary head teachers, literacy coordinators, Key Stage 1 and 2 teachers Status: Recommended Date of issue: 03-2005 Ref: DfES 1310-2005 Department for Education and Skill*.

- Rupp, A. A., & Templin, J. L. (2008). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement: Interdisciplinary Research and Perspectives*, 6(4), 219-262. doi: 10.1080/15366360802490866
- Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications: Guilford Publications*.
- Sawaki, Y., Kim, H.-J., & Gentile, C. (2009). Q-Matrix construction: Defining the link between constructs and test items in large-scale reading and listening comprehension assessments. *Language Assessment Quarterly*, 6(3), 190-209. doi:10.1080/15434300902801917
- Tonekaboni et al., (2021). The Construction and Validation of a Q-matrix for a High-stakes Reading Comprehension Test: A G-DINA Study. Tabaran Institute of Higher Education. *International Journal of Language Testing* Vol. 11(1), 58-87.





ภาคผนวก

มหาวิทยาลัยนครพนม



ภาคผนวก ก

คู่มือการใช้สารสนเทศผลการวิจัยย้อนกลับ

มหาวิทยาลัยสุรินทร์

คู่มือการใช้สารสนเทศผลการวินิจฉัย ย้อนกลับ **Handbook of Diagnostic Feedback**

คู่มือฉบับนี้ประกอบด้วย 4 ส่วน ได้แก่
ส่วนที่ 1 วัตถุประสงค์
ส่วนที่ 2 ขอบเขตเนื้อหา
ส่วนที่ 3 ส่วนประกอบของ Profile
ส่วนที่ 4 การแปลผล และการนำไปใช้



ส่วนที่ 1 วัตถุประสงค์

Profile แสดงสารสนเทศจากการวินิจฉัยแบบย้อนกลับความสามารถด้านการอ่านเพื่อความเข้าใจของนักเรียนหรือผู้สอบรายบุคคลและภาพรวมทั้งในเชิงปริมาณและเชิงคุณภาพ เกี่ยวกับจุดแข็ง จุดอ่อน คุณลักษณะหรือทักษะย่อยที่จำเป็นต้องได้รับการปรับปรุงแก้ไข เพื่อครูผู้สอนหรือผู้เกี่ยวข้องนำไปใช้ออกแบบกิจกรรมการเรียนรู้ ที่ตอบสนองทุกรูปแบบความรอบรู้

ส่วนที่ 2 ขอบเขตเนื้อหา

นักเรียนจะได้รับการวินิจฉัยจุดแข็งและจุดอ่อน อ้างอิงเนื้อหาที่ประเมินตามกรอบมาตรฐาน CEFRในเรื่องต่อไปนี้

1) การอ่านคำแนะนำ (Reading Instruction) สามารถเข้าใจคำแนะนำที่ใช้ภาษาง่ายๆ และขั้นตอนการใช้ เช่น วิธีกดเงินจากตู้ ATM หรือวิธีซื้อเครื่องดื่มจากตู้อัตโนมัติ เครื่องจำหน่ายตัวสารธารณะ ข้อมูลเกี่ยวกับความปลอดภัย และการบอกทิศทาง เช่น Passive Voice ที่ใช้ในโครงสร้างประโยคง่ายๆ is/are + past participle คำสันธาน (conjunction) เช่น and/ but/ or/ before/ after/ because etc. ตัวเชื่อม (connective words) เช่น First, Second, Third, Fourth, Next, Then, Finally,... etc.

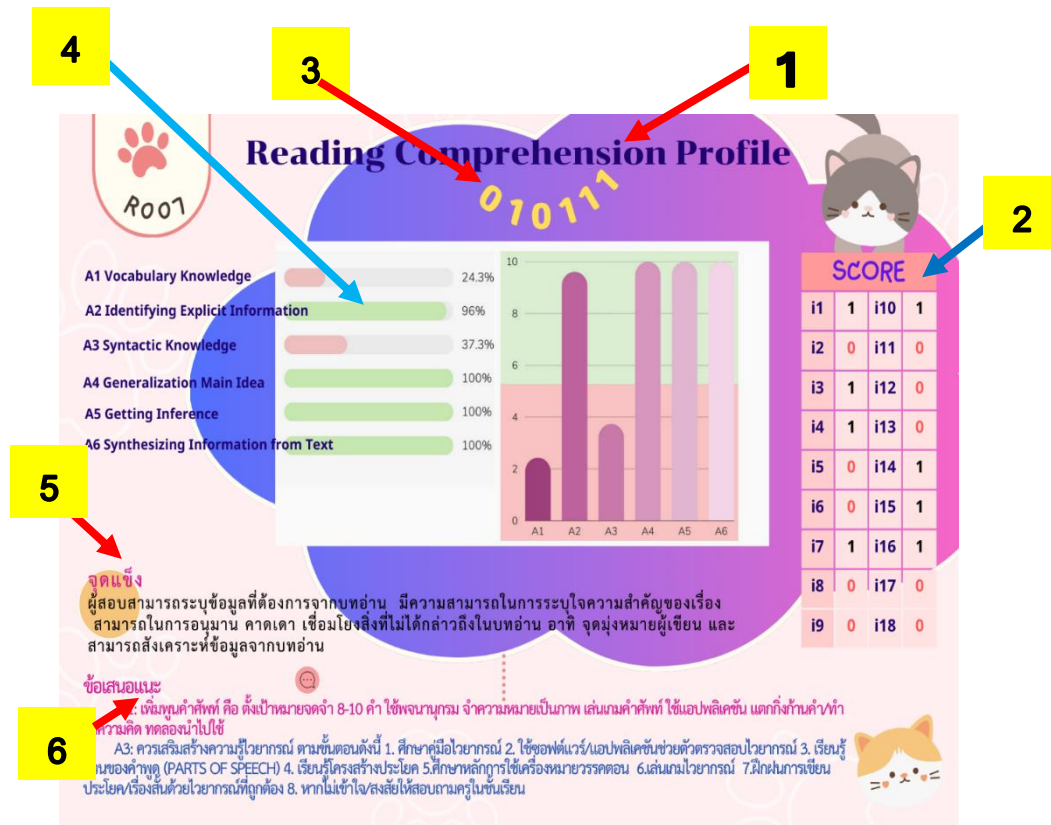
2) การอ่านข้อมูลเบื้องต้น (Reading for Orientation) สามารถค้นหาข้อมูลที่ต้องการและเข้าใจประเด็นหลัก, ใจความสำคัญ (main idea) ของข้อมูลนั้นๆ เช่น ราคา ขนาด ของโฆษณา หน้าเว็บ (Web page) แผ่นพับ หน้าเว็บ (web page) แคตตาล็อก ตารางเวลา เช่น หัวข้อเรื่องใจความสำคัญ รายละเอียดสนับสนุน คำถามเกี่ยวกับใจความสำคัญของเรื่อง เช่น ใคร ทำอะไร ที่ไหน เมื่อไร อย่างไร ทำไม ใช่หรือไม่ - Yes/No Question - Wh-Question etc.

3) การอ่านข้อมูลและข้อโต้แย้ง (Reading for Information & Argument) สามารถเข้าใจประเด็นหลักในบทความสั้น ๆ จากหนังสือพิมพ์ นิตยสาร เกี่ยวกับบุคคลที่มีชื่อเสียงในสังคมหรือเหตุการณ์ปัจจุบัน

4) การอ่านวรรณกรรม (Read Literature) สามารถเข้าใจประเด็นสำคัญ (Main idea) ของงานเขียน เช่น เรื่องสั้น (Story) บทกลอน (Poem) เกี่ยวกับชีวิตประจำวันซึ่งใช้ภาษาง่ายๆและมีสื่อหรือรูปภาพประกอบ

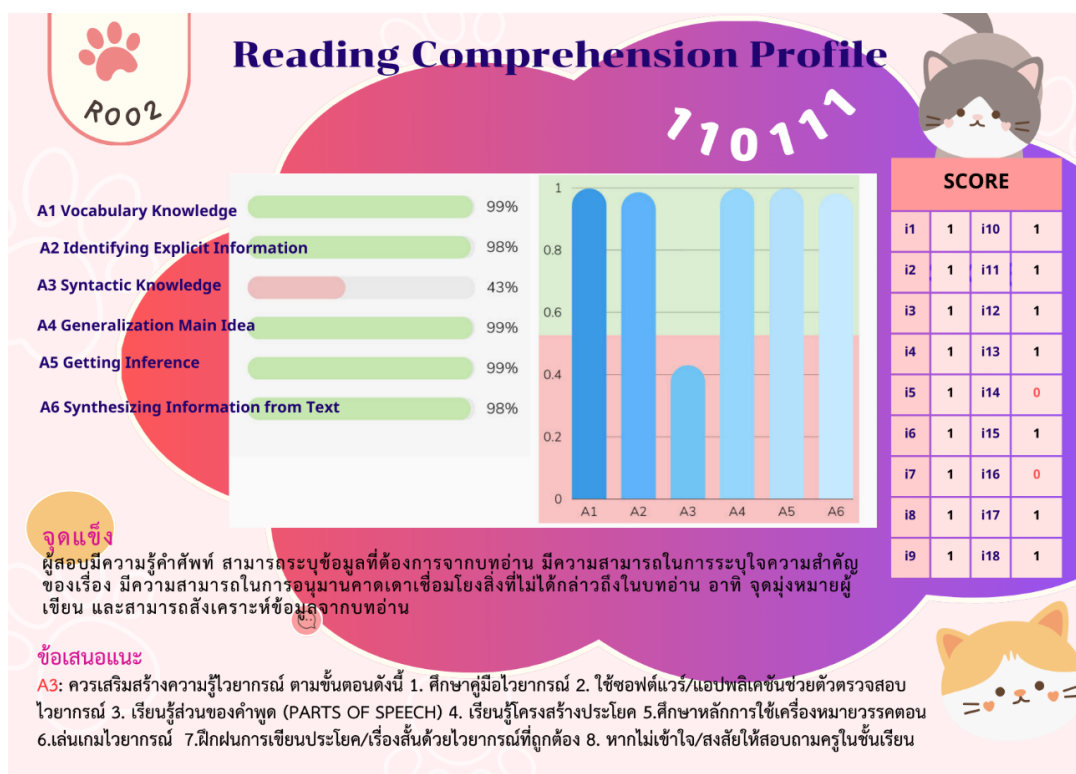
5) การอ่านเอกสารโต้ตอบหรือจดหมาย (Reading Correspondence) สามารถเข้าใจข้อความง่าย ๆ ในจดหมายอิเล็กทรอนิกส์ (E-mail) การสนทนาทางเว็บหรือแอปพลิเคชัน หรือข้อความจากเพื่อนหรือเพื่อนร่วมงานเกี่ยวกับเหตุการณ์ต่างๆในชีวิตประจำวัน เช่น การนัดหมายเวลาไปรับประทานอาหารกลางวัน /อาหารเย็น We shall meet at six p.m. at BN café. /When shall we meet? เป็นต้น

ส่วนที่ 3 ส่วนประกอบของ Profile



- 1 คือ ชื่อโปรไฟล์ (การอ่านเพื่อความเข้าใจ)
- 2 คือ คะแนนดิบจากการทำแบบสอบวินิจฉัยจำนวน 18 ข้อ
- 3 คือ รูปแบบความรอบรู้ (Mastery Pattern)
- 4 คือ สารสนเทศจากการวินิจฉัยของคุณลักษณะ 6 คุณลักษณะของความสามารถด้านการอ่านเพื่อความเข้าใจ
- 5 คือ จุดแข็งของผู้สอบ (ความสามารถตามคุณลักษณะแต่ละตัว)
- 6 คือ ข้อเสนอแนะเพื่อปรับปรุงพัฒนา

ส่วนที่ 4 การแปลผล และการนำไปใช้



จากตัวอย่าง Profile แสดงผลการวินิจฉัยความสามารถการอ่านเพื่อความเข้าใจภาษาอังกฤษ มีส่วนประกอบ 6 ส่วน ได้แก่ 1) รหัสผู้สอบ 2) ผลการทดสอบด้วยแบบสอบแบบทดสอบเลขเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจที่ผู้วิจัยพัฒนาขึ้น จำนวน 18 ข้อ (ตอบถูกได้ 1 ตอบผิดได้ 0) 3) แบบแผนการตอบข้อสอบ 4) แผนภูมิแสดงระดับความรอบรู้ในแต่ละคุณลักษณะของผู้สอบ 5) สารสนเทศแสดงจุดแข็งของผู้สอบ และ 6) สารสนเทศแสดงข้อเสนอแนะเชิงบวกเพื่อการปรับปรุงจุดอ่อนเกิดประโยชน์ในการพัฒนาตนเองให้มีความรอบรู้ต่อไป

ผู้สอบ R002 มีผลการทดสอบด้วยแบบสอบแบบทดสอบเลขเพื่อวินิจฉัยทางพุทธิปัญญาความสามารถการอ่านเพื่อความเข้าใจเท่ากับ 16 คะแนน ดังนี้ 11111111111101011 มีแบบแผนการตอบข้อสอบ “110111” จากแผนภูมิ แปลความหมายได้ว่า ผู้สอบ R002 มีความรอบรู้ใน 5 คุณลักษณะ โดยแท่งสีสูงเกิน แถบสีแดง ซึ่งเป็นสัญลักษณ์ของขอบเขตที่แสดงว่ายังไม่มีความรู้ ดังนั้น ผู้สอบคนนี้มีมีความรอบรู้ในคุณลักษณะ A1: มีความรู้คำศัพท์ เท่ากับร้อยละ 99.89 มีความรอบรู้ในคุณลักษณะ A2: สามารถระบุข้อมูลที่ต้องการจากบทอ่าน เท่ากับร้อยละ 98.62 มีความรอบรู้ในคุณลักษณะ A4: มีความสามารถในการระบุนใจความสำคัญของเรื่อง ร้อยละ 99.82 มีความรอบรู้ในคุณลักษณะ A5: มีความสามารถในการอนุมาน ร้อยละ 99.97 และมีความรอบรู้ในคุณลักษณะ A6: มีความสามารถสังเคราะห์ข้อมูลจากบทอ่านเท่ากับร้อยละ 98.29 R002 มีจุดแข็ง A1: มีความรู้คำศัพท์

A2: สามารถระบุข้อมูลที่ต้องการจากบทอ่าน A4: มีความสามารถในการระบุใจความสำคัญของเรื่อง
 A5: มีความสามารถในการอนุมานคาดเดาเชื่อมโยงสิ่งที่ไม่ได้กล่าวถึงในบทอ่าน อาทิ จุดมุ่งหมาย
 ผู้เขียน A6: สามารถสังเคราะห์ข้อมูลจากบทอ่าน และควนปรับปรุงและได้รับการพัฒนาในด้าน
 ความรู้ไวยากรณ์ ตามขั้นตอนดังนี้ 1. ศึกษาคู่มือไวยากรณ์ 2. ใช้ซอฟต์แวร์/แอปพลิเคชันช่วยตัว
 ตรวจสอบไวยากรณ์ 3. เรียนรู้ส่วนของคำพูด (parts of speech) 4. เรียนรู้โครงสร้างประโยค 5.
 ศึกษาหลักการใช้เครื่องหมายวรรคตอน 6.เล่นเกมไวยากรณ์ 7.ฝึกฝนการเขียนประโยค/เรื่องสั้นด้วย
 ไวยากรณ์ที่ถูกต้อง 8. หากไม่เข้าใจ/สงสัยให้สอบถามครูในชั้นเรียน

การนำไปใช้ประโยชน์

การให้ข้อมูลย้อนกลับ (Feedback) ถือเป็นหัวใจสำคัญของการประเมิน
 ความก้าวหน้าในการเรียนของผู้เรียน (Hattie & Timberley, 2007) และเป็นองค์ประกอบสำคัญใน
 การทำให้นักเรียนเกิดการเรียนรู้ มีประโยชน์ในการเปลี่ยนแปลงพฤติกรรมในการเรียนการสอน
 ทุกรูปแบบครูต้องคำนึงถึงวิธีการและเวลาที่ให้ข้อมูลจึงจะเกิดผลสำเร็จในการจัดการเรียนการสอน
 ข้อมูลย้อนกลับมีผลกระทบในด้านบวกและด้านลบโดยการให้ข้อมูลย้อนกลับจะให้แก่ผู้เรียนภาย
 หลังจากที่ได้ตอบสนองต่อการเรียนรู้ครั้งแรก เมื่อได้ให้ข้อมูลย้อนกลับแก่ผู้เรียนในมิติความรู้
 ทักษะต่างๆ ที่ได้เรียนไปแล้วโดยเฉพาะอย่างยิ่งหากเป็นการให้ข้อมูลย้อนกลับเพื่อชี้แจงความ
 ผิดพลาดที่เกิดขึ้นจะทำให้ผู้เรียนเกิดความรู้ความเข้าใจที่ดียิ่งขึ้น ข้อมูลย้อนกลับเป็นสารสนเทศที่
 ส่งเสริมสนับสนุนความสามารถของแต่ละบุคคล มีวัตถุประสงค์หลัก คือ การลดช่องว่างระหว่าง
 ความรู้ความสามารถในปัจจุบัน และความรู้ความสามารถตามเป้าหมายที่ครู ผู้เรียนหรือผู้ปกครอง
 ต้องการโดยมีการสร้างเป้าหมายร่วมกันตั้งแต่เริ่มต้น และยังสามารถลดความแตกต่างระหว่างความรู้
 ความเข้าใจในสภาพปัจจุบันที่เป็นอยู่และสภาพที่ต้องการให้เกิดขึ้นได้อีกด้วย ซึ่งการที่จะลดความ
 ต่างหรือช่องว่างระหว่างสภาพปัจจุบันและสภาพที่ต้องการให้เกิดขึ้นโดยผ่านการให้ข้อมูล
 ย้อนกลับนั้น หากเป็นการแก้ปัญหาที่ยากหรืองานที่มีความท้าทาย และมีเป้าหมายที่คาดหวังไว้
 ชัดเจน ผู้เรียนจะมีความพยายามมากกว่าการปฏิบัติกิจกรรมเดิมซ้ำๆหรืองานที่ง่ายเกินไป ผู้เรียนยัง
 สามารถพัฒนาทักษะในการติดตามข้อผิดพลาดของตนเองจากข้อมูลย้อนกลับของตนเอง เพื่อให้ไปถึง
 ยังเป้าหมายที่ตั้งไว้ นอกจากนี้ผู้เรียนยังสามารถแสวงหากลยุทธ์ที่ดีขึ้นเพื่อให้งานหรือการเรียนรู้จาก
 การได้รับข้อมูลเพิ่มเติมเพื่อพวกเขาจะสามารถแก้ปัญหาปัญหาหรือใช้เทคนิคที่มีในกำกับ ดูแลตนเอง
 ครูผู้สอนมีบทบาทในการช่วยลดความแตกต่างหรือช่องว่างระหว่างความรู้ ทักษะในปัจจุบัน และ
 ความรู้ ทักษะที่ต้องการให้เกิดขึ้นกับผู้เรียน ซึ่งหมายถึงครูอาจจะต้องช่วยผู้เรียนในการกำหนด
 เป้าหมายการเรียนรู้ที่มีความชัดเจนมากยิ่งขึ้น ซึ่งเป้าหมายที่มีความชัดเจน (Specific Goals) นั้นจะ
 ดีกว่าเป้าหมายที่ไม่เฉพาะเจาะจง เพราะส่วนใหญ่เน้นไปที่ความสนใจของผู้เรียนและเน้นการให้ข้อมูล

ย้อนกลับที่มีข้อมูลสารสนเทศเกี่ยวกับการนำไปใช้แก้ไขปัญหามากกว่าเป้าหมายทั่วไปที่ไม่เฉพาะเจาะจง การที่ผู้อ่านจะเกิดความเข้าใจได้นั้นต้องอาศัยองค์ประกอบดังนี้

1. เข้าใจความหมายของคำ (Word meaning) การเข้าใจความหมายคำศัพท์ช่วยให้เข้าใจประโยคได้ดี

2. เข้าใจหน่วยความคิด (Thought unit) การอ่านทีละคำ จะเป็นอุปสรรคต่อความเข้าใจ เพราะขาดความต่อเนื่อง ดังนั้นผู้เรียนจึงต้องได้รับการฝึกการอ่านเป็นหน่วยความคิด

3. เข้าใจประโยค (Sentence comprehension) เมื่อผู้อ่านสามารถหาหน่วยความคิดย่อยสัมพันธ์กันจนได้ใจความประโยคแล้ว ผู้อ่านจะต้องมองเห็นความสัมพันธ์ระหว่างคำ และระหว่างกลุ่มคำในประโยคด้วย

4. เข้าใจอนุเขต (Paragraph) ผู้อ่านต้องสามารถใจความสำคัญของแต่ละประโยคแล้วนำใจความสำคัญนั้นมาสัมพันธ์กัน ดังนั้นผู้อ่านจึงควรเรียนรู้ประเภท หรือโครงสร้างแบบต่างๆ ของอนุเขตเพื่อให้สามารถระบุประโยคที่เป็นใจความหลักในอนุเขตนั้น จึงมองหาความสัมพันธ์ของประโยคอื่นๆ ซึ่งเป็นรายละเอียดในอนุเขต

5. เข้าใจเนื้อหาทั้งหมด (Comprehension for larger unit) การเข้าใจเรื่องราวต่างๆ ทั้งหมดผู้อ่านต้องความสำคัญระหว่างข้อความสำคัญๆ ในแต่ละตอน เพื่อลำดับความคิดของเรื่องและต้องทราบว่าข้อเขียนแต่ละประเภทมีโครงสร้างข้อความแตกต่างกัน



ภาคผนวก ข

การใช้แพลตฟอร์ม Flexi Quiz ในขั้นตอนวินิจฉัย

มหาวิทยาลัยบูรเวศวร

Flexi Quiz. Com

แสดงตัวอย่าง Platform Online เรียกว่า Flexi Quiz. Com ซึ่งเป็นวิธี/เครื่องมือในการเก็บรวบรวมข้อมูลการวินิจฉัยในระยะที่ 2 ที่ผู้วิจัยใช้ในการวินิจฉัยความสามารถด้านการอ่านเพื่อความเข้าใจ ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 เข้าหน้าเว็บไซต์ได้โดยคลิกลิงก์ <https://www.flexiquiz.com/>

The screenshot shows the Flexi Quiz platform interface. At the top, there is a navigation bar with 'Dashboard', 'Users', 'Groups', 'Reports', and 'Settings'. Below this is a secondary bar with 'Create', 'Configure', 'Publish', and 'Analyze'. The main content area displays a 'Welcome page' for a 'Reading Comprehension Diagnostic Test for Beginner'. The test details include: 'Directions: This diagnostic test is multiple choice that include four passages for answering 18 questions. Testlets form was applied for cognitive diagnostic assessment. The content of test consists of four topics as follows: Topic 1: READ INSTRUCTIONS, Topic 2: READ FOR ORIENTATION & READ CORRESPONDENCE, Topic 3: READ INFORMATIONS & ARGUMENT, Topic 4: READ LITERATURE. >>To choose the best answer.' The browser address bar shows 'flexiquiz.com/SC/N/1ab07d0a-4b4c-4c33-a83b-4e1bbc3902b2?preview=1'. The browser's taskbar at the bottom shows several open tabs, including 'จักษุศาสตร์', 'วิจัยในชั้น', 'การวัดและประเมินผล...', 'การวัดและประเมินผล...', 'PKP Submissions | Politic...', '(1280) โปรแกรม R พ...', and 'PQ Random or fixed tes...'. A large blue banner at the bottom of the page reads 'Reading Comprehension'.

Reading Comprehension Diagnostic Test for Beginner

Directions: This diagnostic test is multiple choice that include four passages for answering 18 questions. Testlets form was applied for cognitive diagnostic assessment. The content of test consists of four topics as follows:

Topic 1: READ INSTRUCTIONS

Topic 2: READ FOR ORIENTATION & READ CORRESPONDENCE

Topic 3: READ INFORMATIONS & ARGUMENT

Topic 4: READ LITERATURE

>>To choose the best answer.

Start Quiz

เมื่อคลิกที่เมนู Start Quiz จะปรากฏข้อสอบข้อที่ 1 และเวลาในการทำแบบทดสอบ
 วินิจฉัยปรากฏมุมมองบนด้านขวามือ

Time remaining
0:29:52






Reading Comprehension

Topic 1: READ INSTRUCTIONS


Read the following text and these pictures for answering items 1-5


1. Read the following instructions and rearrange them according to the correct order of the pictures.

(A) Check the coin chute for any change.
 (B) Swipe your credit card if it's a newer machine.
 (C) Insert the correct amount of money for the item.
 (D) Type in the code or press the button for your item.
 (E) Find the price and code underneath the item you wish to buy.

Picture 1





E C B D A
 B D C E A
 C B D A E
 E B A C D

Page: 1
Answered 0 of 10 (0%)
Next Page >



ผู้สอบสามารถเลือกย้อนกลับและทำข้อสอบหน้าต่อไปได้

Reading Comprehension

Time remaining
0:27:15

4. Where **can't** you generally find the vending machine?

A vending machine is an automated machine that provides items such as snacks, beverages, cigarettes and lottery tickets to consumers after cash, a credit card, or a specially designed card is inserted into the machine. The first modern vending machines were developed in England more than one hundred thirty years and dispensed postcards. It's as simple as putting in money and hitting the button for the item you want. Later, the Sweetmeat Automatic Delivery Company was founded in 1887 in England as the first company to deal primarily with the installation and maintenance of vending machines. It is so popular because people love the convenience.

The first vending machine in the U.S. was built in 1888 by the Thomas Adams Gum Company, selling gum on New York City train platforms. The idea of adding games to these machines as a further incentive to buy came in 1897 when the Pulver Manufacturing Company added small figures, which would move around whenever somebody bought some gum from their machines. The main example of a vending machine giving access to all merchandise after paying for one item is a newspaper vending machine (also called vending box) found mainly in the U.S. and Canada. It contains a pile of identical newspapers.

Source: https://en.wikipedia.org/wiki/Vending_machine

- department store
- airport
- bus terminal
- farm

Page: 4

Answered 0 of 18 (0%)

« Previous Page

Next Page »



ในหน้าสุดท้าย เมื่อทำข้อสอบเสร็จจะมีเมนูให้ผู้สอบกดส่งคำตอบ โดยคลิกที่ Submit จากนั้นโปรแกรมจะประมวลผลการสอบ หากผู้สอบผ่านร้อยละ 70 จะมีเกียรติบัตรส่งไปในอีเมลล์ ที่ผู้สอบลงทะเบียนไว้

Reading Comprehension
Time remaining
0:24:48

18. Which should be the best title for passage?

The unique annual celebration – that sees the local villagers dress in distinctive and boldly colorful masks and headgear to partake in the carnival atmosphere – is a popular date on the festival calendar. The most recognized feature and a highlight of the event is the parade of the costumed villagers through the town.

The Story of the Ghost Festival of Thailand dates back to a long time ago. It is said Prince Vessantara went on a long journey. It was a journey so long that most of his followers believed he was dead. Surprisingly, one day, he returned, alive and in good health. His followers rejoiced, and the celebrations that followed were so strident that it resulted in the awakening of the dead. Since then, the celebrations have continued.

The attractive celebrations have become the most interesting festival. Villagers dressed up in the ghost masks during the three days in a part of Bun Phawet, Buddhist merit-making festival, make merry by participating in parties and parades on the Dan Sai district in Loei province. Skillfully handcrafted from dried sticky husks, these colorful masks have long noses.

The primary attention of every festival-goer would fall on the great monstrous mask, with terrifying teeth and evil eyes. Legend says that it is meant bad luck to keep your mask after the festival gets over, but as times have changed, people now sell their masks and make good money out of it. The masks catch the eye of tourists, They especially by looking at them.


The local population of Loei is always friendly and welcoming. They will be happy to share their culture with you and make sure that you are an integral part of their celebrations. You will also be excited to pose for photos so keep that camera ready to catch the excitement as this is a memory you wouldn't want to forget ever.

Source: <https://www.activity.com/pages/phi-ta-khon-1213.html>

Mask Festival
 Skillfully handcrafted
 The story of Halloween
 Phi Ta Khon in Loei

Page: 18
Answered 0 of 18 (0%)

[← Previous Page](#)
[Submit](#)



Certificate

of achievement

Wilaiporn Khamman

For successful completion of Reading Comprehension.

Awarded on the 27th February 2024