

การเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ของแบบทดสอบ
ความเข้าใจในการอ่านที่มีจำนวนข้อและจำนวนกลุ่มตัวอย่างต่างกัน
ที่วิเคราะห์ด้วยราส์ชโมเดล

ปริญาพนธ์
ของ
เกศมณี พยัคฆ์

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
ตุลาคม 2543
ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

การเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ของแบบทดสอบ
ความเข้าใจในการอ่านที่มีจำนวนข้อและจำนวนกลุ่มตัวอย่างต่างกัน
ที่วิเคราะห์ด้วยวิธีราส์โมเดล

บทคัดย่อ
ของ
เกศมณี พยัคฆ์

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

ตุลาคม 2543

1395016

1395016 2543

เกศมณี พยัคฆ์. (2543). การเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ของแบบทดสอบความเข้าใจในการอ่านที่มีจำนวนข้อและจำนวนกลุ่มตัวอย่างต่างกันที่วิเคราะห์ด้วยราล์ชโมเดล. ปรินซ์ตันเพนน์ กทม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. คณะกรรมการควบคุม : รองศาสตราจารย์ อังคณา สายยศ, รองศาสตราจารย์ชูศรี วงศ์รัตน์นะ.

การวิจัยครั้งนี้มีจุดมุ่งหมาย เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ และค่าความสามารถของผู้สอบของแบบทดสอบความเข้าใจในการอ่านที่มีข้อสอบ 20 และ 40 ข้อ และจำแนกตามกลุ่มตัวอย่าง 100 , 150, 200, 250 และ 300 คน ซึ่งได้จากการสุ่มแบบใส่คืน จากกลุ่มตัวอย่าง 606 คน ที่สุ่มมาแบบแบ่งชั้นของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2542 ของโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบแบบเลือกตอบความเข้าใจในการอ่านภาษาไทย และวิเคราะห์ข้อมูลด้วย โปรแกรม BICAL วัดตามทฤษฎีการตอบสนองข้อสอบแบบราล์ชโมเดล ผลการวิจัยสรุปดังนี้

1. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ระหว่าง 20 และ 40 ข้อ เมื่อกลุ่มตัวอย่าง เท่ากันคือ 100 , 150, 200, 250 และ 300 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ
2. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างข้อสอบที่มีจำนวน 20 และ 40 ข้อ เมื่อกลุ่มตัวอย่าง เท่ากัน คือ 100, 150, 200, 250 และ 300 คน มีค่า แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ
3. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 20 และ 40 ข้อ ระหว่างกลุ่มตัวอย่างที่มีจำนวน 100 คน, 150 คน, 200 คน, 250 คน และ 300 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ ระดับ .01
4. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ของข้อสอบ 20 และ 40 ข้อ ระหว่างกลุ่มตัวอย่างที่มีจำนวน 100 คน , 150 คน , 200 คน, 250 คน และ 300 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

A COMPARISON OF STANDARD ERROR OF PARAMETER ESTIMATION OF
READING COMPREHENSION TEST BETWEEN DIFFERENCE TEST LENGTHS
AND SAMPLE SIZES ANALYZE BY RASCH MODEL

AN ABSTRACT
BY
KETMANEE PAYUK

Presented in partial fulfillment of the requirements
for the Master of Education degree in Educational Measurement
at Srinakharinwirot University

October 2543

Ketmanee Payuk. (2000). *A comparison of standard error of parameter estimation of reading comprehension test between difference test lengths and sample sizes analyze by Rasch Model*. Master thesis M.Ed. (Educational Measurement). Bangkok : Graduate School Srinakharinwirote University. Adviser Committee : Assoc. Prof. Aungkana Saiyos, Assoc. Prof. Chusri Wongrattana.

The purpose of this research were to compare the standard errors of items difficulty and abilities estimations of reading comprehension test with length of 20 and 40 items. The sample sizes had 100, 150, 200, 250 and 300 students with repeated random sampling from 606 students of Prathomsuksa VI in the second semester of the academic year 1999 in Prachinburi Province were selected by Stratified Random Sampling technique. The research instrument used were reading comprehension test. The Rasch Model in Item Response Theory by the BICAL Computer Program was used to estimate item difficulty and examinee's ability.

1. The standard error of items difficulty estimations between test length of 20 and 40 items with the same sample sizes about 100, 150, 200, 250 and 300 were not statistical significant.

2. The standard error of abilities estimations between test length of 20 and 40 items with the same sample sizes about 100, 150, 200, 250 and 300 were not statistical significant..

3. The standard error of items difficulty estimations of test length about 20 and 40 items between 100, 150, 200, 250 and 300 sample size were statistical significant at .01 level.

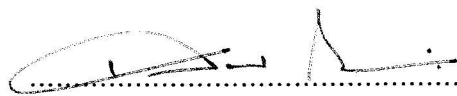
4. The standard error of abilities estimations of test length about 20 and 40 items between 100, 150, 200, 250 and 300 sample size were not statistical significant.

ปริญญานิพนธ์
เรื่อง

การเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ของแบบทดสอบ
ความเข้าใจในการอ่านที่มีจำนวนข้อและจำนวนกลุ่มตัวอย่างต่างกัน
ที่วิเคราะห์ด้วยราส์ชโมเดล

ของ
นางสาวเกศมณี พยัคฆ์

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
มหาวิทยาลัยศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย

(ศาสตราจารย์ ดร. เสริมศักดิ์ วิศาลาภรณ์)

วันที่ ๕ เดือน ตุลาคม พ.ศ. 2543

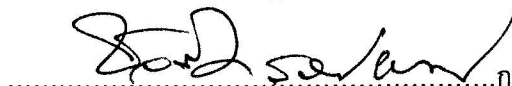
คณะกรรมการสอบปริญญานิพนธ์

.....ประธาน

(รองศาสตราจารย์อังคณา สายยศ)

.....กรรมการ

(รองศาสตราจารย์ชูศรี วงศ์รัตน)

.....กรรมการที่แต่งตั้งเพิ่มเติม

(อาจารย์ชวลิต รวยอาจิน)

.....กรรมการที่แต่งตั้งเพิ่มเติม

(รองศาสตราจารย์วิบูลา วิศาลาภรณ์)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จลงได้ด้วยความกรุณา และความอนุเคราะห์อย่างยิ่งจาก รองศาสตราจารย์อังคณา สายยศ ประธานกรรมการควบคุมปริญญานิพนธ์ รองศาสตราจารย์ชูศรี วงศ์รัตนะ กรรมการควบคุมปริญญานิพนธ์ อาจารย์ชวลิต รวยอาจิณ และรองศาสตราจารย์วิญญา วิศาลาภรณ์ คณะกรรมการแต่งตั้งเพิ่มเติม ในการสอบปากเปล่าปริญญานิพนธ์ที่ได้กรุณาในการให้ คำแนะนำและข้อคิดเห็นต่าง ๆ อันเป็นประโยชน์ยิ่งในการปรับปรุงแก้ไขให้การวิจัยครั้งนี้สมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณอย่างสูงไว้ ณ ที่นี้

ขอกราบขอบพระคุณรองศาสตราจารย์นิภา ศรีไพโรจน์, ดร.สุพร เข้มเฮง, ผู้ช่วยศาสตราจารย์ระวีวรรณ พันธุ์พานิช, รองศาสตราจารย์วิญญา วิศาลาภรณ์ และอาจารย์ ชวลิต รวยอาจิณ ที่ให้ความกรุณาเป็นผู้เชี่ยวชาญตรวจสอบความเที่ยงตรงของเครื่องมือวิจัยครั้งนี้

ขอกราบขอบพระคุณผู้บริหารโรงเรียน ผู้ช่วยฝ่ายวิชาการ ครูประจำชั้นประถมศึกษาปีที่ 6 ของโรงเรียนประถมศึกษา สังกัดสำนักงานคณะกรรมการการประถมศึกษาจังหวัดปราจีนบุรี ที่ เอื้อเฟื้อให้ความร่วมมือ และความอนุเคราะห์ในการเก็บรวบรวมข้อมูลเพื่อการวิจัยในครั้งนี้

ผู้วิจัยขอขอบพระคุณเพื่อนๆ ชาววัดผลการศึกษาที่ให้ความช่วยเหลือและเป็นกำลังใจแก่ ผู้วิจัยโดยสม่ำเสมอ

สุดท้ายผู้วิจัยขอน้อมรำลึกถึงพระคุณบิดา-มารดา ครู-อาจารย์ ที่ให้คำแนะนำ ให้การ สนับสนุนทุกสิ่งทุกอย่างเท่าที่จะทำได้หรือให้ได้ จนผู้วิจัยสำเร็จการศึกษา

เกษมณี พยัคฆ์

สารบัญ

บทที่

หน้า

1 บทนำ.....	1
ภูมิหลัง.....	1
ความมุ่งหมายของการวิจัย.....	3
ความสำคัญของการวิจัย.....	3
ขอบเขตของการวิจัย.....	4
นิยามคำศัพท์เฉพาะ.....	4
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	6
ราล์ฟโมเดล.....	16
วิธีการประมาณค่าพารามิเตอร์.....	18
คุณภาพของแบบทดสอบ.....	19
งานวิจัยที่เกี่ยวข้อง.....	21
สมมติฐานการวิจัย.....	25
3 วิธีดำเนินการศึกษาค้นคว้า.....	26
ประชากร.....	26
กลุ่มตัวอย่าง.....	26
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	27
การสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	28
ลักษณะของแบบทดสอบ.....	30
วิธีดำเนินการเก็บรวบรวมข้อมูล.....	31
ขั้นตอนในการวิเคราะห์ข้อมูล.....	31
สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	32
4 ผลการวิเคราะห์ข้อมูล.....	35
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล.....	35
การเสนอผลวิเคราะห์ข้อมูล.....	35
ผลการวิเคราะห์ข้อมูล.....	36

สารบัญ (ต่อ)

บทที่	หน้า
5 สรุป อภิปรายและข้อเสนอแนะ.....	44
ความมุ่งหมายในการศึกษาค้นคว้า.....	44
กลุ่มตัวอย่าง.....	44
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	44
วิธีดำเนินการเก็บรวบรวมข้อมูล.....	44
สรุปผลการวิเคราะห์ข้อมูล.....	45
อภิปราย.....	47
 บรรณานุกรม.....	 48
 ภาคผนวก.....	 51
 ประวัติย่อผู้วิจัย.....	 73

บัญชีตาราง

ตาราง

หน้า

1 แสดงคุณสมบัติของแบบทดสอบ 5 ข้อ.....	20
2 จำนวนกลุ่มตัวอย่างตามโรงเรียน.....	27
3 ค่าความยากของข้อสอบ ค่าความยากเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน และค่าความสามารถของผู้สอบ ของแบบทดสอบความเข้าใจในการอ่าน จำนวน 20 ข้อ และ 40 ข้อ ที่มีจำนวนกลุ่มตัวอย่าง 100, 150, 200, 250, และ 300 คน.....	36
4 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากและค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบทดสอบความเข้าใจในการอ่าน จำนวน 20 ข้อ และ 40 ข้อ ที่มีจำนวนกลุ่มตัวอย่าง 100, 150, 200, 250, และ 300 คน.....	38
5 เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างข้อสอบ 20 ข้อ และ 40 ข้อ เมื่อกลุ่มตัวอย่าง 100, 150, 200, 250 และ 300 คน	39
6 เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างข้อสอบ 20 ข้อ และ 40 ข้อ เมื่อกลุ่มตัวอย่าง 100, 150, 200, 250 และ 300 คน	40
7 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 20 ข้อ ระหว่างกลุ่มตัวอย่าง 100, 150, 200, 250 และ 300 คน.....	40
8 เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 20 ข้อ เป็นรายคู่.....	41
9 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างกลุ่มตัวอย่าง 100, 150, 200, 250 และ 300 คน เมื่อข้อสอบ 40 ข้อ.....	41
10 เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 40 ข้อ เป็นรายคู่.....	42
11 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างกลุ่มตัวอย่าง 100, 150, 200, 250 และ 300 คน เมื่อข้อสอบ 20 ข้อ.....	43
12 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างกลุ่มตัวอย่าง 100, 150, 200, 250 และ 300 คน เมื่อข้อสอบ 40 ข้อ.....	43
13 แสดงค่าความยากและค่าอำนาจจำแนกของแบบทดสอบเลือกตอบความเข้าใจในการอ่าน จำนวน 60 ข้อ ในการทดสอบครั้งที่ 1.....	54

บัญชีตาราง (ต่อ)

ตาราง	หน้า
14 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 20 ข้อ กลุ่มตัวอย่าง 100 คน.....	60
15 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 20 ข้อ กลุ่มตัวอย่าง 150 คน.....	61
16 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 20 ข้อ กลุ่มตัวอย่าง 200 คน.....	62
17 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 20 ข้อ กลุ่มตัวอย่าง 250 คน.....	63
18 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 20 ข้อ กลุ่มตัวอย่าง 300 คน.....	64
19 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 40 ข้อ กลุ่มตัวอย่าง 100 คน.....	65
20 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 40 ข้อ กลุ่มตัวอย่าง 150 คน.....	67
21 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 40 ข้อ กลุ่มตัวอย่าง 200 คน.....	69
22 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 40 ข้อ กลุ่มตัวอย่าง 250 คน.....	71
23 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 40 ข้อ กลุ่มตัวอย่าง 300 คน.....	73

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 แสดงความสัมพันธ์ระหว่าง ระดับความสามารถของผู้สอบกับโอกาสที่จะตอบถูก.....	8
2 ความหมายของค่าพารามิเตอร์ของข้อสอบของรูปแบบลอจิสต์ แบบ 3 พารามิเตอร์.....	9
3 แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์ตัวเดียว.....	11
4 แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์สองตัว.....	12
5 แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์สามตัว.....	12
6 แสดงความหมายของค่าItem Parameter ของ One-Parameter Logistic Model.....	13
7 แสดงความหมายของค่าItem Parameter ของ Two-Parameter Logistic Model.....	14
8 แสดงความหมายของค่าItem Parameter ของ Three-Parameter Logistic Model.....	15
9 การเปรียบเทียบความสามารถกับความยาวของข้อสอบ.....	16
10 แสดงความแตกต่างระหว่างความสามารถของคนกับความยากของข้อสอบ.....	17
11 แสดง Information Curve ของแบบทดสอบความถนัดด้านภาษาของสถาบัน ETS.....	20
12 แสดงขั้นตอนการสร้างแบบทดสอบเลือกตอบ.....	27

บทที่ 1

บทนำ

ภูมิหลัง

การวัดผลการศึกษาและจิตวิทยาเป็นการวัดความสามารถ หรือคุณลักษณะแฝง (Ability or Latent Trait) ของบุคคล เช่นความสามารถ ความถนัด และเจตคติ เป็นต้น ซึ่งไม่สามารถวัดได้โดยตรง เหมือนกับการวัดทางกายภาพ ต้องอาศัยการวัดทางอ้อม โดยใช้ข้อคำถามในแบบทดสอบเป็นสื่อ เพื่อกระตุ้นให้บุคคลแสดงพฤติกรรมออกมาด้วยการตอบแบบทดสอบ แล้วนำผลการสอบหรือคะแนน (Performance or Score) ไปอธิบายหรือทำนายความสามารถที่แท้จริงของบุคคลที่ถูกวัด (พิชญา เกตุสกุล. 2539) การวัดหรือการประมาณค่าความสามารถโดยอาศัยการวัดทางอ้อมนี้ มักจะมีความคลาดเคลื่อนในการวัดรวมอยู่ด้วย ซึ่งทำให้ข้อสรุปที่ได้ขาดความถูกต้อง ดังนั้นนักวัดผลจึงได้พยายามสร้างเครื่องมือและพัฒนาทฤษฎีการวัดต่างๆขึ้นมา เพื่อที่จะได้สรุปอ้างอิงผลของการวัดซึ่งเป็นตัวแทนของความสามารถ หรือคุณลักษณะแฝงของบุคคล ให้มีความถูกต้อง และเชื่อถือได้มากที่สุด ทฤษฎีการวัดที่นิยมแพร่หลายได้แก่ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory : CTT) และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT)

ทฤษฎีการทดสอบแบบดั้งเดิม เป็นทฤษฎีการวัดที่ใช้กันอย่างแพร่หลายมาตั้งแต่อดีตจนถึงปัจจุบัน ซึ่งเน้นการประมาณค่าคะแนนจริง หรือความสามารถที่แท้จริง โดยใช้คะแนนของกลุ่มผู้สอบเป็นหลัก ซึ่งถูกมองว่าเป็นทฤษฎีที่ไม่เพียงพอต่อการประมาณค่าความสามารถที่แท้จริง และทฤษฎีนี้ไม่สามารถทำนายได้ว่า บุคคลจะตอบข้อสอบอย่างไร ยกเว้นแต่ได้ใช้ข้อสอบเหล่านั้นกับบุคคลที่คล้ายคลึงกันมาก่อน (ผจงจิต อินทสุวรรณ. 2525 : 51-69) และถ้ากลุ่มตัวอย่างมีขนาดเล็ก ก็จะทำให้ค่าความยากสูงเกินความจำเป็น และทำให้ค่าอำนาจจำแนกมีความคลาดเคลื่อน (Henrysson. 1976) ซึ่งแฮมเบิลตัน (Hambleton. 1989 : 147-148) ได้กล่าวถึงจุดอ่อนของทฤษฎีแบบดั้งเดิมไว้หลายประการคือ

1. ค่าสถิติของข้อสอบ ได้แก่ค่าความยาก และค่าอำนาจจำแนกของข้อสอบมีค่าไม่คงที่ จะขึ้นอยู่กับกลุ่มตัวอย่างที่ใช้วิเคราะห์ ถ้ากลุ่มตัวอย่างมีลักษณะต่างกัน ค่าความยากและค่าอำนาจจำแนกจะมีค่าไม่เท่ากัน ความไม่เป็นอิสระจากกลุ่มตัวอย่าง ทำให้ประโยชน์ของการใช้ค่าสถิตินี้ในการพัฒนาแบบทดสอบและประยุกต์ใช้ในกรณีอื่นลดลง

2. การเปรียบเทียบความสามารถของผู้สอบแต่ละคน จะใช้ผลการสอบจากแบบทดสอบฉบับเดียวกัน หรือแบบทดสอบคู่ขนานกันเท่านั้น ถ้าใช้ข้อคำถามไม่เหมือนกันจะไม่สามารถนำมาเปรียบเทียบกันได้ เพราะคะแนนของผู้สอบแต่ละคนขึ้นอยู่กับค่าสถิติของข้อสอบ

3. ความเชื่อมั่นของแบบทดสอบ ที่นิยามในลักษณะของแบบทดสอบคู่ขนานนั้นเป็นไปได้ยากที่จะสัมฤทธิ์ผลในทางปฏิบัติ เนื่องจากไม่สามารถสร้างแบบทดสอบให้คู่ขนานกันได้อย่างแท้จริง และผลการตอบของผู้สอบอาจไม่เหมือนเดิมในการตอบครั้งที่สอง

4. ไม่มีพื้นฐานในการที่จะบอกว่าผู้สอบคนหนึ่งมีความน่าจะเป็นในการตอบข้อสอบข้อหนึ่งได้ถูกต้องเพียงใด ความน่าจะเป็นในการตอบสนองได้ถูกต้องมีประโยชน์ในการปรับปรุงแบบทดสอบ

ให้เหมาะสมกับความสามารถระดับของผู้สอบ ซึ่งจะทำให้สามารถวัดความสามารถของผู้สอบได้อย่างแท้จริง

5. การกำหนดข้อตกลงเบื้องต้นที่เกี่ยวกับความแปรปรวนของความคลาดเคลื่อนในการวัดของผู้สอบทุกคนมีค่าเท่ากัน ซึ่งขัดแย้งกับลอร์ดและโนวิก (พิชญา เกตุสกุล. 2539 : 2 ; อ้างอิงมาจาก Lord and Novick. 1968. *Statistical Theory of Mental Test Scores* : p.183) ที่พบว่าความคลาดเคลื่อนมาตรฐานในการวัดสำหรับผู้สอบที่มีความสามารถต่ำ จะมีค่ามากกว่าผู้สอบที่มีความสามารถปานกลางและความสามารถสูง

ดังนั้นด้วยขีดจำกัดต่าง ๆ ดังกล่าวแล้ว ขณะนี้ได้มีการวิเคราะห์ข้อทดสอบอีกแบบหนึ่งเรียกว่า ทฤษฎีการตอบสนองข้อสอบ ซึ่งทฤษฎีนี้มีชื่อเรียกต่าง ๆ กัน เช่น ทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) ทฤษฎีโค้งลักษณะข้อสอบ (Item Characteristic Curve Theory) หรือทฤษฎีการสอบแบบใหม่ (Modern Test Theory) เป็นต้น มีทั้งแบบ 3 พารามิเตอร์ (3-parameter model) แบบ 2 พารามิเตอร์ (2-parameter model) และแบบ 1 พารามิเตอร์ (1-parameter model) ซึ่งในรูปแบบหลังนี้อาจเรียกอีกอย่างหนึ่งได้ว่าแบบราสช์ (Rasch model) ในที่นี้จะขอกล่าวถึงเฉพาะแบบราสช์ ซึ่งนักทดสอบเชื่อว่าสามารถแก้ปัญหาทั้ง 5 ข้อดังกล่าวแล้วได้เพราะว่า (สุพัฒน์ สุขมณลสันต์. 2542 : 40)

1. การวิเคราะห์ข้อทดสอบแบบราสช์ มีสมมติฐานว่าตัวเลือกของข้อทดสอบทุก ๆ ข้อมีอำนาจจำแนกเท่ากัน ดังนั้นเรื่องการเดาคำตอบเป็นสิ่งที่ไม่ต้องคำนึงถึง
2. ผลของการวิเคราะห์ข้อทดสอบแบบราสช์จะทราบได้ว่าข้อทดสอบใดบ้างที่วัดเนื้อหาเดียวกันได้ (Pollitt 1980 : 3)
3. ผลของการวิเคราะห์ข้อทดสอบแบบราสช์ ทำให้ทราบได้ว่าลำดับที่มาตรฐานของความยากของข้อทดสอบแต่ละข้อเป็นอย่างไร ลำดับที่เหล่านี้นักไม่เปลี่ยนแปลง แม้ว่าแบบทดสอบนั้นจะใช้กับกลุ่มตัวอย่างที่มีธรรมชาติแตกต่างกันก็ตาม และยังเป็นอันดับที่มาตรฐานต่อเนื่องเป็นขั้นบันไดด้วย
4. สมมติฐานของการวิเคราะห์แบบทดสอบด้วยราสช์ มีเหตุผลที่น่าเชื่อถือมากกว่าการวิเคราะห์ข้อทดสอบด้วยวิธีทั่วไปกล่าวคือ เชื่อว่าการที่ผู้สอบจะทำข้อทดสอบแต่ละข้อได้ถูกต้องหรือไม่ขึ้นอยู่กับระดับความสามารถของผู้สอบ และระดับความยากของข้อทดสอบนั้นเท่านั้น ไม่ได้ขึ้นอยู่กับปัจจัยอื่นอีกเลย เช่น ไม่ต้องคำนึงถึงความยาวของแบบทดสอบ หรือขนาดของกลุ่มตัวอย่างแต่อย่างใด

นอกจากนี้ โกวิท ประวาลพุกษ์ (2525 : 63) ได้กล่าวถึงคุณสมบัติของข้อสอบที่วิเคราะห์ด้วยทฤษฎีการตอบข้อคำถามว่า สามารถเปรียบเทียบค่าความยากง่ายได้ โดยอิสระและเป็นสถิติที่เพียงพอสำหรับการใช้อธิบายคุณภาพของแบบทดสอบ นอกจากนี้ แฮมเบิลตัน (Hambleton.) เบรินบอม (Birnbaum.) ลอร์ด (Lord.) ราสช์ (Rasch.) วาร์ม (Warm.) ต่างยืนยันว่า ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบข้อคำถามนั้น ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบที่แตกต่างกัน (สุพล นิลกลาง 2541: 3 ; อ้างอิงจาก Hambleton. 1985. *Item response Theory*. ; Birnbaum. 1958. *Statistical Theories of Mental Test Scores* ; Lord. 1989. *Statistical Theory of Mental Test Scores* ; Rasch. 1982. *PROBABILISTIC models for Some Intelligence and Attainment tests* ; Warm. 1978. *Technical Report 941278*.) สำหรับ คูก และทาล์ฟ (Cook and Talf. 1988 : 31-34) ได้ทำการศึกษาความคงที่ของพารามิเตอร์ ที่วิเคราะห์ด้วยทฤษฎีการตอบข้อคำถาม ผลการศึกษาพบว่า ค่าพารามิเตอร์ จะเปลี่ยนเมื่อระยะเวลา หรือความยาวของแบบทดสอบที่มีความยาวต่างกัน ส่วนบุญชิต รอดแก้ว (2533 : 75) และ

สมพร บุญอ้อม (2529 : 92) ได้ศึกษาค่าพารามิเตอร์ของข้อสอบ พบว่าค่าพารามิเตอร์ของข้อสอบจะแปรปรวน เมื่อลักษณะของกลุ่มผู้สอบต่างกัน

จากคำกล่าวและผลการศึกษาข้างต้นจะเห็นว่า ผลการศึกษาและทฤษฎียังมีการขัดแย้งกันอยู่ ในประเทศไทยก็มีผู้ทำการศึกษาความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ของแบบทดสอบอยู่น้อยมาก ผู้วิจัยจึงสนใจที่จะศึกษาความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ของข้อสอบ ที่มีจำนวนข้อ และจำนวนกลุ่มตัวอย่างต่างกัน โดยจำนวนข้อที่ใช้ในการทดลองผู้วิจัยใช้ 20 ข้อ และ 40 ข้อ ซึ่งผู้วิจัยได้มาจากผลการทดลองของ Andre De Champlsin and Mare E. Gessaroli. ที่กล่าวว่าจำนวนข้อสอบในแบบทดสอบโดยค่าความคลาดเคลื่อนมาตรฐานต่ำมีความสัมพันธ์กับแบบทดสอบที่มีขนาดยาว (Champlsin and E. Gessaroli .1998, 231-253) และในการวิจัยใช้กลุ่มตัวอย่าง 100 คน 150 คน 200 คน 250 คน และ 300 คน เพราะต้องการทราบว่าในการวิเคราะห์แบบ รัสซโสมเดลนั้น ใช้จำนวนกลุ่มตัวอย่างต่ำกว่า 200 คนแล้วจะทำให้ ค่าความคลาดเคลื่อนในการประมาณค่าของแบบทดสอบเปลี่ยนไปหรือไม่ โดยปกติการวิเคราะห์แบบ รัสซโสมเดลจะต้องใช้กลุ่มตัวอย่าง 200 คนขึ้นไป ดังนั้นผู้วิจัยจึงมีความสนใจที่จะทำการศึกษาเรื่องนี้

ความมุ่งหมายของการวิจัย

1. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างข้อสอบที่มีจำนวนข้อต่างกัน เมื่อมีกลุ่มตัวอย่างจำนวนเท่ากัน
2. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อมีข้อสอบจำนวนข้อเท่ากัน
3. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างข้อสอบที่มีจำนวนข้อต่างกัน เมื่อมีกลุ่มตัวอย่างจำนวนเท่ากัน
4. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่าง กลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อมีข้อสอบจำนวนข้อเท่ากัน

ความสำคัญของการวิจัย

ผลการศึกษาครั้งนี้ ทำให้ทราบว่า การวิเคราะห์แบบทดสอบวัดความเข้าใจในการอ่านภาษาไทยที่วิเคราะห์ด้วย รัสซโสมเดล จะต้องใช้ขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบอย่างน้อยที่สุดเท่าใดจึงจะสามารถประมาณค่าพารามิเตอร์ความยากของข้อสอบให้สัมพันธ์กับค่าความสามารถของผู้สอบ เมื่อวิเคราะห์ด้วย รัสซโสมเดล อันจะเป็นประโยชน์ต่อการพัฒนาจำนวนข้อสอบให้เหมาะสมกับขนาดกลุ่มตัวอย่าง ทั้งนี้เพื่อเป็นการประหยัดค่าใช้จ่ายและเวลา ตลอดจนให้ผลการวิเคราะห์ของแต่ละบุคคลคงที่และไม่เปลี่ยนแปลงไปตามข้อสอบที่ใช้ อันจะเป็นประโยชน์ต่อการวัดผลการศึกษาต่อไป

ขอบเขตของการวิจัย

1. ประชากรที่ใช้ในการศึกษาครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 โรงเรียนในสังกัดสำนักงาน การประถมศึกษาจังหวัดปราจีนบุรี จำนวน 273 โรงเรียน จำนวนห้องเรียน 328 ห้องเรียน รวมนักเรียน 5,290 คน

2. กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนในสังกัด สำนักงานการประถมศึกษาจังหวัดปราจีนบุรี จำนวน 20 โรงเรียน จำนวนนักเรียน 606 คน ซึ่งเลือกมาโดย การสุ่มแบบแบ่งชั้น

3. ตัวแปรที่ศึกษา

1) ตัวแปรอิสระ มีดังนี้

1.1 จำนวนข้อสอบ จำแนกดังนี้

1.1.1 40 ข้อ

1.1.2 20 ข้อ

1.2 จำนวนกลุ่มตัวอย่าง จำแนกดังนี้

1.2.1 กลุ่มผู้สอบจำนวน 100 คน

1.2.2 กลุ่มผู้สอบจำนวน 150 คน

1.2.3 กลุ่มผู้สอบจำนวน 200 คน

1.2.4 กลุ่มผู้สอบจำนวน 250 คน

1.2.5 กลุ่มผู้สอบจำนวน 300 คน

2) ตัวแปรตาม มีดังนี้

2.1 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความยากของ ข้อสอบ

2.2 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถ ของผู้สอบ

นิยามศัพท์เฉพาะ

1. ทฤษฎีการตอบข้อคำถาม (Item Response Theory:IRT) หมายถึงทฤษฎีที่อธิบายถึงความ น่าจะเป็นในการตอบข้อสอบได้ถูกต้องในรูปของความสัมพันธ์ของความสามารถที่แท้จริงของผู้สอบกับการ ตอบข้อสอบ โดยทฤษฎีนี้มีความเชื่อว่าพฤติกรรมการตอบสนองต่อข้อสอบของบุคคลจะถูกกำหนดโดย ลักษณะความสามารถอยู่ภายในตัวบุคคลไม่สามารถสังเกตได้

2. ราสช์โมเดล (Rasch Model) หมายถึง รูปแบบการวิเคราะห์หาค่าพารามิเตอร์ของราสช์ที่ใช้ รูปแบบทางคณิตศาสตร์คำนวณหาค่าความสามารถของผู้สอบ(θ) กับความยากของข้อสอบ (b) ซึ่งเป็น ส่วนหนึ่งของทฤษฎีการตอบข้อคำถาม แบบหนึ่งพารามิเตอร์ ในการศึกษาค้นคว้าครั้งนี้จะใช้โปรแกรม BICAL

3. การวิเคราะห์ด้วยราส์โมเดล หมายถึง การวิเคราะห์ข้อสอบด้วยโมเดลของราส์ ที่ใช้ทฤษฎีการตอบข้อคำถาม ที่ใช้ ค่าพารามิเตอร์ 1 ตัว คือค่าความยาก (b) ในการศึกษาครั้งนี้วิเคราะห์ด้วยโปรแกรม BICAL

4. ค่าพารามิเตอร์ หมายถึง ตัวเลขที่แสดงคุณลักษณะของข้อสอบ ซึ่งในการวิจัยครั้งนี้ศึกษา ลักษณะค่าพารามิเตอร์ของข้อสอบ คือ ค่าความยากของข้อสอบ และค่าความสามารถของผู้สอบ ซึ่งได้จากการวิเคราะห์ด้วยราส์โมเดล

5. ค่าความยากของข้อสอบ (Item Difficulty : b) หมายถึงค่าความสามารถตรงจุดเปลี่ยนโค้ง คุณลักษณะข้อสอบแต่ละข้อ (Item Characteristic Curve : ICC) เป็นระดับความสามารถที่จะตอบข้อนั้นถูกต้อง ด้วยความน่าจะเป็น .50 เมื่อไม่มีการเดา ซึ่ง มีค่า อยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติมีค่า อยู่ระหว่าง -3.00 ถึง 3.00 ค่า -3.00 แสดงว่าข้อสอบนั้นง่ายมาก ค่า 3.00 แสดงว่าข้อสอบนั้นยากมาก ซึ่งในการวิจัยครั้งนี้ ประเมินค่าตามทฤษฎีการตอบข้อสอบโดยใช้ราส์โมเดล โดยใช้โปรแกรม BICAL

6. ค่าความสามารถของผู้สอบ (Examinee's Ability : θ) หมายถึงระดับความสามารถของผู้สอบแต่ละคน ซึ่งวัดในหน่วย θ ที่มีค่าเฉลี่ยเท่ากับ 0 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 1 สำหรับค่าความสามารถมีค่าเป็นไปได้ตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติจะมีค่าอยู่ระหว่าง -3 ถึง +3 ค่า-3 แสดงว่ามีความสามารถต่ำสุด ค่า +3 แสดงว่ามีความสามารถสูงสุด

7. ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Standard Error of Estimate) หมายถึง ค่าที่เบี่ยงเบนไปจากความเป็นจริงของคะแนนของผู้สอบ ซึ่งมีลักษณะดังนี้คือ

7.1 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากข้อสอบ คือ ค่าประมาณความยากของข้อสอบที่เบี่ยงเบนไปจากค่าความยากที่แท้จริงของข้อสอบ

7.2 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ (Standard Error of Ability Estimation) คือ ค่าประมาณความสามารถของผู้สอบที่เบี่ยงเบนไปจากค่าความสามารถที่แท้จริง

8. จำนวนข้อสอบต่างกัน หมายถึงจำนวนข้อของแบบทดสอบที่มีจำนวนไม่เท่ากัน การวิจัยครั้งนี้แบบทดสอบที่ใช้เป็นแบบทดสอบวัดความเข้าใจในการอ่าน จำนวน 2 ฉบับที่มีจำนวนข้อไม่เท่ากัน ได้แก่แบบทดสอบที่มีจำนวนข้อ 20 ข้อ และ 40 ข้อ

9. จำนวนกลุ่มตัวอย่าง หมายถึง กลุ่มตัวอย่างที่ทดสอบด้วยแบบทดสอบความเข้าใจในการอ่าน จำนวน 5 กลุ่ม ที่ได้มาโดยการสุ่มแบบแบ่งชั้น ที่มี จำนวน 100 คน 150 คน 200 คน 250 คน และ 300 คน

10. แบบทดสอบความเข้าใจในการอ่าน หมายถึง ชุดของข้อคำถามวิชาภาษาไทยแบบเลือกตอบ ที่ผู้วิจัยสร้างขึ้นเอง เพื่อวัดความสามารถในการแปลความ การตีความ การขยายความ และจับใจความของสถานการณ์ที่กำหนดให้อ่าน อันได้แก่ ข้อความ บทสนทนา โคลง กลอน ที่มีความยาก-ง่ายเหมาะสมกับนักเรียนชั้นประถมศึกษาปีที่ 6 ซึ่งตรวจสอบความเหมาะสมของข้อความโดยผู้เชี่ยวชาญ

11. ผู้เชี่ยวชาญ หมายถึง ผู้เชี่ยวชาญทางด้านการวัดผลการศึกษา และภาษาไทยจำนวน 5 ท่าน เพื่อพิจารณาความเหมาะสมของการใช้ภาษาที่ใช้ในข้อความที่กำหนดให้ และพิจารณาความเที่ยงตรงความเป็นปรนัยของข้อสอบ

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้ศึกษาในเรื่องต่าง ๆ ดังนี้

1. ทฤษฎีการตอบข้อคำถาม
 - 1.1 พื้นฐานความเข้าใจในทฤษฎีการตอบข้อคำถาม
 - 1.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม
 - 1.3 พิสัยของค่าพารามิเตอร์
 - 1.4 โมเดลของทฤษฎีการตอบข้อคำถาม
2. ราส์ชโมเดล
3. วิธีการประมาณค่าพารามิเตอร์
4. คุณภาพของแบบทดสอบ
5. งานวิจัยที่เกี่ยวข้อง

1. ทฤษฎีการตอบข้อคำถาม

1.1 พื้นฐานความเข้าใจในทฤษฎีการตอบข้อคำถาม

ทฤษฎีการตอบข้อคำถาม (Item Response Theory) เป็นทฤษฎีที่ศึกษาเกี่ยวกับความสัมพันธ์ของการตอบข้อคำถามและคุณลักษณะภายใน (Latent Trait) ของผู้สอบ กล่าวคือ การตอบข้อสอบของผู้สอบสามารถได้รับการทำนายได้ด้วยคุณลักษณะเฉพาะของผู้ตอบ กล่าวอีกนัยหนึ่ง คือเราสามารถนำระดับความสามารถหรือคุณลักษณะที่กะประมาณได้ ไปทำนายหรืออธิบายการตอบข้อสอบถูกของผู้สอบได้ (Lord and Novick, 1968 ; Hambleton and Swaminathan, 1985 : 9) เป้าหมายหลักของทฤษฎีการตอบข้อคำถาม คือต้องการหาคุณภาพรายข้อของแบบทดสอบและการกะประมาณความสามารถของผู้สอบซึ่งต้องอาศัยความสอดคล้องกันระหว่างโมเดลและข้อมูล (Baker, 1987 :121 ; citing Yen, 1981. *Review of educational Research.*)

ผู้ริเริ่มให้แนวคิดและหลักการเบื้องต้นของทฤษฎีการตอบข้อคำถาม คือ เฟอร์ดิงสัน และ ลอเลย์ (Lord and Novick, 1968 : 369 ; citing Ferguson and Lawley, 1942, 1943. *Educational and Psychological Measurement*) ต่อมาปี ค.ศ. 1952 ลอร์ด (Ferderic M. Lord) ได้เสนอทฤษฎีในรูปโค้งลักษณะข้อสอบ (Item Characteristic Curve Theory) ใจความสำคัญของทฤษฎีนี้กล่าวว่าโอกาสในการตอบข้อสอบถูกหรือไม่ ขึ้นอยู่กับความยากของข้อสอบข้อนั้นกับระดับความสามารถของผู้สอบ (Trait or Ability) ซึ่งกำหนดด้วยฟังก์ชันทางคณิตศาสตร์ที่เรียกว่าเส้นโค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) โค้งลักษณะของข้อสอบแต่ละข้อมีลักษณะเป็นโค้งปกติสะสม หรือที่เรียกว่า โมเดลโค้งปกติสะสม (Normal Ogive Model) โมเดลนี้จะกล่าวถึงพารามิเตอร์ของข้อสอบ 2 ค่า คือ ค่าความยาก (b) และค่าอำนาจจำแนก (a) แต่ความก้าวหน้าของทฤษฎีนี้ในระยะแรกเป็นไปอย่างเชื่องช้ามาก เนื่องจากความยุ่งยากในการคำนวณและขาดแคลนเครื่องคอมพิวเตอร์ที่ใช้วิเคราะห์ข้อมูล ทำให้ทฤษฎีนี้ชะงักการพัฒนอยู่ระยะหนึ่ง

ในปี ค.ศ. 1960 ราล์ฟ (Georg Rasch) ได้ศึกษาและเสนอแนวคิดในรูปของพารามิเตอร์เพียงตัวเดียว คือ ค่าความยาก (b) ที่เรียกว่า ราล์ฟโมเดล (Rasch Model) และในปี ค.ศ. 1965 ลอร์ด ได้กลับมาสนใจทฤษฎีนี้ใหม่ (Warm. 1978 : 19)

ในปี ค.ศ. 1968 เบิร์นบอม (Birnbaum) ได้ศึกษาและเสนอแนวคิดในรูปของโลจิสติกโมเดล (Logistic Model) ที่ใช้พารามิเตอร์ 2 ตัว คือค่าความยากและค่าอำนาจจำแนก ซึ่งเป็นโมเดลที่ง่ายกว่าโมเดลของลอร์ด จึงทำให้เป็นที่นิยมใช้กันอย่างแพร่หลายและมีการพัฒนาขึ้นมาเรื่อยๆ จนสามารถพัฒนาใช้กับโมเดลพารามิเตอร์ตัวเดียว และโมเดลพารามิเตอร์สามตัว (Warm. 1979 : 19-21) ส่วนแม็คโดแนลด์ บาดัน และลอร์ด (บุญชิต รอดแก้ว. 2533 : 7-8 ; อ้างอิงจาก McDonald. 1967. *Educational and Psychological Measurement* ; Barton and Lord. 1981. *Educational Research Methodology' and Measurement.*) ได้พัฒนาโลจิสติกโมเดลที่ใช้กับโมเดลพารามิเตอร์สี่ตัว (Hambleton and Swaminathan. 1985 : 48-49)

จากการที่ได้ศึกษาเกี่ยวกับทฤษฎีการตอบข้อคำถามซึ่งได้มีนักการศึกษาได้ให้ถึงความหมาย และลักษณะของทฤษฎีการตอบข้อคำถามดังนี้

ลอร์ดและโนวิก (Lord and Novick. 1968 : 358) กล่าวว่าทฤษฎีการตอบข้อคำถามเป็นทฤษฎีที่อธิบายหรือทำนายความสัมพันธ์ระหว่างลักษณะ หรือความสามารถที่มีอยู่ในตัวบุคคลกับพฤติกรรมการตอบสนองต่อข้อสอบของบุคคลนั้น โดยทฤษฎีนี้เชื่อว่าพฤติกรรมการตอบสนองต่อข้อสอบของบุคคลจะถูกกำหนดโดยลักษณะหรือความสามารถที่อยู่ภายใน ซึ่งไม่สามารถสังเกตได้

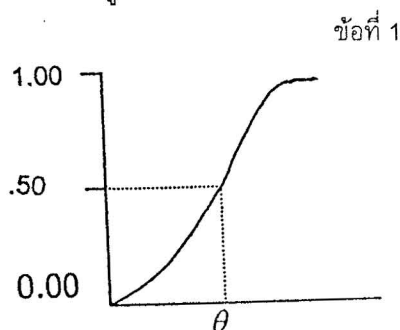
แฮมเบิลตัน และสวามินาธาน (Hambleton and Swaminathan. 1985 : 9) เชื่อว่าทฤษฎีการตอบข้อคำถามเป็นทฤษฎีที่ใช้ทำนาย หรืออธิบายคุณลักษณะหรือความสามารถของผู้สอบที่ไม่สามารถวัดได้โดยตรง ซึ่งใช้โมเดลทางคณิตศาสตร์แสดงความสัมพันธ์ระหว่างค่าที่สังเกตได้จากการสอบ อ้างอิงไปสู่ความสามารถที่แท้จริงของผู้สอบ

สำเร็จ บุญเรืองรัตน์ (2525: 2) กล่าวถึงใจความสำคัญของทฤษฎีการตอบข้อคำถาม คือ ผลการสอบของผู้ตอบจากแบบทดสอบใดๆ นั้นขึ้นอยู่กับความสามารถของผู้สอบหรือคะแนนของผู้สอบจากแบบทดสอบใดๆ สามารถพยากรณ์หรืออธิบายได้จากความสามารถของคนที่นั้นผู้จริงจิต อินทสุวรรณ (2525 : 55) กล่าวว่า ทฤษฎีคุณลักษณะภายใน (Latent Trait Theory) เป็นข้อความทางคณิตศาสตร์ที่กล่าวถึงการตอบข้อสอบของผู้สอบ ซึ่งขึ้นอยู่กับระดับความสามารถของผู้สอบเอง ความสัมพันธ์นี้กำหนดได้ด้วยฟังก์ชันทางคณิตศาสตร์ ที่เรียกว่าฟังก์ชันของลักษณะข้อคำถาม (Item Characteristic Function) หรือเรียกว่าโค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC)/สงบ ลักษณะ (2525 : 49) ได้กล่าวถึงทฤษฎีการตอบข้อคำถามว่ามีความสำคัญคือมุ่งหาความสัมพันธ์ระหว่างความสามารถที่แท้จริงกับการตอบข้อสอบ ซึ่งความสามารถที่แท้จริงได้แก่ คุณลักษณะที่วัดได้ด้วยแบบทดสอบ เช่น ความรู้ ความเข้าใจในด้านต่างๆ ส่วนการตอบข้อสอบหมายถึงข้อสอบที่ให้คะแนนแบบตอบถูกให้ 1 คะแนน ตอบผิดให้ 0 คะแนน นอกจากนี้ สงบ ลักษณะ ยังกล่าวว่าถ้าข้อตกลงเบื้องต้นของการสอบตรงตามที่ทฤษฎีต้องการและเราสามารถคำนวณค่าพารามิเตอร์ได้เหมาะกับโมเดลแล้วจะมีผลดีคือ

1. ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนก ค่าความยาก และค่าการเดา จะเป็นค่าที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบไม่ว่าจะนำไปสอบกับผู้ใด

2. เมื่อทราบลักษณะการตอบข้อสอบแต่ละข้อของผู้สอบคนใดเราสามารถคำนวณหาความสามารถที่แท้จริงของบุคคลนั้นได้ ค่าความสามารถที่แท้จริงนี้มีความสัมพันธ์โดยตรงกับคะแนนจริง เรียก ลักษณะเช่นนี้ว่าความเป็นอิสระของข้อสอบ

ทฤษฎีการตอบข้อคำถามเป็นการวิเคราะห์ข้อสอบ โดยอาศัยข้อมูลจากตัวข้อสอบเอง ข้อสอบที่นำมาวิเคราะห์ในทฤษฎีการตอบข้อคำถาม ต้องเป็นข้อสอบที่มีให้ทุกคนตอบผิดหมดหรือทุกคนตอบถูกหมด ข้อสอบทั้งสองแบบเป็นข้อสอบที่ไม่ชี้บ่งระดับโอกาสของการตอบถูกของบุคคลที่มีความสามารถแตกต่างกันอย่างใด การวิเคราะห์ข้อทดสอบที่ใช้ค่าความสามารถจริงของบุคคลนี้อาศัยข้อมูลที่เรียกว่า ลักษณะการตอบรายข้อเป็นสำคัญ โคง์ลักษณะการตอบรายข้อนี้คือภาพที่แสดงความสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับโอกาสที่จะตอบถูก



ภาพประกอบ 1 แสดงความสัมพันธ์ระหว่างความสามารถของผู้สอบกับโอกาสที่จะตอบถูก

จากภาพกำหนดให้แกนตั้งแทนค่าโอกาสตอบถูกและแกนนอนแทนความสามารถของผู้ตอบ ถ้าข้อสอบเป็นข้อที่มีคุณลักษณะที่ดี โอกาสของการตอบถูกของผู้ที่มีความสามารถต่ำกว่าย่อมมีค่าน้อยกว่าเสมอ บนโค้งนี้ถ้าลากเส้นจากแกนตั้งที่จุดค่า P เท่ากับ 0.50 ออกไปตัดโค้งลักษณะการตอบรายข้อ แล้วจากจุดตัดลากตั้งฉากมาตัดแกนนอน ค่าที่ได้ก็จะเป็นค่าอย่างง่ายของข้อทดสอบ จากโค้งรายข้อนี้สามารถบ่งบอกถึงค่าพารามิเตอร์รายข้อ ได้แก่ ค่าดังนี้

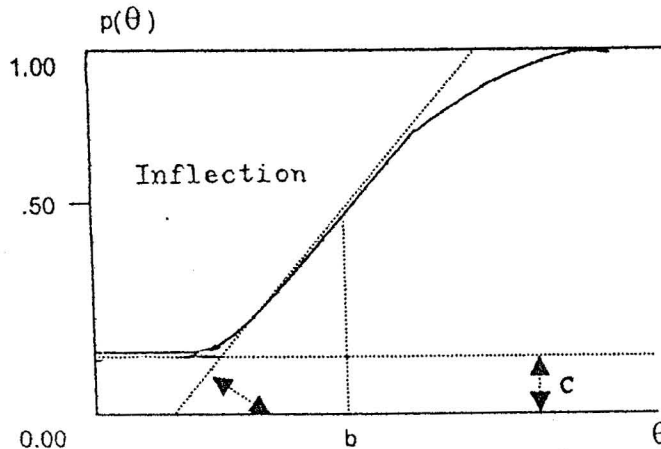
1. พารามิเตอร์ c เป็นความน่าจะเป็นที่บุคคลหนึ่งซึ่งปราศจากความสามารถโดยสิ้นเชิงจะตอบข้อนั้นได้ถูก พารามิเตอร์ตัวนี้มีชื่อเรียกว่า พารามิเตอร์ของการเดา ข้อใดที่ไม่สามารถตอบถูกได้ด้วยการเดาข้อนั้นมีพารามิเตอร์ของการเดาเป็นศูนย์ ($c=0$)

2. พารามิเตอร์ b เป็นพารามิเตอร์บอกตำแหน่งของโค้งตามมาตราความสามารถ พารามิเตอร์นี้มีชื่อเรียกว่า ความยากของข้อสอบ (Item Difficulty) ข้อคำถามยิ่งยากโค้งยิ่งเลื่อนไปทางขวามากขึ้น โคง์โลจิสติกมีจุดเปลี่ยนโค้งอยู่ที่ $\theta = b$ เมื่อไม่มีการเดา b จะเป็นระดับความสามารถของบุคคลที่มีความน่าจะเป็นของการตอบข้อนั้นถูกเป็น .5

3. พารามิเตอร์ a เป็นสัดส่วนกับความชันของโค้ง (ความชันมีค่าเป็น $.425 a (1-c)$) ดังนั้น a จึงแทนอำนาจจำแนก (Discriminating Power) ซึ่งจะบอกว่าฟังก์ชันการตอบข้อคำถามแปรไปตามระดับความสามารถมากน้อยเพียงใด

4. ค่าความสะเพร่า (γ) เป็นค่าที่แสดงถึงความสะเพร่าในการตอบข้อสอบข้อหนึ่ง ซึ่งมีค่าน้อยกว่า 1

5. ระดับความสามารถของผู้สอบ (θ) มีค่าอยู่ระหว่าง -3.0 ถึง 3.0 ค่าที่เป็นลบแสดงว่ามีความสามารถต่ำ และค่าที่เป็นบวกแสดงว่ามีความสามารถสูง



ภาพประกอบ 2 ความหมายของค่าพารามิเตอร์ของข้อสอบรูปแบบลอจิสต์ แบบ 3 พารามิเตอร์

ทฤษฎีการตอบข้อคำถามมีความเชื่อว่า ค่าพารามิเตอร์ต่างๆของข้อสอบไม่ว่าจะเป็นค่าความยาก ค่าอำนาจจำแนก หรือค่าการเดาของข้อสอบแต่ละข้อเป็นคุณลักษณะที่มีอยู่ประจำ และคงที่พอสมควรในตัวข้อสอบนั้นจริง แต่เนื่องจากความสามารถของผู้สอบเป็นลักษณะแฝง ไม่สามารถที่จะวัดหรือสังเกตได้โดยตรง แต่จะเป็นตัวพยากรณ์ผลการสอบหรือคะแนน เป็นสิ่งที่เราสามารถสังเกตหรือวัดได้ นักวัดผลจึงได้พยายามหาความสัมพันธ์ระหว่างปริมาณการกระทำข้อสอบหรือคะแนนกับปริมาณความสามารถ โดยสามารถเขียนในรูปความสัมพันธ์ได้ดังนี้

$$P = f(\mu_i / \theta_1, \theta_2, \theta_3, \dots, \theta_k; \beta_j)$$

เมื่อ P แทน ผลการสอบ

f แทน ฟังก์ชัน

μ_i แทน ผลการตอบข้อสอบ ตอบถูกให้ 1 ตอบผิดให้ 0

$\theta_1, \theta_2, \theta_3, \dots, \theta_k$ แทน ความสามารถที่ 1, 2, 3, ..., k

β_j แทน ค่าพารามิเตอร์ของข้อสอบข้อที่ j

ความสามารถดังกล่าวยังต้องอาศัยข้อตกลงเบื้องต้นต่างๆ ดังต่อไปนี้

1.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อสอบ

ลอร์ด และ โนวิค (Lord and Novick, 1968 : 359) กล่าวว่าข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถามประกอบด้วยลักษณะดังนี้

1. ความเป็นมิติเดียว (Unidimension) หมายความว่า ความสามารถของคนมีอยู่ทั้งหมด k อย่าง ซึ่งความสามารถแต่ละอย่างต่างก็ส่งผลต่อการตอบข้อสอบต่างๆ ที่รวมกันเป็นแบบทดสอบ ถ้าผลการตอบข้อสอบหรือคะแนนของผู้สอบนั้น สามารถอธิบายได้เพียงความสามารถเดียวก็เพียงพอแล้ว เรียกเงื่อนไขข้อนี้ว่ามีความเป็นมิติเดียว ในการปฏิบัติจริงนั้นสามารถกระทำโดยการวิเคราะห์องค์ประกอบ แล้วสังเกตค่า Eigen Value ค่าสูงสุดว่าแตกต่างจากค่าอื่นอย่างชัดเจนหรือไม่

2. ความเป็นอิสระในการตอบข้อสอบ หมายถึง การตอบข้อสอบใดข้อหนึ่งจะไม่มีผลต่อการตอบข้อสอบนั้น ถ้าข้อสอบชุดนั้นมีความเป็นมิติเดียวแล้วความเป็นอิสระในการตอบข้อสอบก็จะมีเอง ซึ่งสามารถตรวจสอบได้จากการคำนวณค่าไคสแควร์ เพื่อทดสอบความเป็นอิสระของการตอบคำถาม

ถ้าให้ (1) U_g แทน ผลการตอบคำถามข้อที่ g ของผู้สอบคนหนึ่ง

$$g = 1, 2, 3, \dots, n$$

โดยถือหลักว่าถ้าตอบถูกให้ 1 และตอบผิดให้ 0

(2) P_g เป็นโอกาสที่ผู้สอบตอบคำถามข้อ g ถูกต้อง

$$(3) Q_g = 1 - P_g$$

ข้อตกลงเกี่ยวกับ Local Independence กล่าวว่า

$$\begin{aligned} \text{Prob} \{U_1 = u_1, U_2 = u_2, \dots, U_n = u_n\} \\ = \prod_{g=1}^n P_g^{u_g} Q_g^{1-u_g} \dots\dots\dots(2) \end{aligned}$$

สมการสองนั้นอ่านว่าโอกาสของการตอบ (Prob ซึ่งย่อมาจาก Probability) แบบทดสอบใดถูกต้องทั้งหมดทุกข้อคือผลคูณ (π) ของโอกาสของการตอบคำถามแต่ละข้อได้ถูกต้อง

จากข้อตกลงดังกล่าวนี้ ทำให้สามารถสร้างสมการแสดงความถี่ของคะแนนของกลุ่มผู้สอบที่มีความสามารถ θ ได้ดังนี้

$$f(x / \theta) = \sum_{g=1}^n \pi P_g^{u_g} Q_g^{1-u_g} \dots\dots\dots(3)$$

เมื่อ x แทน คะแนนของผู้สอบซึ่งมีค่าตั้งแต่ 0 ถึง n

3. โค้งแสดงลักษณะของข้อสอบ (Item Characteristic Curve :ICC)

การแจกแจงความถี่ของคะแนนของกลุ่มผู้สอบที่มีระดับความสามารถ θ นั้นหาได้จากสมการ

(3) ถ้าแบบทดสอบวัดความสามารถเดียวสมการที่ (3) เรียกว่าโค้งลักษณะแบบทดสอบ (Test Characteristic Curve) แต่ถ้าวัดความสามารถหลายอย่างเรียกว่า ฟังก์ชันลักษณะข้อสอบ (Test Characteristic Function)

สำหรับคำถามแต่ละข้อนั้นก็สามารถแสดงลักษณะของโค้งผลการตอบได้เช่นกัน เรียกว่า โค้งลักษณะข้อสอบ (Item Characteristic Curves) สำหรับคำถามที่ข้อถูกเป็น 1 และผิดเป็น 0 นั้น สมการแจกแจงความถี่สามารถแสดงได้ดังนี้

$$f_g(U_g / \theta) = P_g^{U_g} Q_g^{(1-U_g)} \dots\dots\dots(4)$$

$$\begin{aligned} \text{นั่นคือ } f_g(U_g / \theta) &= P_g \text{ ถ้า } U_g = 1 \\ &= Q_g \text{ ถ้า } U_g = 0 \end{aligned}$$

แสดงถึงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบข้อนั้นถูกต้องกับระดับความสามารถที่วัดได้ โดยใช้ชุดของข้อสอบ โดยที่โอกาสที่ผู้สอบตอบข้อสอบถูกต้องจะขึ้นอยู่กับโค้งแสดงลักษณะข้อสอบ ในแต่ละโมเดลที่ใช้ ด้วยเหตุนี้โอกาสของการตอบถูกข้อสอบข้อนั้นจะไม่ขึ้นอยู่กับความสามารถของกลุ่มประชากร

1.3 พิสัยของค่าพารามิเตอร์

ค่าพารามิเตอร์ในทฤษฎีการตอบข้อคำถามแบ่งออกเป็นพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก (b_g) ค่าอำนาจจำแนก (a_g) ค่าสัมประสิทธิ์การเดา (c_g) และค่าพารามิเตอร์ของผู้สอบ ได้แก่ ระดับความสามารถของผู้สอบ (θ) พิสัยของพารามิเตอร์ต่างๆ เหล่านี้มีดังนี้ (Hambleton and Cook, 1977: 81)

b_g คือ ค่าที่แสดงระดับความสามารถที่จุดเปลี่ยนโค้งของโค้งลักษณะข้อสอบในกรณีที่ไม่มีการเดา ค่าความยากก็คือค่า ณ จุด $P_g = 0.50$ ดังนั้นค่าความยากจึง มีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติ นั้นจะมีค่าอยู่ระหว่าง -3 ถึง +3 ค่า -3 แสดงว่า ข้อสอบข้อนั้นง่ายมาก และค่า +3 แสดงว่าข้อสอบนั้นยากมาก

a_g คือ ค่าที่เป็นสัดส่วนโดยตรงกับค่าความชัน (Slope) ของโค้งลักษณะข้อสอบ ณ จุดเปลี่ยนโค้ง หรือบนจุดโค้งที่ $\theta = b_g$ ในทางทฤษฎีค่า a_g มีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติมีค่าตั้งแต่ 0 ถึง +2 เพราะค่า a_g ที่เป็นลบแสดงว่า ข้อสอบไม่ดีใช้ไม่ได้ จึงตัดข้อสอบนั้นทิ้งไป ค่า 0 แสดงว่าข้อสอบไม่มีอำนาจจำแนกหรือมีค่าอำนาจจำแนกต่ำ ค่า +2 แสดงว่าข้อสอบมีค่าอำนาจจำแนกสูง

c_g คือ ค่าความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำมาก มีโอกาสทำข้อสอบข้อ g ถูก

θ คือ ระดับความสามารถของผู้สอบที่คำนวณจากคะแนนจริง มีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติมีค่าอยู่ระหว่าง -3 ถึง +3 ค่า -3 แสดงว่ามีความสามารถต่ำ +3 แสดงว่ามีความสามารถสูง

1.4 โมเดลของทฤษฎีการตอบข้อคำถาม

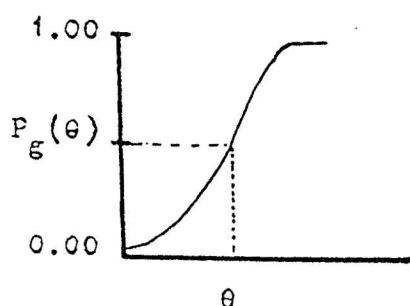
โมเดลของทฤษฎีการตอบข้อคำถามเป็นโมเดลทางคณิตศาสตร์ที่กล่าวถึงกันมาก มีอยู่ 2 โมเดล คือ (Warm. 1978 : 23-25 ; Hambleton and Cook. 1977 : 81-82)

1. โมเดลโค้งปกติสะสม (Normal Ogive Model) มี 3 รูปแบบคือ

1.1 แบบ 1-พารามิเตอร์ (One-Parameter) เป็นรูปแบบที่มีค่าพารามิเตอร์ตัวเดียว คือค่า b ที่แสดงลักษณะของข้อสอบ เมื่อ θ คือความสามารถของผู้สอบ แสดงได้ด้วยฟังก์ชันดังนี้

$$p_g(\theta) = \int_{-\infty}^{\theta - b_g} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

และแสดงโค้งลักษณะข้อสอบดังนี้

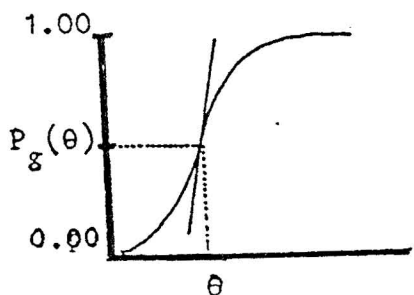


ภาพประกอบ 3 แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์ตัวเดียว

1.2 แบบ 2-พารามิเตอร์ เป็นรูปแบบที่ประกอบด้วย 2 พารามิเตอร์คือ a และ b ที่แสดงลักษณะข้อสอบ เมื่อ θ คือความสามารถของผู้สอบ มีฟังก์ชันดังนี้

$$p_g(\theta) = \int_{-\infty}^{a(\theta - b_g)} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

และแสดงโค้งลักษณะข้อสอบดังนี้

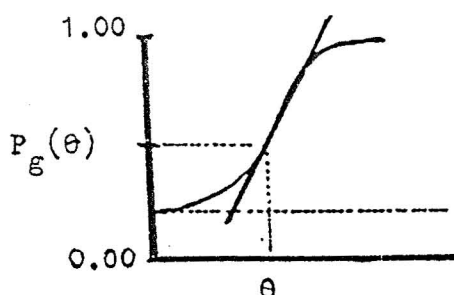


ภาพประกอบ 4 แสดงโมเดลโค้งปกติสะสมที่มีพารามิเตอร์สองตัว

1.3 แบบ 3-พารามิเตอร์ เป็นแบบที่มีค่าพารามิเตอร์ลักษณะข้อสอบ 3 ค่า คือ a , b และ c เมื่อ θ คือความสามารถของผู้สอบ มีฟังก์ชันดังนี้

$$p_g(\theta) = c_g + (1 - c_g) \int_{-\infty}^{a_g(\theta - b_g)} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

และแสดงโค้งลักษณะข้อสอบดังนี้



ภาพประกอบ 5 แสดงโมเดลโค้งปกติที่มีพารามิเตอร์สามตัว

เมื่อ a แทน ค่าอำนาจจำแนกหรือสัดส่วนของความชัน ณ จุดเปลี่ยนโค้ง

b แทน ค่าความยากของข้อสอบหรือตำแหน่งของข้อสอบบนเส้นความสามารถ

c แทน ค่าโอกาสการเดาถูก

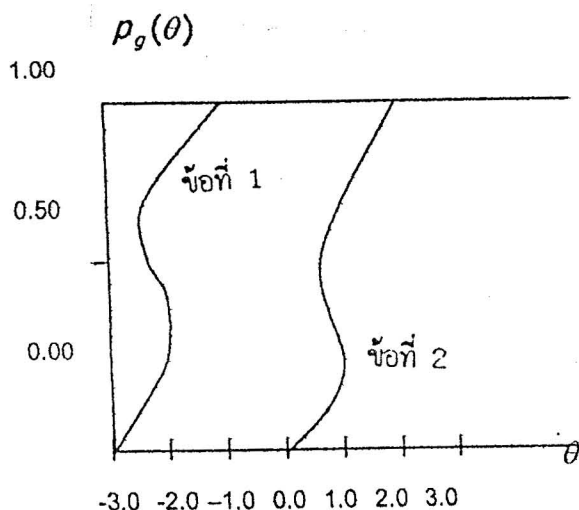
2. โมเดลโลจิสติก เป็นโมเดลที่พัฒนาเพื่อพัฒนาให้เหมาะสมกับสถานการณ์การทดสอบในกรณี ที่ตอบโดยการเดาได้ ซึ่งจะทำให้ส่วนปลายที่สุดของโค้งของ $P_i(\theta)$ ดังนั้นพารามิเตอร์ของการเดาจึงจำเป็นต้องพิจารณาด้วย นอกจากนี้หากผู้สอบทำข้อสอบด้วยความสะเพร่า โมเดลนี้จะเพิ่มความรอบคอบขึ้นอีกค่าหนึ่ง ซึ่งจะเห็นได้จากปลายสูงสุดของโค้ง ICC มีค่าน้อยกว่า 1 โมเดลโค้งปกติสะสมกับโมเดลโลจิสติก จึงมีลักษณะของโค้งใกล้เคียงกัน แต่โมเดลโลจิสติกมีความง่ายในการประมาณค่ามากกว่า จึงทำให้แพร่หลายมากกว่ารูปแบบโมเดลโลจิสติกมี 4 รูปแบบ

2.1 แบบ 1-พารามิเตอร์ หรือรูปแบบราสซัสโมเดล ราสซัสได้พัฒนาโมเดลนี้ โดยถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันหมด และไม่มีการเดาได้ถูกต้อง จึงมีค่าพารามิเตอร์แสดงลักษณะข้อสอบเพียงตัวเดียว คือค่าความยากเขียนเป็นฟังก์ชันได้ดังนี้

$$P_g(\theta) = \frac{e^{Da(\theta - b_g)}}{1 + e^{Da(\theta - b_g)}}; g = 1, 2, 3, \dots, n$$

หรือ

$$= \frac{e^{(\theta - b_g)}}{1 - e^{(\theta - b_g)}} = \frac{\text{EXP}(\theta - b_g)}{1 + \text{EXP}(\theta - b_g)}$$



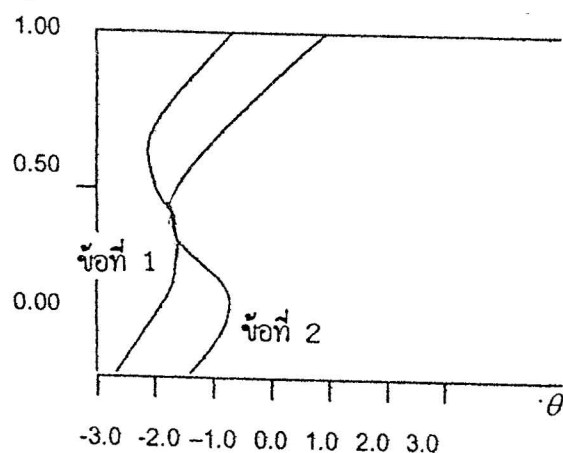
ภาพประกอบ 6 แสดงความหมายของค่า Item Parameter ของ One-Parameter Logistic Model

2.2 แบบ 2-พารามิเตอร์ (Two-Parameter) หรือรูปแบบที่เบิร์นบอม (Birnbbaum) ได้พัฒนาขึ้น ในปี 1968 มีค่าพารามิเตอร์แสดงลักษณะข้อสอบ 2 ค่า คือค่าอำนาจจำแนก และค่าความยาก แสดงได้โดยฟังก์ชันนี้

$$P_g(\theta) = \frac{e^{Da_g(\theta - b_g)}}{1 + e^{Da_g(\theta - b_g)}}; g = 1, 2, 3, \dots, n$$

หรือ

$$p_g(\theta) = \frac{\text{EXP}(Da_g(\theta - b_g))}{1 + \text{EXP}(Da_g(\theta - b_g))}$$



ภาพประกอบ 7 แสดงความหมายของค่า Item Parameter ของ Two-Parameter Logistic Model

2.3 แบบ 3-พารามิเตอร์ (Three-Parameter) หรือรูปแบบที่เบิร์ตโนมได้พัฒนาโดยมีค่าพารามิเตอร์แสดงคุณลักษณะข้อสอบ 3 ค่า โดยเพิ่มค่าการเดา (C) ใน 2-พารามิเตอร์ เขียนเป็นฟังก์ชันได้ดังนี้

$$P_g(\theta) = c_g + (1 - c_g) \frac{e^{Da_g(\theta - b_g)}}{1 + e^{Da_g(\theta - b_g)}}; g = 1, 2, 3, \dots, n$$

หรือ

$$= c_g + (1 - c_g) \frac{\text{EXP}(Da_g(\theta - b_g))}{1 + \text{EXP}(Da_g(\theta - b_g))}$$

เมื่อ $P_g(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบคนหนึ่ง มีระดับความสามารถจะตอบข้อสอบข้อ g ได้ถูกต้อง

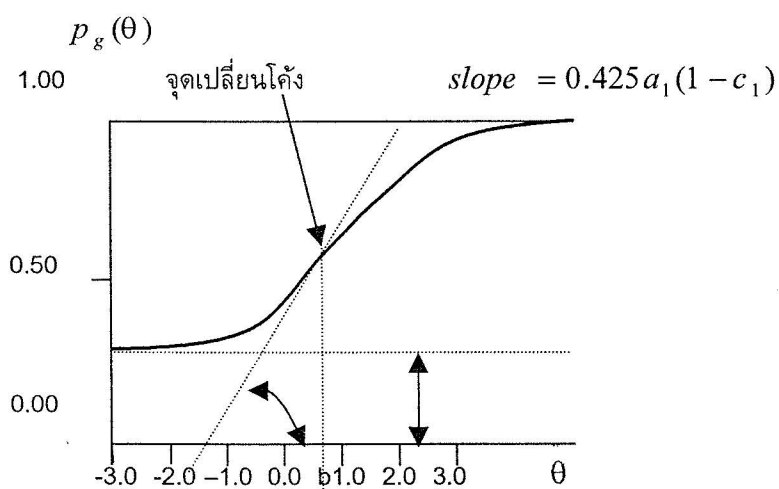
c_g แทน ค่าความน่าจะเป็นของผู้สอบที่มีความสามารถต่ำมาก มีโอกาสจะทำข้อสอบข้อนี้ได้ถูกต้อง หรือค่าการเดา

b_g แทน ค่าความยากของข้อสอบข้อที่ g ที่แสดงระดับความสามารถที่แท้จริงที่จุดโค้งชันที่สุด คือ θ ณ จุดความน่าจะเป็น $(1+C)/2$

a_g แทน ค่าอำนาจจำแนกข้อสอบข้อ g ที่เป็นสัดส่วนกับความชันของ $P_g(\theta)$ ณ ตำแหน่ง $\theta = b_g$

D แทน ค่าคงที่ของ Scaling factor มีค่าเท่ากับ 1.702

e แทน ค่าคงที่มีค่าเท่ากับ 2.7182818



ภาพประกอบ 8 แสดงความหมายของค่า Item Parameter ของ Three-Parameter Logistic Model

2.4 แบบ 4-พารามิเตอร์ โมเดลนี้เสนอโดยแม็คโดแนลด์ในปี ค.ศ.1967 ลอร์ด และ บาร์ตัน ได้พัฒนาในปี ค.ศ.1981 โดยเพิ่มพารามิเตอร์ค่าความสะเพร่า (Careless) ในโมเดลที่ใช้พารามิเตอร์ 3 ตัว เป็นค่าที่แสดงให้เห็นถึงกรณีที่ผู้สอบมีความสามารถสูงมักตอบข้อสอบผิด อาจเป็นเพราะความไม่รอบคอบทำให้ผู้นั้นไม่เลือกคำตอบที่ถูก จะเห็นได้ว่าปลายสูงสุดของโค้ง ICC ต่ำกว่า 1 เขียนเป็นสมการได้ดังนี้ (Hambleton and Swaminathan.1985 : 48)

$$P_g(\theta) = c_g + (\gamma_g - c_g) \frac{e^{D_{a_g}(\theta - b_g)}}{1 + e^{D_{a_g}(\theta - b_g)}}; g = 1, 2, 3, \dots, n$$

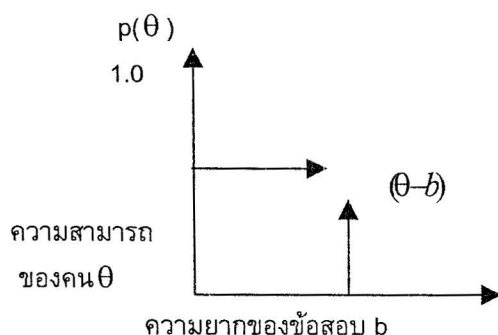
เมื่อ γ_g แทน ค่าความรอบคอบของข้อสอบข้อ g ซึ่งมีค่าน้อยกว่า 1

2. ราล์ชโมเดล

ราล์ชโมเดล เป็นโมเดลหนึ่งในโมเดลของทฤษฎีการตอบข้อคำถามที่ใช้พารามิเตอร์ตัวเดียว กล่าวคือ พิจารณาค่าความยากตัวเดียว โดยถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันหมดและไม่มีการเดาได้ถูกต้อง

ในการสรุปโมเดลเกี่ยวกับราล์ชโมเดล อาจกล่าวได้ว่า เมื่อคนหนึ่งๆ ทำข้อสอบข้อหนึ่ง เหตุการณ์ขณะนั้นก็คือ คนพยายามใช้ความสามารถที่จะเอาชนะข้อสอบด้วยการแก้ปัญหา โจทย์ปัญหาข้อนั้น ถ้าคนนั้นมีความสามารถมากกว่าความยากของข้อสอบข้อนั้น โอกาสที่คนจะทำข้อสอบถูกก็มีน้อยไปด้วย หรืออาจพูดอีกอย่างว่า โอกาสทำผิดก็มีมาก และถ้าความสามารถของคนมีพอ ๆ กันกับความยากของข้อสอบ โอกาสของการทำข้อถูกก็จะมีอยู่ครึ่งต่อครึ่ง คืออาจจะทำถูกก็ได้หรืออาจจะทำผิดก็ได้เช่นเดียวกัน

เหตุการณ์ที่บุคคลใช้ความสามารถแก้ปัญหา โจทย์ข้อหนึ่งๆ อาจเขียนเป็นรูปได้ดังนี้



ภาพประกอบ 9 การเปรียบเทียบความสามารถของคนกับความยากของข้อสอบ

เมื่อกำหนดให้ θ คือความสามารถของคน และ b คือ ความยากของข้อสอบ การเปรียบเทียบความสามารถกับความยากของข้อสอบผลลัพธ์ อาจเขียนในเชิงคณิตศาสตร์ได้ดังนี้

ถ้า $\theta > b$ และ $(\theta - b) > 0, p > .5$

อธิบายได้ว่า เมื่อความสามารถของคนมีมากกว่าความยากของข้อสอบ คือ b_g และถ้านำเอาความสามารถของคนลบด้วยความยากของข้อสอบแล้ว ผลลัพธ์จะได้มากกว่า 0 โอกาสที่คน ๆ นั้นจะทำข้อสอบข้อนั้นถูกก็จะมีค่ามากกว่า 0.50

ถ้า $\theta < b$ และ $(\theta - b) < 0, p < .5$

อธิบายได้ว่า เมื่อความสามารถของคนลบด้วยความยากของข้อสอบได้ผลลัพธ์น้อยกว่าศูนย์ โอกาสที่คน ๆ นั้นจะทำข้อสอบข้อนั้นได้ถูกต้อง ก็จะมีค่าน้อยกว่า 0.50

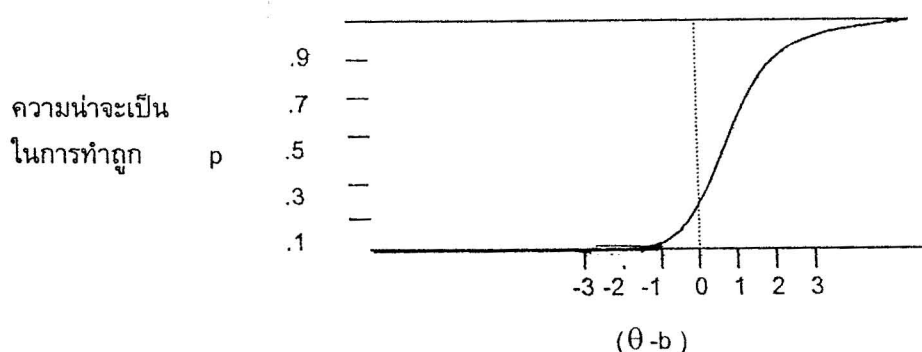
ถ้า $\theta = b$ และ $(\theta - b) = 0, p = .5$

อธิบายได้ว่า เมื่อความสามารถของคนลบด้วยความยากของข้อสอบได้เท่ากับศูนย์พอดีแล้ว โอกาสที่คน ๆ นั้นจะทำข้อสอบข้อนั้นได้ก็จะมีค่าเท่ากับ .50 พอดี ซึ่งอาจจะทำถูกหรือผิดก็ได้

จะเห็นได้ว่า เมื่อคน ๆ หนึ่งทำข้อสอบข้อหนึ่ง บุคคลนั้นต้องใช้ความสามารถต่อสู้กับความยากของข้อสอบ ผลของการทำข้อสอบจึงเท่ากับความแตกต่างระหว่างความสามารถของคนกับความยากของข้อสอบ ความแตกต่างระหว่าง θ กับ b หรือ $(\theta - b)$ นั้นเอง แต่ค่า $\theta - b$ นี้อาจจะให้ผลลัพธ์ได้ตั้งแต่ $-\infty$ จนถึง $+\infty$ จึงต้องคิดแปลงให้เป็นค่าความน่าจะเป็นซึ่งจะมีค่าอยู่ระหว่าง .00 ถึง .99 หรือจาก 0 ถึง 1 เท่านั้น และเพื่อให้ค่าเปลี่ยนแปลงแต่น้อย จึงใช้ความน่าจะเป็นค่าอัตราส่วนของลอการิทึม (Natural log) ของค่าคงที่ตัวหนึ่งคือ ค่า $e^{(\theta - b)}$ และค่าอัตราส่วนนั้นก็คือ

$$P(\theta) = \frac{e^{(\theta - b)}}{1 + e^{(\theta - b)}}$$

ซึ่งจะทำให้ได้เส้นความถี่สะสมปกติหรือ Normal Ogive ที่รัสซเรียกว่า Logistic ที่อาจแสดงได้ดังต่อไปนี้



ภาพประกอบ 10 แสดงความแตกต่างระหว่างความสามารถของคนกับความยากของข้อสอบ

3. วิธีประมาณค่าพารามิเตอร์

จากที่กล่าวมาจะเห็นว่า สามารถแยกคะแนนของคนออกจากคะแนนของข้อสอบ ทำให้สามารถวิเคราะห์ค่าความสามารถของคนโดยไม่ขึ้นกับข้อสอบได้ในทำนองเดียวกันก็อาจวิเคราะห์ค่าของข้อคำถามโดยไม่ต้องขึ้นอยู่กับผู้สอบได้ เป็นลักษณะพิเศษของการวิเคราะห์แบบราสช์โมเดล ที่ทำให้สามารถหาค่าวิเคราะห์ของคน และข้อสอบได้เป็นอิสระจากกัน

ในการวิจัยครั้งนี้ผู้วิจัยเลือกใช้โลจิสติกโมเดลแบบ 1-พารามิเตอร์ หรือราสช์โมเดลที่มีจุดเด่น 3 ประการคือ

1. โมเดลนี้มีจำนวนพารามิเตอร์ในข้อคำถามน้อยกว่าโมเดลอื่นๆ จึงทำให้ง่ายต่อการนำมาใช้งาน
2. มีปัญหาอันเกิดจากการประมาณค่าพารามิเตอร์ของข้อคำถามน้อยกว่าใช้โมเดลอื่นๆ
3. คุณสมบัติในการประมาณค่าพารามิเตอร์ของข้อคำถาม และค่าพารามิเตอร์ความสามารถของผู้สอบ แยกเป็นอิสระจากกันอย่างสมบูรณ์

ในทฤษฎีการตอบข้อคำถามเมื่อมีการนำไปประยุกต์ใช้ นอกจากการที่จะเลือกโมเดลต่างๆ ให้เหมาะสมแล้ว วิธีประมาณค่าความสามารถของผู้สอบก็เป็นอีกประเด็นหนึ่งที่เป็นขั้นตอนสำคัญที่ต้องเลือกใช้ให้เหมาะสมกับสภาพแต่ละครั้ง เพื่อผลการประมาณค่าความสามารถของผู้สอบจะได้มีประสิทธิภาพสูงสุด วิธีการประมาณค่าในทฤษฎีการตอบข้อคำถามที่มีการให้คะแนนเป็น 2 ค่า คือ ถูกให้ 1 และผิดให้ 0 มีอยู่หลายวิธี พอสรุปได้ดังนี้ (ต่าย เชียงฉี. 2534 : 49-50 ; อ้างอิงมาจาก Hambleton and Swaminath. 1985 : 300)

1. วิธีประมาณค่าความสามารถของผู้สอบตามค่าความยากของข้อที่ยากที่สุดที่ทำถูกเป็นการให้คะแนน โดยดูจากข้อสอบการทดสอบที่ผู้สอบทำมาทั้งหมด ข้อใดที่มีค่าความยากสูงสุดที่ผู้สอบทำถูกถือว่าเป็นค่าความสามารถของผู้สอบคนนั้น ค่าความยากที่ว่านี้ควรจะเป็นค่าความยากที่วิเคราะห์มาจากทฤษฎีการตอบข้อคำถาม
2. วิธีประมาณค่าความสามารถของผู้สอบตามค่าเฉลี่ยความยากของข้อที่ทำถูก การประมาณค่าตามค่าความยากเฉลี่ยของข้อสอบข้อที่ตอบถูก เป็นวิธีที่ช่วยขจัดข้อบกพร่องของการประมาณค่า ตามค่าความยากเฉลี่ย โดยนำเอาเฉพาะข้อสอบที่ผู้สอบตอบถูกเท่านั้นมาหาค่าเฉลี่ย ซึ่งจะให้ความหมายของคะแนนเพื่อวัดระดับความสามารถของผู้สอบได้มากขึ้น คะแนนที่ได้โดยการประมาณค่าวิธีนี้จะมากพอ ๆ กับการประมาณค่าตามผลการตอบข้อสอบข้อสุดท้ายหรือพอ ๆ กับการประมาณค่าตามค่าความยากของข้อสอบที่มากที่สุด แต่คะแนนจะมีค่าน้อยกว่า เมื่อใช้วิธีประมาณค่าตามผลการตอบข้อสอบข้อสุดท้ายได้รวมเอาค่าความยากของข้อสอบที่ใช้ในการแยกทางไว้ด้วย วิธีการประมาณค่าตามค่าความยากเฉลี่ยของข้อสอบที่ตอบถูกจะมีความคง ที่ของคะแนน และให้รายละเอียดของคะแนนมากกว่าวิธีการประมาณค่าตามผลการตอบข้อสอบข้อสุดท้าย อย่างไรก็ตาม วิธีการประมาณค่าตามค่าความยากเฉลี่ยของข้อสอบที่ตอบถูกจะมีความไวต่อการเดาตอบของผู้สอบมากกว่าวิธีการประมาณค่าตามค่าความยากเฉลี่ยอาจจะมีผลทำให้มีค่าความเชื่อมั่นต่ำลง

3. วิธีประมาณค่าความสามารถโดยวิธีพริกซ์ เป็นวิธีที่แสดงให้เห็นถึงหลักการสำคัญต่าง ๆ ภายใต้การวิเคราะห์หาคุณลักษณะประจำข้อ และค่าคุณลักษณะของผู้ตอบจากราสช์โมเดล โดยค่าทั้งสองที่ประมาณได้จากวิธีการนี้ถูกกำหนดให้อยู่บนมาตราวัดเชิงเส้นเดียวกัน และเป็นอิสระจากกลุ่มตัวอย่าง

และเป็นความยาวของแบบทดสอบที่ใช้โดยที่วิธีการประมาณค่าความสามารถวิธีนี้มีข้อตกลงเบื้องต้นอยู่คือ ในแบบทดสอบต้องมีการกระจายในลักษณะเบ้หรือมีแนวโน้มเป็นโค้งปกติ

4. วิธีประมาณค่าความสามารถตามค่าเฉลี่ยความยากของทุกข้อที่ทำ (Average Difficulty) วิธีการประมาณค่าแบบนี้เสนอโดยลอร์ด (Weiss.1974:33;citing Lord. 1972) วิธีการคือ นำค่าความยากของข้อสอบทุกข้อที่ผู้สอบได้ทำ โดยไม่คำนึงว่าจะตอบถูกหรือไม่มาหาค่าเฉลี่ย ค่าเฉลี่ยนี้จะถือเป็นความสามารถของผู้สอบผู้นั้น

5. วิธีประมาณค่าความสามารถตามค่าความยากของข้อสุดท้ายที่ทำ (Difficulty of the Final Item) นั่นคือ การประมาณค่าความสามารถวิธีนี้มักนำไปใช้ในการทดสอบเทเลอร์ โดยใช้ข้อสุดท้าย ซึ่งมีค่าความยากสูงสุด โดยไม่คำนึงว่าจะตอบถูกหรือไม่ก็ตาม จะถือว่าค่าความยากของการทดสอบเทเลอร์ขั้นสุดท้ายที่ผู้สอบได้ทำ คือค่าความสามารถของผู้สอบผู้นั้น

6. วิธีประมาณค่าความสามารถตามค่าความยากของข้อสอบที่ต่อจากขั้นสุดท้ายที่ทำ (Difficulty of (N+1)th Item) ซึ่งมีลักษณะคล้ายกับวิธีที่ 5 แต่เป็นการตรวจสอบต่อไปอีกขั้นหนึ่ง กล่าวคือ ถ้าขั้นสุดท้ายที่ทำถูก คะแนนในขั้นต่อไปก็จะเพิ่มขึ้น และถ้าขั้นสุดท้ายที่ทำผิด ขั้นต่อไปคะแนนจะลดลง

ในการวิจัยครั้งนี้ ผู้วิจัยได้เลือกใช้วิธีการประมาณค่าความสามารถโดยวิธีพร็อกซ์ เนื่องจากเป็นวิธีที่มีความคงที่ในการประมาณค่ามาก ถึงแม้จะใช้ค่าสถิติที่ซับซ้อน แต่ค่าที่ประมาณค่าออกมาได้ผลมีประสิทธิภาพ เป็นวิธีประมาณค่าที่เริ่มจากการแปลงคะแนนรวม ของผู้สอบแต่ละคนและข้อคำถามในเมตริกซ์ข้อมูล ผลการตอบของกลุ่มตัวอย่างให้อยู่ในรูปคะแนนโลจิท ทั้งนี้เพื่อให้คะแนนของบุคคลและของข้อคำถามอยู่ในมาตราอันตรภาค เป็นหน่วยเดียวกันบนมาตราเชิงเส้น จากนั้นจึงคำนวณหาค่าเฉลี่ยคะแนนโลจิทของแต่ละรายการ เพื่อแยกอิทธิพลของผู้ตอบออกมา และปรับขยายค่าที่คำนวณได้อีกครั้งหนึ่งเพื่อลดผลที่เกิดจากการกระจายของคะแนนข้อคำถาม และของผู้ตอบ หลักการของการประมาณค่าพารามิเตอร์นี้เป็นของโคเฮน แล้วมีการนำมาปรับปรุงโดยไรท์และมาสเตอร์ สามารถที่จะวิเคราะห์ผลได้โดยการใช้โปรแกรม BICAL(Wright and Mead 1979)

4. คุณภาพของแบบทดสอบ

ตามทฤษฎีการตอบข้อคำถามการหาคุณภาพของแบบทดสอบโดยการพิจารณาจากฟังก์ชันอินฟอร์เมชันของแบบทดสอบ (Test Information Function) ซึ่งได้จากผลรวมของอินฟอร์เมชันของข้อสอบแต่ละข้อ (Item Information Function) ซึ่งจะเป็นตัวบอกความแน่นอนในการประมาณค่าความสามารถที่แท้จริง (LORD. 1980 :72)

ฟังก์ชันอินฟอร์เมชันของข้อสอบแต่ละข้อกำหนดดังนี้

$$I(\theta, U_g) = (P'_g)^2 / P_g Q_g; g = 1, 2, \dots, n$$

เมื่อ $I(\theta, U_g)$ แทน ฟังก์ชันอินฟอร์เมชันของข้อสอบแต่ละข้อ

P_g แทน ความน่าจะเป็นที่ผู้สอบที่มีความสามารถ θ จะตอบข้อสอบข้อที่ g ถูก

Q_g แทน $(1 - P_g)$

P'_g แทน ความชันของโค้งลักษณะข้อสอบที่ระดับความสามารถ θ ของข้อสอบ

ข้อที่ g หรือค่าอนุพันธ์ (Derivative) ของโอกาสในการตอบข้อที่ g ถูก

จากสมการ จะเห็นได้ว่าฟังก์ชันอินฟอร์เมชันของข้อสอบแต่ละข้อขึ้นอยู่กับความชันของโค้งลักษณะข้อสอบ ถ้าโค้งลักษณะข้อสอบชันมากขึ้นในขณะที่ความแปรปรวนของการตอบข้อสอบถูกยิ่งน้อยลง โค้งอินฟอร์เมชันของข้อสอบที่ระดับความสามารถนั้นก็ยิ่งสูงขึ้น ซึ่งความสูงของโค้งอินฟอร์เมชันของข้อสอบอยู่ที่ระดับความสามารถใดแสดงว่า สามารถจำแนกระดับความสามารถของผู้สอบได้ดี ณ ระดับความสามารถนั้น (ผจงจิต อินทสุวรรณ.2525 : 64)

ฟังก์ชันอินฟอร์เมชันของแบบทดสอบ

ฟังก์ชันอินฟอร์เมชันของแบบทดสอบ เป็นสัดส่วนกลับของกำลังสองของความยาวของช่วงความเชื่อมั่น เป็นผลมาจากการประมาณค่าความสามารถของผู้สอบ จากการทำข้อสอบของเข้า โค้งอินฟอร์เมชันของแบบทดสอบจึงเป็นเครื่องแสดงถึงความถูกต้องแน่นอนของค่าความสามารถที่ประมาณได้ ซึ่งแสดงในรูปสมการได้ดังนี้

$$I(\theta) = \sum_{g=1}^n I(\theta, U_g)$$

$$I(\theta) = \sum_{g=1}^n \frac{(P'_g)^2}{P_g Q_g}$$

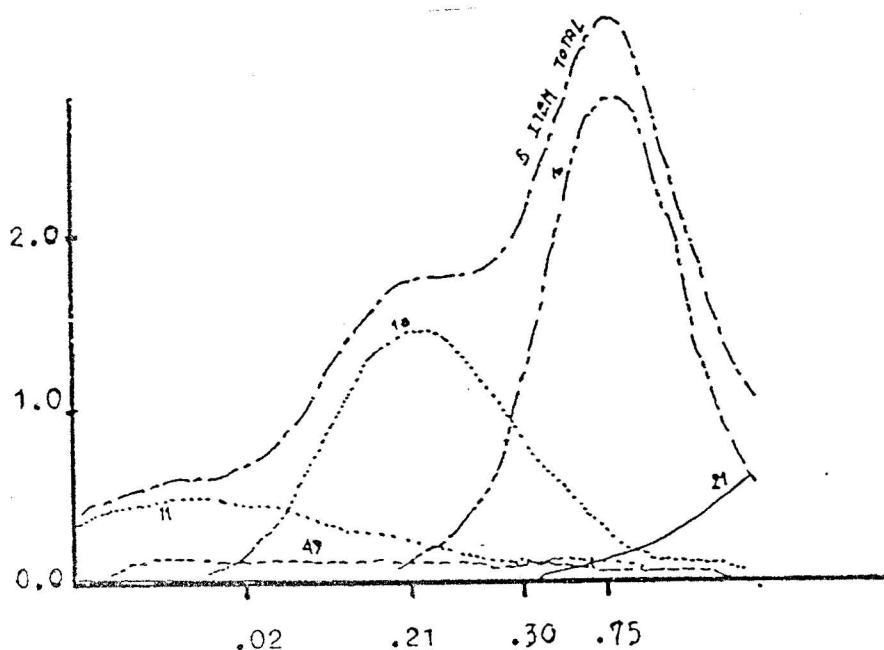
เมื่อ $I(\theta)$ คือ ฟังก์ชันอินฟอร์เมชันของแบบทดสอบ

ประโยชน์ที่ได้รับจากการทราบค่าอินฟอร์เมชันของข้อสอบแต่ละข้อ เราสามารถสร้างแบบทดสอบให้มีโค้งอินฟอร์เมชันของแบบทดสอบ ณ ระดับหนึ่งของความสามารถผู้สอบที่ต้องการได้ค่า I นี้ใช้เป็นดัชนีชี้ให้เป็นคุณภาพของการวัดของแบบทดสอบ ถ้าโค้งของข้อสอบ (Information Curve หรือ Test Information Curve) สูงที่ระดับความสามารถใดก็แสดงว่า แบบทดสอบหรือข้อสอบที่มีคุณภาพดีมากสำหรับวัดความสามารถของบุคคลที่มีระดับความสามารถนั้นๆ

ตาราง 1 แสดงคุณสมบัติของแบบทดสอบ 5 ข้อ

ข้อที่	b_g	a_g	c_g
10	1.1	2.0	.05
11	-1.5	.9	.20
13	0.1	1.6	.16
30	2.4	1.1	.09
47	-0.4	.4	.20

ค่าต่างๆ ในตารางนี้เป็นคุณสมบัติของข้อสอบความถนัดทางการเรียนด้านภาษา (Lord. 1968) 5 ข้อ เมื่อนำมาคำนวณค่า I ของแต่ละข้อ และคำนวณค่า I รวมทั้ง 5 ข้อแล้วก็สามารถเขียนกราฟได้ดังภาพ



ภาพประกอบ 11 แสดง Information Curve ของแบบทดสอบความถนัดด้านภาษาของสถาบัน ETS จากกราฟให้ความหมายว่า แบบทดสอบทั้ง 5 ข้อเหมาะกับการทดสอบกับนักเรียนที่มีความสามารถสูงๆ

5. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวกับการศึกษาเรื่องนี้ มีผู้ทำการวิจัยหลายท่านทั้งต่างประเทศและในประเทศ สำหรับงานวิจัยต่างประเทศมีผู้ทำการวิจัยดังนี้

ทินส์เลย์ และเดวิส (ไอลดา คล้ายสำริด. 2538 ; อ้างอิงจาก Tinsley and Dawis. 1975.

Educational and Psychological Measurement : p. 325-339) ได้ศึกษาความคงที่ของความยากของข้อสอบด้วยวิธีสโมเดล โดยใช้แบบทดสอบ 4 ฉบับ คือ Word Analogy, Number Analogy, Picture Analogy, Symbol Analogy มีจำนวนข้อ 66, 60, 50 และ 40 ข้อตามลำดับ นำไปทดสอบกับนักศึกษาในมหาวิทยาลัย และนักเรียนระดับมัธยมศึกษา วิเคราะห์ค่าความยากด้วยวิธีสโมเดล 10 ครั้ง แล้วนำค่าความยากของแต่ละครั้งมาหาความสัมพันธ์ ผลปรากฏว่า แบบทดสอบทั้ง 4 ฉบับ ที่มีจำนวนข้อมากกว่า 30 ข้อขึ้นไปมีความยากสัมพันธ์กันสูงมาก และถ้าตัดข้อที่ไม่เหมาะสมกับโมเดลออก จะทำให้ค่าความยากมีความสัมพันธ์กันสูงมากขึ้น

เรนซ์และบอชอว์ (สุพล นิลกลาง. 2541 ; อ้างอิงจาก Rentz and Bashaw. 1975. *Educational and Psychological Measurement*) ได้ทำการศึกษาและพบว่า ขนาดของกลุ่มตัวอย่าง 500-1,000 คน จะ

ทำให้การประมาณค่าพารามิเตอร์ของรูปแบบของราสช์มีความคงที่เพิ่มขึ้นอย่างรวดเร็ว และจะค่อย ๆ เพิ่มขึ้นเมื่อกลุ่มตัวอย่างที่ใช้ประมาณค่าพารามิเตอร์ มีจำนวน 2,000 คน และ 4,000 คน

สไลด์และลินด์ (บุญชิต รอดแก้ว.2533 ; อ้างอิงจาก Slide and Linn. 1979. *Applied Measurement in education* : p.159-165) ได้ศึกษาการเทียบคะแนนใช้ข้อสอบ 60 ข้อ แบ่งเป็น 2 ฉบับ ฉบับง่ายและฉบับยาก ฉบับละ 30 ข้อ ทดสอบกับกลุ่มตัวอย่าง 1,638 คน วิเคราะห์โดยใช้คอมพิวเตอร์ โปรแกรม BICAL แบ่งกลุ่มตัวอย่างโดยอาศัยคะแนนจากการทดสอบออกเป็น 3 กลุ่ม (กลุ่มสูง 506 คน กลุ่มกลาง 509 คน และกลุ่มต่ำ 565 คน) ซึ่งตัดกลุ่มที่สอบได้คะแนนศูนย์หรือได้คะแนนเต็มออก ผลการเปรียบเทียบแบบทดสอบฉบับยาก ฉบับง่ายกับกลุ่มตัวอย่างทั้ง 3 กลุ่ม พบว่าค่าเฉลี่ยของความสามารถของแต่ละกลุ่มในแบบทดสอบแต่ละฉบับมีค่าใกล้เคียงกัน คือค่าความสามารถเฉลี่ยของกลุ่มสูงที่สอบด้วยแบบทดสอบฉบับยาก 1.583 ฉบับง่าย 1.644 ค่าความสามารถเฉลี่ยของกลุ่มต่ำที่สอบด้วยแบบทดสอบฉบับยาก -.766 ฉบับง่าย -.782 เมื่อนำค่าความสามารถของกลุ่มสูงกับกลุ่มต่ำไปเทียบคะแนนโดยการเขียนกราฟ ให้แกนนอนเป็นความสามารถของกลุ่มสูง แกนตั้งเป็นความสามารถของกลุ่มต่ำ แล้วใช้เส้นตรงที่ผ่านมุม 45 องศา กับแกนนอนหรือเส้นตรงที่มีความชันเป็น 1 เป็นหลักในการพิจารณา จะได้ว่าแบบทดสอบฉบับง่ายมีแนวเป็นเส้นตรงอยู่เหนือเส้นหลัก แบบทดสอบฉบับยากก็มีแนวเป็นเส้นตรงอยู่ล่างเส้นหลัก และเส้นตรงทั้ง 3 เส้น เกือบจะขนานกันซึ่งเป็นการชี้ให้เห็นว่าความสามารถของแต่ละบุคคล ไม่ขึ้นอยู่กับกลุ่มผู้สอบ

ค็อค (Koch. 1983 : 15-32) ได้ศึกษาโดยใช้ Graded Response Latent Trait Model ในการวัดเจตคติโดยมุ่งที่จะวิจัยถึงความสามารถในการประยุกต์และการใช้ทฤษฎีคุณลักษณะภายในโดยใช้แบบทดสอบวัดเจตคติตามแนวทางลิเกิร์ตประมาณค่าพารามิเตอร์ของแบบทดสอบรายข้อและค่าความสามารถของผู้สอบ แล้วเปรียบเทียบการประมาณค่า 2 วิธี คือวิธีคุณลักษณะภายในกับวิธีแบบดั้งเดิม ผลการวิจัยพบว่า การประมาณค่าพารามิเตอร์และค่าความสามารถของทั้งสองวิธีมีความสัมพันธ์กันสูง นอกจากนี้พบว่า Graded Response Latent Trait Model มีความเหมาะสมกับแบบทดสอบวัดเจตคติ การใช้คุณลักษณะภายในจะมีผลการวัดในแต่ละระดับของเจตคติมีลักษณะต่อเนื่อง และยืนยันว่าค่าพารามิเตอร์และค่าความสามารถของผู้สอบมีคุณสมบัติไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง

ดักลาส (Douglass. 1981 : 4000-A) ได้ศึกษาเปรียบเทียบการใช้รูปแบบโครงสร้างข้อคำถามกับการสอนในระบบห้องเรียน โดยใช้รูปแบบทั้ง 3 รูปแบบของทฤษฎีการตอบข้อคำถาม ใช้ข้อสอบ 100 ข้อ จากกลุ่มตัวอย่างนักเรียนระดับวิทยาลัย จำนวน 1,082 คน สร้างข้อสอบ 4 ฉบับละ 43-53 ข้อ โดยสุ่มมาจากข้อสอบ 100 ข้อ แล้วสอบกับกลุ่มตัวอย่างที่สุ่มมา กลุ่มละ 200,600 และ 800 คนตามลำดับ ปรากฏว่าการวิเคราะห์ข้อสอบด้วยราสช์โมเดลมีความคงที่ของค่าพารามิเตอร์ดีกว่ารูปแบบ 2 และ 3 พารามิเตอร์ นอกจากนี้ ฮัทเทน (Hutten. 1982 : 4799-A) ได้ศึกษาความเหมาะสมของข้อมูลจริงกับรูปแบบของทฤษฎีการตอบข้อคำถามในการประมาณค่าพารามิเตอร์ความสามารถ และความยากโดยใช้กลุ่มตัวอย่าง 1,000 คน และ 250 คน ผลการศึกษาพบว่า รูปแบบของราสช์โมเดล และรูปแบบ 3 พารามิเตอร์ มีความเหมาะสมกับข้อมูล 80 เปอร์เซนต์ทั้งสองรูปแบบ ส่วนกลุ่มตัวอย่าง 250 คน มีความเหมาะสมสำหรับราสช์โมเดลมากกว่ารูปแบบ 3 พารามิเตอร์ ผลการศึกษานี้สนับสนุนให้ใช้รูปแบบ 3 พารามิเตอร์ กับกลุ่มตัวอย่างตั้งแต่ 1,000 คนขึ้นไป จะทำให้มีความแม่นยำในการประมาณค่าสูง

ราสช์ (Rasch. 1983 : 1429-A) ได้ศึกษาการใช้รูปแบบของราสช์ในแบบทดสอบหลายๆ คุณลักษณะ โดยแปลงให้เป็นคุณลักษณะเดียวกันหลายๆกลุ่ม วิธีการใช้ประมาณค่าใช้คะแนนจริงจากแบบ

ทดสอบจากการศึกษาพบว่าค่าพารามิเตอร์ที่ได้จากการวัดแต่ละกลุ่มมีการกระจายที่ไม่เปลี่ยนแปลงไปตามกลุ่มผู้สอบและแบบทดสอบ ส่วนแคเพสัน (Mcpheson, 1983) ได้ศึกษาการใช้รูปแบบของราล์ชในการประเมินข้อสอบโดยใช้แบบทดสอบ 4 ฉบับ สำหรับนักเรียนเกรด 4 และเกรด 7 ที่สอบในโปรแกรมการประเมินทางการศึกษาของมิชิแกน ซึ่งเป็นแบบทดสอบอิงเกณฑ์ในรูปแบบการอ่านที่ใช้ข้อสอบในปี ค.ศ.1976-1979 ผู้วิจัยต้องการทดสอบว่า รูปแบบของราล์ชจะเหมาะสมหรือไม่กับแบบทดสอบที่ใช้เวลามาก ในการที่จะเกิดผลสัมฤทธิ์ ผลปรากฏว่าแบบทดสอบทั้ง 4 ฉบับ เหมาะสมกับรูปแบบของราล์ชเพียง 1 ฉบับ ผู้วิจัยจึงได้เสนอแนะว่า การนำรูปแบบของราล์ชมาใช้นั้น การวัดควรจะวัดในจุดประสงค์ที่แท้จริง ทั้งยังต้องพิจารณาถึงข้อตกลงเบื้องต้นของรูปราลซ์ ในการวิเคราะห์ข้อสอบด้วย

โกลด์แมนและราชู (Goldman and Raju, 1988 : 11-19) ได้ศึกษาผลของขนาดกลุ่มตัวอย่างที่แตกต่างกัน 3 กลุ่ม คือ 250 คน, 500 คน และ 1000 คน เมื่อเทียบกับจำนวนของกลุ่มตัวอย่างมาตรฐาน 3,000 คน ในการกะประมาณค่าพารามิเตอร์ ผลการศึกษาพบว่า กลุ่มตัวอย่าง 250 คน มีความเหมาะสมสำหรับราล์ชโมเดล แต่ถ้านำสำหรับโมเดล 2 พารามิเตอร์ จะต้องใช้กลุ่มตัวอย่าง 1,000 คนขึ้นไป จึงจะสามารถกะประมาณค่าพารามิเตอร์ได้ถูกต้อง

สำหรับ งานวิจัยในประเทศ มีผู้ทำการวิจัยหลายท่านดังรายละเอียดของงานวิจัยแต่ละคนดังนี้
อวยพร วิบูลย์กาญจน์ (2526 : 60) ได้เปรียบเทียบผลการวิเคราะห์ข้อสอบอุปมาอุปไมยด้วย คลาสสิกคอล กับ ราล์ชโมเดล โดยใช้กลุ่มตัวอย่างเป็นนักเรียนประถมศึกษาปีที่ 6 จำนวน 1,884 คน เครื่องมือที่ใช้เป็นแบบทดสอบอุปมาอุปไมย และแบบทดสอบมาตรฐานวิชาวิทยาศาสตร์ของสำนักงานทดสอบทางการศึกษา และจิตวิทยามหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร วิเคราะห์ข้อสอบด้วยคอมพิวเตอร์ โปรแกรม BICAL และโปรแกรมวิเคราะห์ข้อสอบทดสอบความแตกต่างระหว่างจำนวนข้อสอบ โดยใช้ค่าสถิติทดสอบซี (Z-test) แล้วพบว่าวิธีคลาสสิกคอลได้ 34 ข้อ และราล์ชโมเดลได้ 26 ข้อ หากค่าความเชื่อมั่นโดยสูตร KR-20 ได้ค่าความเชื่อมั่น .85 และราล์ชโมเดลได้ .80 ส่วนความเที่ยงตรงเชิงสภาพแบบ คลาสสิกคอลได้ .49 และแบบราล์ชโมเดลได้ .46

ปนัดดา วัชรวัฒน์ (2528:71) ได้ศึกษาผลการวิเคราะห์ข้อสอบด้วยวิธีแบบเดิมกับวิธีโลจิสติกโมเดล ทั้งสามพารามิเตอร์ คือ โมเดลที่ใช้พารามิเตอร์ตัวเดียว พารามิเตอร์ 2 ตัว และพารามิเตอร์ 3 ตัว กลุ่มตัวอย่างเป็นนักเรียน ชั้นมัธยมศึกษาปีที่ 1 จำนวน 1,687 คน เครื่องมือเป็นแบบทดสอบคณิตศาสตร์ชนิดเลือกตอบ 5 ตัวเลือก จำนวน 50 ข้อ พบว่า จำนวนข้อสอบที่ได้รับการคัดเลือกจากการวิเคราะห์แบบ โลจิสติกโมเดล ทั้ง 3 พารามิเตอร์ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ระดับ .05 ความเชื่อมั่นของข้อสอบที่ได้รับการคัดเลือกจากวิธีโลจิสติกที่ใช้พารามิเตอร์ 1 ตัว และ 3 ตัว เมื่อปรับจำนวน ข้อให้เท่ากัน แล้ว มีค่าไม่แตกต่างกัน และเมื่อปรับจำนวนข้อของข้อสอบให้เท่ากันแล้ว ค่าความเชื่อมั่นของการวิเคราะห์แบบดั้งเดิมกับวิธีโลจิสติกไม่แตกต่างกัน

บุญชิต รอดแก้ว (2533) ได้ศึกษาความไม่แปรปรวนของค่าพารามิเตอร์ของข้อสอบที่วิเคราะห์ด้วยโมเดลโลจิสติก 3 พารามิเตอร์ กับกลุ่มตัวอย่างที่มีลักษณะต่างกันพบว่าค่าอำนาจจำแนกของข้อสอบของกลุ่ม 1,052 2,104 และ 3,155 คน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติสำหรับค่าความยากของกลุ่ม 1,052 2,104 และ 3,155 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.1

วลีมาส แซ่ฮ้ง (2537) ศึกษาความคลาดเคลื่อนมาตรฐานและประสิทธิภาพสัมพัทธ์ของการประมาณค่าเจตคติต่อวิชาคณิตศาสตร์ระหว่าง มาตราวัดแบบลิเคิร์ตกับมาตราวัดแบบอสมุต ผลการวิจัยพบว่าค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าเจตคติต่อวิชาคณิตศาสตร์ระหว่างมาตราวัดแบบลิเคิร์ตกับมาตราวัดแบบอสมุต มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าเจตคติต่อวิชาคณิตศาสตร์ ระหว่างมาตราวัดแบบลิเคิร์ตเทียบกับมาตราวัดแบบอสมุต มีค่าต่ำกว่า 1 ในกลุ่มผู้ตอบที่มีระดับเจตคติปานกลาง แต่มีค่าสูงกว่า 1 ในกลุ่มผู้ตอบที่มีระดับเจตคติต่ำมาก ระดับเจตคติต่ำ ระดับเจตคติสูงและระดับเจตคติสูงมาก

รุ่งฤดี ทรัพย์นิธิ (2539) ศึกษาเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ระหว่างการวิเคราะห์โดยโมเดลโลจิสติก แบบหนึ่ง สอง สาม พารามิเตอร์และขนาดกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1000 คน ใช้โปรแกรมไบล็อก (BILOG) พบว่า

1. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากัน เมื่อประมาณค่าด้วย โมเดล 1 พารามิเตอร์ มีค่าแตกต่างกัน เมื่อประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก มีค่าไม่แตกต่างกัน ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนกมีค่าแตกต่างกันในบางกลุ่มตัวอย่าง เมื่อประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ความคลาดเคลื่อนในการประมาณค่าความยาก และการเดา มีค่าไม่แตกต่างกัน ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าแตกต่างกันในบางกลุ่มตัวอย่าง

2. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากัน เมื่อประมาณค่าด้วย โมเดล 1 พารามิเตอร์ มีค่าแตกต่างกัน เมื่อประมาณค่าด้วยโมเดล 2 และ 3 พารามิเตอร์ ปรากฏผลเช่นเดียวกันคือ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ มีค่าแตกต่างกันในบางกลุ่มตัวอย่าง

3. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างโมเดล เมื่อวิเคราะห์ด้วยกลุ่มตัวอย่างที่มีจำนวนเท่ากัน ความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าแตกต่างกัน ทุกขนาดกลุ่มตัวอย่าง ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก มีค่าแตกต่างกันในบางกลุ่มตัวอย่าง

4. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างโมเดล เมื่อวิเคราะห์ด้วยกลุ่มตัวอย่างที่มีจำนวนเท่ากัน มีค่าแตกต่างกัน จากทุกขนาดกลุ่มตัวอย่าง โดยโมเดล 3 พารามิเตอร์ แตกต่างจากโมเดล 1 และ 2 พารามิเตอร์

จากการศึกษาพบว่ากลุ่มตัวอย่างที่นำมาใช้ในการวิเคราะห์ และให้ค่าประมาณพารามิเตอร์ที่ดีที่สุด คือ กลุ่มตัวอย่าง 750 คน

สุพล นิลกลาง (2541) ได้เปรียบเทียบคุณลักษณะ ของแบบทดสอบเลือกตอบระหว่างกลุ่มตัวอย่างที่มีจำนวน 6 กลุ่ม คือ 200 คน 300 คน 450 คน 675 คน 1,015 คน และ 1,610 คน วิเคราะห์ด้วยวิธีซิมูเลชัน โดยศึกษาเกี่ยวกับกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2540 พบว่า ค่าความยากของแบบทดสอบ ค่าความสามารถของผู้สอบ และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

จากการศึกษางานวิจัยที่เกี่ยวข้องทั้งภายในประเทศและต่างประเทศ พบว่าส่วนใหญ่จะศึกษาโดยใช้กลุ่มตัวอย่างจำนวนมาก อย่างต่ำสุดแค่ 200 คน และใช้จำนวนข้อสอบไม่แน่นอน และงานวิจัยก็เป็นเพียงเพื่อใช้รูปแบบการวิเคราะห์ข้อมูลแบบบราสซ์ไปเปรียบเทียบกับกรวิเคราะห์ข้อมูลกับรูปแบบอื่นๆ ผู้วิจัยจึงสนใจที่จะศึกษาความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ที่มีจำนวนข้อและกลุ่มตัวอย่างต่างกัน คือมีจำนวนข้อเพียง 20 ข้อ และ 40 ข้อ และกลุ่มตัวอย่างน้อยๆตั้งแต่ 150 คน 200 คน 250 คน จนถึงสูงสุดเพียง 300 คน ซึ่งยังไม่มียานวิจัยใดศึกษาเรื่องนี้ ซึ่งตั้งสมมติฐานได้ดังนี้

สมมติฐานการวิจัย

1. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างข้อสอบที่มีจำนวนข้อต่างกัน เมื่อกลุ่มตัวอย่างมีจำนวนเท่ากัน มีค่าแตกต่างกัน
2. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อข้อสอบมีจำนวนข้อเท่ากัน มีค่าแตกต่างกัน
3. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างข้อสอบที่มีจำนวนข้อต่างกัน เมื่อกลุ่มตัวอย่างที่มีจำนวนเท่ากัน มีค่าแตกต่างกัน
4. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบที่ ระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อข้อสอบมีจำนวนข้อเท่ากัน มีค่าแตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนในสังกัดสำนักงานการประถมศึกษาปราชินบุรี จำนวน 273 โรงเรียน จำนวนห้องเรียน 328 ห้องเรียน รวมจำนวนนักเรียน 5,290 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนในสังกัดสำนักงานการประถมศึกษาปราชินบุรี ซึ่งได้มาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีขนาดโรงเรียนเป็นชั้น (Strata) มีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) ดังนี้

ชั้นที่ 1 จำแนกโรงเรียนที่มีอยู่ 273 โรงเรียน แบ่งออกเป็น 3 ขนาด คือ ขนาดใหญ่ ขนาดกลาง และขนาดเล็ก (โดยแบ่งตามขนาดโรงเรียนตามเกณฑ์ของสำนักงานการประถมศึกษา) จากการสำรวจพบว่า มีโรงเรียนขนาดใหญ่ 45 โรงเรียน มี 100 ห้องเรียน จำนวนนักเรียน 1,000 คน โรงเรียนขนาดกลาง 97 โรงเรียน มี 111 ห้อง จำนวนนักเรียน 1,940 คน โรงเรียนขนาดเล็ก 131 โรงเรียน มี 131 ห้อง จำนวนนักเรียน 1,350 คน

ชั้นที่ 2 สุ่มโรงเรียนตามขนาดโรงเรียนโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยใช้ขนาดโรงเรียนเป็นชั้น (Strata) มีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) สุ่มมาประมาณ 5 เปอร์เซ็นต์ของแต่ละขนาด ได้โรงเรียนขนาดใหญ่ 4 โรงเรียน ขนาดกลาง 10 โรงเรียน และขนาดเล็ก 6 โรงเรียน ดังแสดงรายละเอียดในตาราง

ตาราง 2 จำนวนกลุ่มตัวอย่าง จำแนกตามขนาดโรงเรียน

กลุ่มตัวอย่างที่ใช้สอบข้อสอบ 40 ข้อ			กลุ่มตัวอย่างที่ใช้สอบข้อสอบ 20 ข้อ		
ขนาดโรงเรียน	โรงเรียน	จำนวนนักเรียน	ขนาดโรงเรียน	โรงเรียน	จำนวนนักเรียน
ใหญ่	1. ชุมชนวัดบางแตน	41	ใหญ่	1. ชุมชนวัดศรีประจันตะคาม	50
	2. ชุมชนบ้านทุ่งโพธิ์	38		2. อนุบาลเมืองปราจีนบุรี	70
กลาง	3. ชุมชนบ้านขอนแก่น	32	กลาง	3. บ้านเขาไม้แก้ว	45
	4. บ้านโคกสั่น	34		4. บ้านวังดินสอ	31
	5. บ้านเนินหอม	36		5. บ้านเขาด้วน	35
	6. บ้านโคกอุดม	32		6. บ้านดงบัง	14
	7. อนุบาลกบินทร์บุรี	37		7. บ้านวังตะเคียน	33
เล็ก	8. วัดมะกอกแก้ว	20	เล็ก	8. บ้านเขากระแด	16
	9. วัดจันทร์าราม	20		9. บ้านเขาปูน	5
	10. บ้านวังกวาง	13		10. วัดบ้านยาง	4
รวม		303	รวม		303

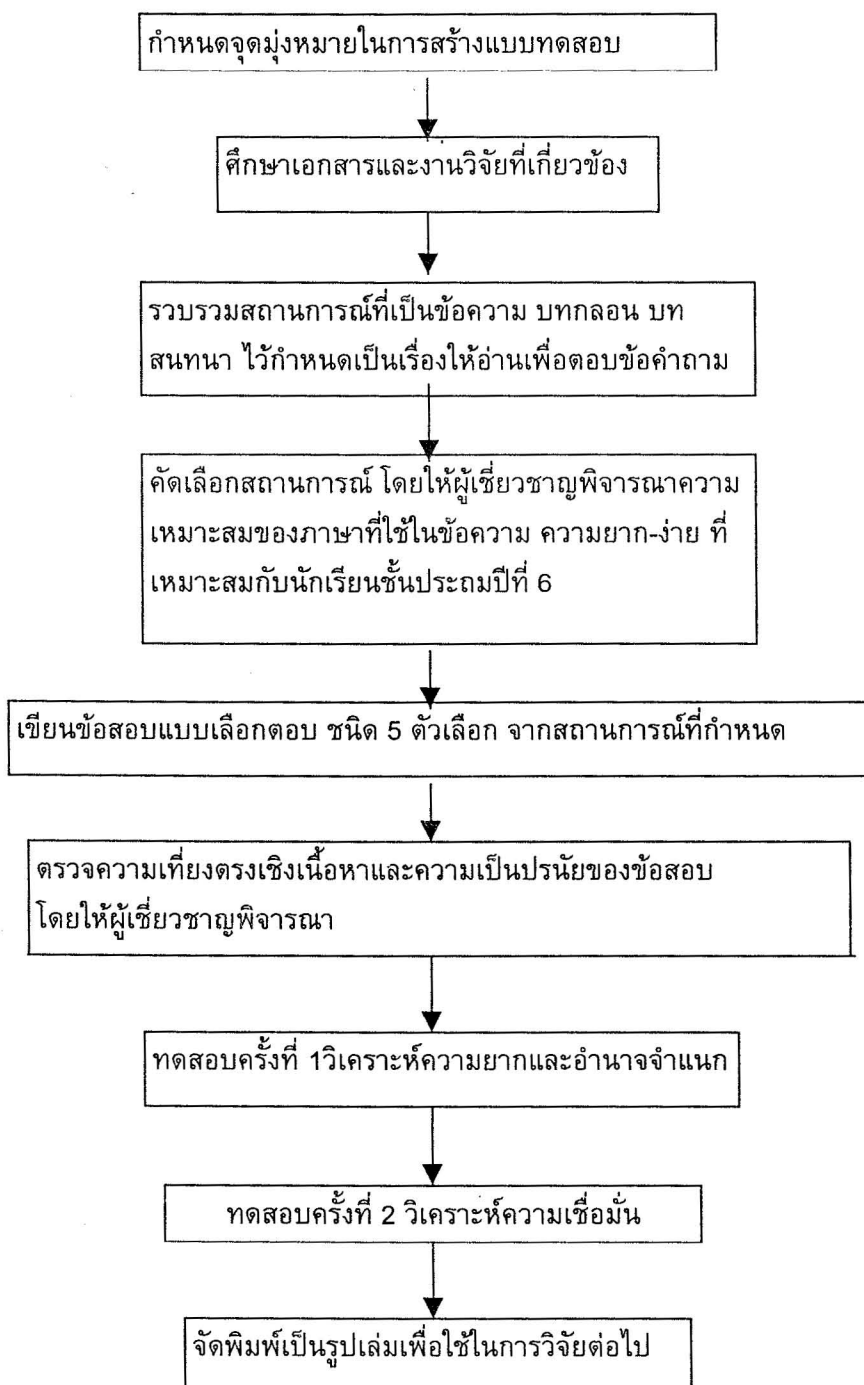
เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบแบบเลือกตอบความเข้าใจในการอ่านภาษาไทย ที่ผู้วิจัยสร้างขึ้นเป็นชนิด 5 ตัวเลือก ที่มี จำนวนข้อต่างกัน 2 ฉบับคือ

1. แบบทดสอบเลือกตอบ จำนวน 40 ข้อ
2. แบบทดสอบเลือกตอบ จำนวน 20 ข้อ

การสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า

ในการศึกษาครั้งนี้ ผู้วิจัยจะสร้างแบบทดสอบวัดความเข้าใจในการอ่านโดยมีขั้นตอนดังนี้



จากภาพประกอบ 12 แสดงขั้นตอนในการสร้างเครื่องมือซึ่งมีรายละเอียดในการสร้างดังนี้

1. กำหนดจุดมุ่งหมายในการสร้าง ในการสร้างแบบทดสอบครั้งนี้มีจุดมุ่งหมาย เพื่อศึกษาความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความยากของแบบทดสอบความเข้าใจในการอ่านภาษาไทย ที่มีจำนวนข้อและจำนวนกลุ่มตัวอย่างต่างกัน

2. ศึกษาเอกสารที่เกี่ยวข้องกับการอ่าน การเขียนข้อสอบวัดความเข้าใจในการอ่านและวิธีการสร้างแบบทดสอบเลือกตอบ และศึกษางานวิจัยที่เกี่ยวข้องกับความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ความยากของแบบทดสอบที่มีจำนวนข้อและจำนวนกลุ่มตัวอย่างต่างกัน

3. รวบรวมสถานการณ์ที่ใช้ในการสร้างข้อสอบ โดยพิจารณาจากเนื้อหาสาระและภาษาที่เหมาะสมกับนักเรียนระดับประถมศึกษาปีที่ 6 ที่วิเคราะห์จากหลักสูตรภาษาไทย จำนวน 30 สถานการณ์

4. คัดเลือกสถานการณ์โดยให้ผู้เชี่ยวชาญพิจารณาความเหมาะสมของภาษาที่ใช้ในข้อความความยาก-ง่าย เหมาะสมกับนักเรียนชั้นประถมศึกษาปีที่ 6 ซึ่งคัดเลือกได้สถานการณ์ทั้งหมด 25 สถานการณ์

5. เขียนข้อสอบเลือกตอบความเข้าใจในการอ่าน โดยเขียนข้อคำถามเพื่อวัดความสามารถในการแปลความ การตีความ และการขยายความของเรื่องที่กำหนดให้อ่าน จากสถานการณ์ที่คัดเลือกได้ในข้อ 4 จำนวน 25 สถานการณ์ โดยเขียนข้อสอบแบบเลือกตอบชนิด 5 ตัวเลือกสถานการณ์ละ 4 ข้อ จะได้ข้อสอบ 100 ข้อ มี ตัวเลือก

6. นำแบบทดสอบวัดความเข้าใจในการอ่านภาษาไทยที่สร้างขึ้นไปให้ผู้เชี่ยวชาญทางด้านวัดผลและผู้เชี่ยวชาญภาษาไทยจำนวน 5 ท่าน ตรวจสอบความเที่ยงตรงเชิงเนื้อหา ความเป็นปรนัยของข้อสอบ มีความชัดเจนในการสื่อความหมาย ข้อคำถามตรงกับสถานการณ์และวัดความสามารถในการแปลความ การตีความ การขยายความและการจับประเด็นใจความและมีค่า IOC ไม่ต่ำกว่า 0.5 ผลปรากฏว่า มีข้อสอบอยู่บางข้อที่ต้องปรับปรุงแก้ไขและมีอยู่ 40 ข้อที่ต้องตัดทิ้ง จนได้แบบทดสอบแบบทดสอบวัดความเข้าใจในการอ่านภาษาไทยจำนวน 60 ข้อ จาก 17 สถานการณ์

7. นำแบบทดสอบในขั้นที่ 6 ไปทดลองสอบครั้งที่ 1 กับนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนอนุบาลนาดี โรงเรียนอนุบาลจังหวัดปราจีนบุรี จำนวน 300 คน แล้วนำผลการทดลองสอบครั้งที่ 1 มาทำการตรวจให้คะแนนด้วยวิธี 0-1 จากนั้นวิเคราะห์ค่าความยากและค่าอำนาจจำแนกโดยใช้โปรแกรมวิเคราะห์ข้อสอบCTIA ซึ่งพัฒนาโดย สุพัฒน์ สุกมลสันต์ (2538) แล้วพิจารณาค่าความยากโดยดูที่ดัชนีค่าความยากและพิจารณาค่าอำนาจจำแนกโดยดูที่สหสัมพันธ์ไบซีเรียล ปรากฏว่า แบบทดสอบมีดัชนีค่าความยากอยู่ระหว่าง 0.051 - 0.824 และค่าอำนาจจำแนกอยู่ระหว่าง -0.241 - 0.707 (รายละเอียดแสดงในตารางภาคผนวก) คัดเลือกข้อสอบที่มีดัชนีค่าความยากอยู่ระหว่าง 0.20-0.80 และค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป ปรากฏว่า แบบทดสอบได้ที่มีคุณภาพอยู่ในเกณฑ์ 40 ข้อจาก 16 สถานการณ์ ซึ่งมีดัชนีค่าความยากระหว่าง 0.209 - 0.780 และมีค่าอำนาจจำแนกอยู่ระหว่าง 0.215 - 0.707

8. นำข้อสอบที่มีค่าความยากและอำนาจจำแนกที่ถึงตามเกณฑ์ที่กำหนดไว้มาทดลองสอบครั้งที่ 2 กับนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนเมืองปราจีนบุรี โรงเรียนบ้านโป่งใหญ่ โรงเรียนบ้านคลองสิบสอง จำนวน 300 คน จากนั้นนำมาวิเคราะห์ความเชื่อมั่นของข้อสอบทั้งฉบับ ได้ค่าความเชื่อมั่น 0.807 สำหรับข้อสอบฉบับนี้เป็นแบบทดสอบที่มี 40 ข้อ ส่วนฉบับที่มี 20 ข้อ ได้จากการสุ่มจากฉบับ 40 ข้อ จาก 16 สถานการณ์ ได้แบบทดสอบฉบับ 20 ข้อ จาก 9 สถานการณ์

9. นำข้อสอบที่คัดเลือกได้มาจัดพิมพ์เป็นรูปเล่มเพื่อใช้ในการวิจัยต่อไป

ลักษณะแบบทดสอบที่ใช้ในการวิจัย

แบบทดสอบที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นแบบทดสอบเลือกตอบชนิด 5 ตัวเลือก วัดความเข้าใจในการอ่านภาษาไทย

ตัวอย่างแบบทดสอบ

อ่านบทสนทนาแล้วตอบคำถามข้อ 1-4 (เป็นสถานการณ์ที่ไม่มีในบทเรียนของนักเรียน)

ต้นไม้ : นี่นะ ! ท่านสิงโต ตั้งแต่ท่านมาอาศัยอยู่ได้ร่มเงาของเรา...ยังไม่เคยเห็นท่านทำ

ประโยชน์อะไรให้กับเราบ้างเลย

สิงโต : ท่านกล่าวหนักเกินไปแล้ว รู้ไหมที่ท่านอยู่รอดปลอดภัยจนถึงทุกวันนี้ ก็เพราะเราช่วย

ปก ป้องท่าน

ต้นไม้ : มากไป! เจ้าไม่สำนึกบุญคุณของข้าแล้ว ยังมาพูดจาโอหัง เจ้าออกไปได้แล้ว

สิงโต : ได้เมื่อท่านไล่ข้าก็จะไป

หลายวันต่อมา มีชาวบ้านมาโค่นต้นไม้ต้นนั้นเอาไปทำฟืนเผาถ่าน

(นิตยสาร สู้ชีวิต)

1. ต้นไม้มีนิสัยอย่างไร

- ก. ใจร้อน
- ข. ใจแคบ
- ค. ใจร้าย
- ง. ใจน้อย
- จ. ใจแข็ง

คำตอบถูก ข

3. ส่วนใดมีความหมายตรงกับข้อความนี้

- ก. หมาในรางหญ้า
- ข. ปลาตกน้ำตัวโต
- ค. แกงจืดจึงรู้คุณเกลือ
- ง. ปลาหมอตายเพราะปาก
- จ. คับที่อยู่ได้คับใจอยู่ยาก

คำตอบถูก ค

2. ข้อความในเรื่อง ให้ข้อคิดในเรื่องใด

- ก. ความรัก
- ข. ความอดทน
- ค. ความสามัคคี
- ง. ความเสียสละ
- จ. ความเอื้ออารี

คำตอบถูก จ

4. ถ้าต้นไม้ไม่กล่าวกับสิงโตเช่นนี้ อะไรจะเกิดขึ้น

- ก. ต้นไม้กลัวสิงโต
- ข. ต้นไม้ถูกโค่นไปทำฟืน
- ค. สิงโตเอาเปรียบต้นไม้ต่อไป
- ง. ต้นไม้และสิงโตจะทะเลาะกันไม่สิ้นสุด
- จ. ต้นไม้และสิงโตอยู่กันอย่างมีความสุข

คำตอบถูก จ

การตรวจให้คะแนน

การตรวจให้คะแนนของแบบทดสอบวิชาภาษาไทย ชนิด 5 ตัวเลือกนี้ใช้เกณฑ์ การตรวจให้คะแนนคือ ถ้าตอบถูกให้ 1 คะแนน ถ้าตอบผิดหรือเว้นช่องว่างไว้ในข้อนั้นๆให้ 0 คะแนน

วิธีดำเนินการเก็บรวบรวมข้อมูล

1. ติดต่อกับหัวหน้าสถานศึกษาที่ถูกเลือกเป็นกลุ่มตัวอย่าง เพื่อขอความร่วมมือในการตอบแบบทดสอบ โดยกำหนดวัน เวลาและสถานที่ในการสอบ
2. จัดเตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่สอบ
3. ดำเนินการเก็บรวบรวมข้อมูลโดยนำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่าง โดยดำเนินการสอบดังนี้
 - 3.1 ชี้แจงนักเรียนที่เป็นกลุ่มตัวอย่างให้ทราบวัตถุประสงค์ของการสอบและขอความร่วมมือในการสอบเพื่อให้ผลที่ได้เป็นจริงที่สุด
 - 3.2 แบบทดสอบทั้ง 2 ฉบับ ดำเนินการสอบโดยสอบตามโรงเรียนที่สุ่มได้ในกลุ่มตัวอย่างที่แยกไว้แล้วว่านักเรียนโรงเรียนใดสอบแบบทดสอบฉบับ 20 ข้อ หรือ 40 ข้อ
4. นำกระดาษคำตอบที่ได้มาตรวจให้คะแนน ผู้ตอบตอบเพียงตัวเลือกเดียว ตรวจให้คะแนนแบบ 0-1 หากคำตอบของผู้สอบตรงกับเฉลยให้ 1 คะแนน หากคำตอบไม่ตรงกับเฉลยหรือไม่ตอบให้ 0 คะแนน
5. นำคะแนนที่ได้มาวิเคราะห์หาค่าสถิติต่อไป

ขั้นตอนในการวิเคราะห์ข้อมูล

ขั้นตอนในการวิเคราะห์ข้อมูล ผู้วิจัยได้ดำเนินการตามขั้นตอนต่อไปนี้

1. วิเคราะห์ข้อมูลในขั้นตอนการทดลองเครื่องมือ ได้ดำเนินการดังนี้
 - 1.1 ตรวจให้คะแนนตามเกณฑ์การให้คะแนนที่ตั้งไว้
 - 1.2 ค่าความยาก ค่าอำนาจจำแนกข้อสอบเป็นรายข้อ และความเชื่อมั่นของข้อสอบโดยใช้โปรแกรมสำเร็จรูป CTIA
2. วิเคราะห์ข้อมูลในขั้นตอนการวิจัย ได้ดำเนินการดังนี้
 - 2.1 ตรวจให้คะแนนตามเกณฑ์การให้คะแนนที่ตั้งไว้
 - 2.2 ในแบบทดสอบที่มี 40 ข้อ โดยนำกระดาษคำตอบของนักเรียน 300 คน มาสุ่มออกเป็น 5 กลุ่ม โดยสุ่มออกทีละกลุ่มแบบไม่ใส่คืน เมื่อครบจำนวนแล้วใส่กระดาษคำตอบคืนแล้วจึงสุ่มจับกลุ่มต่อไป ได้กลุ่มตัวอย่างที่มีขนาดต่างกัน ดังนี้
 - กลุ่มที่ (1) 100 คน
 - กลุ่มที่ (2) 150 คน
 - กลุ่มที่ (3) 200 คน
 - กลุ่มที่ (4) 250 คน
 - กลุ่มที่ (5) 300 คน

คำนวณหาค่าพารามิเตอร์ความยากของผู้สอบ ค่าความสามารถของผู้สอบ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ

2.3 ในแบบทดสอบฉบับที่มี 20 ข้อ จะนำกระดาษคำตอบของนักเรียน 300 คน มาทำ
เช่นเดียวกับข้อ 2.2

3. ทดสอบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบและค่าความสามารถของผู้สอบระหว่างที่มีจำนวนข้อต่างกัน เมื่อกลุ่มตัวอย่างมีจำนวนเท่ากันใช้ t-test

4. ทดสอบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบและค่าความสามารถของผู้สอบที่มีกลุ่มตัวอย่างต่างกันเมื่อมีข้อสอบจำนวนเท่ากันใช้ One-Way ANOVA ถ้าพบความแตกต่าง ทดสอบเป็นรายคู่ด้วยวิธีของ Tukey-Kramer Test

สถิติที่ใช้วิเคราะห์ข้อมูล

1. สถิติที่ใช้ในการหาคุณภาพเครื่องมือ

1.1 ค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน

1.2 ค่าความยากหาโดยใช้สูตร (สุพรรณิ สุขมลสันต์. 2542 :20)

$$p = \frac{N_a}{N_t}$$

N_a แทน จำนวนผู้สอบทั้งหมดที่ตอบตัวเลือก g

N_t แทน จำนวนผู้สอบทั้งหมด

p แทน ค่าความยากของตัวเลือกแต่ละข้อสำหรับผู้สอบทั้งหมด

1.3 ค่าอำนาจจำแนก หาโดยสูตร (สุพรรณิ สุขมลสันต์. 2542 :23)

$$r_{pb} = \frac{\bar{X}_r - \bar{X}_w}{S.D._t} \sqrt{p(1-p)}$$

r_{pb} แทน ดัชนีค่าอำนาจจำแนก

$\bar{X}_r$ แทน ค่าเฉลี่ยของคะแนนผู้สอบที่ตอบข้อสอบนั้นถูก

$\bar{X}_w$ แทน ค่าเฉลี่ยของคะแนนผู้สอบที่ตอบข้อสอบนั้นผิด

$S.D._t$ แทน ค่า S.D. ของคะแนนรวมของผู้สอบทั้งหมด

p แทน อัตราส่วนของผู้ที่ตอบข้อสอบข้อนั้นถูก

1.4 ค่าความเชื่อมั่นของแบบทดสอบ หาโดยสูตร Kuder-Richardson Formula 20

(สัพฒน์ สุกมลสันต์. 2542)

$$r_{ii} = \frac{k}{k-1} \left[1 - \frac{\sum p_g q_g}{S.D.^2_i} \right]$$

k แทน จำนวนข้อสอบทั้งหมด

p_g แทน อัตราส่วนของผู้ที่ตอบถูกแต่ละข้อ

q_g แทน $(1 - p_g)$

$S.D.^2_i$ แทน ความแปรปรวนของคะแนนสอบของผู้สอบทั้งหมด

ในการวิจัยข้อ 1.1, 1.2, 1.3, และ 1.4 วิเคราะห์หาคูณภาพโดยใช้โปรแกรม CTIA ซึ่งพัฒนาโดย สัพฒน์ สุกมลสันต์ (2542)

2. สถิติที่ใช้ในการวิจัย

2.1 ค่าพารามิเตอร์ความยาก (b) ของข้อสอบ (สัพฒน์ สุกมลสันต์ .2542: 141)

$$b_g = M + Y \ln \left(\frac{N - S_g}{S_g} \right)$$

M แทน ค่าเฉลี่ยของความสามารถ (θ) ของบุคคล

y แทน ค่าแก้ไขความคลาดเคลื่อนของ θ ซึ่งได้จาก $Y = \sqrt{1 - \frac{\sigma^2}{2.89}}$

σ^2 แทน ความแปรปรวนของความสามารถ

N แทน จำนวนผู้สอบทั้งหมด

S_g แทน จำนวนคะแนนที่มีผู้ตอบข้อทดสอบนั้นได้

2.2 ค่าความสามารถของบุคคล (θ) (สัพฒน์ สุกมลสันต์ .2542: 141)

$$\theta_g = H + X \ln [r_v | (L - r_v) |]$$

H แทน ค่าเฉลี่ยของระดับความยากทั้งหมด

X แทน ค่าแก้ไขความคลาดเคลื่อนของ b ซึ่งได้มาจาก

$$X = \sqrt{1 + \frac{W^2}{2.89}}$$

W^2 แทน ความแปรปรวนของระดับความยาก

L แทน ความยาวของข้อสอบทั้งหมด

r_v แทน คะแนนของผู้ที่มีความสามารถระดับ v

2.3 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ

2.4 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก

ในการวิจัยครั้งนี้ การวิเคราะห์ข้อ 2.1-2.4 ใช้โปรแกรม BICAL

3. ทดสอบความแตกต่างของค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบและค่าความสามารถของผู้สอบของกลุ่มตัวอย่างระหว่างที่มีจำนวนข้อต่างกันเมื่อกลุ่มตัวอย่างมีจำนวนเท่ากัน ใช้ t-test ทดสอบ (ล้วน สายยศ และอังคณา สายยศ.2540.)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

$\bar{x}_1, \bar{x}_2$ แทน ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าของข้อสอบฉบับที่ 1,2 ตามลำดับ

s_1^2, s_2^2 แทน ความแปรปรวนของข้อสอบฉบับที่ 1,2 ตามลำดับ

n_1, n_2 แทน ขนาดของข้อสอบฉบับที่ 1,2 ตามลำดับ

4. ทดสอบความแตกต่างของค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบและค่าความสามารถของผู้สอบของกลุ่มตัวอย่างระหว่างกลุ่มตัวอย่างต่างกันเมื่อข้อสอบมีจำนวนเท่ากัน โดยวิเคราะห์ความแปรปรวนแบบทิศทางเดียว (One-Way ANOVA) และถ้าพบความแตกต่างอย่างมีนัยสำคัญทางสถิติ จะตรวจสอบความแตกต่างรายคู่ด้วยวิธีของ Tukey-Kramer Test (ล้วน สายยศ และอังคณา สายยศ.2540.321-324)

$$F = \frac{MS_b}{MS_w}$$

$$qTK = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{MS_E \left\langle \frac{1}{n_1} + \frac{1}{n_2} \right\rangle / 2}}$$

MS_b แทน ความแปรปรวนระหว่างกลุ่ม

MS_w แทน ความแปรปรวนภายในกลุ่ม

F แทน การแจกแจงของ F ที่ α

$\bar{x}_1$ แทน ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าของข้อสอบกลุ่มที่ 1

$\bar{x}_2$ แทน ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าของข้อสอบกลุ่มที่ 2

n_1 แทน จำนวนสมาชิกของกลุ่มที่ 1

n_2 แทน จำนวนสมาชิกของกลุ่มที่ 2

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการศึกษาครั้งนี้ เพื่อให้เกิดความเข้าใจตรงกันในการเสนอและแปลผลการวิเคราะห์ข้อมูล ผู้วิจัยจึงได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูลดังนี้

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

b	แทน	ค่าความยากของข้อสอบ
$\bar{b}$	แทน	ค่าความยากเฉลี่ยของข้อสอบ
θ	แทน	ค่าความสามารถของผู้สอบ
$\bar{\theta}$	แทน	ค่าความสามารถเฉลี่ยของผู้สอบ
SD_b	แทน	ค่าเบี่ยงเบนมาตรฐานของความยากของข้อสอบ
SS	แทน	ผลบวกกำลังสองของคะแนนแบบทดสอบ
MS	แทน	ความแปรปรวนของคะแนนแบบทดสอบ
df	แทน	ชั้นของความเป็นอิสระ

การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลดังนี้

1. ค่าสถิติพื้นฐาน
 - 1.1 ค่าความยากของข้อสอบ และค่าความสามารถของผู้สอบ
 - 1.2 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากและค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ
2. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ระหว่างที่มีจำนวนข้อต่างกัน เมื่อกลุ่มตัวอย่างที่มีจำนวนเท่ากัน
3. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างที่มีจำนวนข้อต่างกัน เมื่อกลุ่มตัวอย่างมีจำนวนเท่ากัน
4. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อข้อสอบมีจำนวนข้อเท่ากัน
5. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อข้อสอบมีจำนวนข้อเท่ากัน

ผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐาน

จากการนำแบบทดสอบความเข้าใจในการอ่านจำนวน 40 ข้อ และแบบทดสอบความเข้าใจในการอ่านจำนวน 20 ข้อ ไปทดสอบกับนักเรียน อย่างละ 300 คน แล้วแบ่งแต่ละกลุ่มตัวอย่างออกเป็น 5 กลุ่ม ใช้วิธีสุ่มอย่างง่ายโดยสุ่มออกทีละกลุ่มแบบไม่ใส่คืน เมื่อครบจำนวนแล้วใส่คืนแล้วจึงสุ่มจับกลุ่มต่อไป ซึ่งได้แก่กลุ่ม 100, 150, 200, 250 และ 300 คน ได้ค่าสถิติพื้นฐาน จำแนกตามตัวแปรของข้อสอบดังนี้

1.1 ค่าความยากของข้อสอบ และค่าความสามารถของผู้สอบ

ในการวิเคราะห์หาค่าความยากของข้อสอบ และค่าความสามารถของผู้สอบนั้นได้จากการทดสอบนักเรียนชั้นประถมศึกษาปีที่ 6 มาคำนวณหาค่า ความยากของข้อสอบ ค่าความยากเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน ค่าความสามารถของผู้สอบ และค่าความสามารถเฉลี่ย ที่วิเคราะห์โดยวิธีราล์ช โมเดล ได้ผลดังตาราง 3

ตาราง 3 ค่าความยากของข้อสอบ ค่าความยากเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน และค่าความสามารถของผู้สอบ ของแบบทดสอบความเข้าใจในการอ่าน จำนวน 20 ข้อ และ 40 ข้อ ที่มีจำนวนกลุ่มตัวอย่าง 100, 150, 200, 250, และ 300 คน

กลุ่มตัวอย่าง	แบบทดสอบ 20 ข้อ				แบบทดสอบ 40 ข้อ			
	b	$\bar{b}$	SD	θ	b	$\bar{b}$	SD	θ
100	-1.999ถึง1.212	0.002	1.19	-3.33ถึง3.33	-2.263ถึง1.900	-0.001	0.78	-4.23ถึง4.23
150	-1.717ถึง1.726	-0.008	1.23	-3.37ถึง3.37	-1.869ถึง1.897	-0.001	0.94	-4.21ถึง4.21
200	-1.797ถึง1.723	-0.003	1.10	-3.30ถึง3.30	-2.154ถึง2.010	-0.001	1.07	-4.23ถึง4.23
200	-1.797ถึง1.723	-0.003	1.10	-3.30ถึง3.30	-2.154ถึง2.010	-0.001	1.07	-4.23ถึง4.23
250	-1.380ถึง1.440	0.003	1.14	-3.32ถึง3.32	-1.403ถึง1.869	0.000	1.07	-4.14ถึง4.14
300	-1.507ถึง1.486	0.003	1.12	-3.26ถึง3.26	-1.871ถึง1.929	-0.001	1.01	-4.18ถึง4.18

จากตาราง 3 ผลการวิเคราะห์แสดงให้เห็นว่า แบบทดสอบความเข้าใจในการอ่าน 20 ข้อของกลุ่มตัวอย่าง 100,150, 200, 250และ300 คน มีค่าความยากระหว่าง -1.999ถึง1.212, -1.717ถึง1.726, -1.797ถึง1.723,-1.380ถึง1.440, -1.507ถึง1.486 และ มีค่าความยากเฉลี่ยเท่ากับ 0.002 ,-0.008, -0.003, 0.003, 0.003 ตามลำดับ แสดงว่าข้อสอบมีค่าความยากอยู่ระดับปานกลาง ในกลุ่มตัวอย่างแต่ละ

กลุ่ม สำหรับการวัดการกระจายคะแนนที่พิจารณาจากคะแนนส่วนเบี่ยงเบนมาตรฐาน ของกลุ่มตัวอย่างแต่ละกลุ่มมีค่าเท่ากับ 1.19 , 1.23, 1.10, 1.14, 1.12 ตามลำดับแสดงว่ามีการกระจายคะแนนในแต่ละกลุ่มตัวอย่างใกล้เคียงกัน ส่วนค่าความสามารถของผู้สอบ มีค่าอยู่ระหว่าง -3.33ถึง3.33, -3.37 ถึง 3.37, -3.30 ถึง 3.30, -3.32ถึง3.32, -3.26ถึง3.26 ในกลุ่มตัวอย่างแต่ละกลุ่มแสดงว่าความสามารถของผู้สอบมีค่าต่างกัน คือมีความสามารถในระดับน้อยถึงค่อนข้างมีความสามารถมาก ในแบบทดสอบความเข้าใจในการอ่าน 40 ข้อของกลุ่มตัวอย่าง 100,150, 200, 250และ300 คน มีค่าความยากระหว่าง -2.263 ถึง 1.900, -1.869 ถึง 1.897, -2.154 ถึง 2.010, -1.403 ถึง 1.869, -1.871 ถึง 1.929 มีค่าความยากเฉลี่ย -0.001, -0.001, -0.001, 0.000, -0.001 ตามลำดับ แสดงว่าข้อสอบมีค่าความยากอยู่ระดับปานกลางในกลุ่มตัวอย่างแต่ละกลุ่ม สำหรับการวัดการกระจายคะแนนที่พิจารณาจากคะแนนส่วนเบี่ยงเบนมาตรฐาน ของกลุ่มตัวอย่างแต่ละกลุ่มมีค่า เท่ากับ 0.78, 0.94, 1.07, 1.07, 1.01 ตามลำดับ แสดงว่ามีการกระจายคะแนนในแต่ละกลุ่มตัวอย่างใกล้เคียงกัน ส่วนค่าความสามารถของผู้สอบมีค่าอยู่ระหว่าง -4.23 ถึง 4.23, -4.21 ถึง 4.21, -4.23 ถึง 4.23, -4.14 ถึง 4.14, -4.18 ถึง 4.18 ในกลุ่มตัวอย่างแต่ละกลุ่มแสดงว่าความสามารถของผู้สอบมีค่าต่างกัน คือมีความสามารถในระดับน้อยถึงค่อนข้างมีความสามารถมาก

1.2 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากและค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ

ในการวิเคราะห์หาค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากและ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ นั้นได้จากการทดสอบนักเรียนชั้นประถมศึกษาปีที่ 6 มาคำนวณหาค่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากเฉลี่ย ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถเฉลี่ย ที่วิเคราะห์โดยวิธีราล์ช โมเดล ได้ผลดังตาราง 4

ตาราง 4 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากและค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบทดสอบความเข้าใจในการอ่าน จำนวน 20 ข้อ และ 40 ข้อ ที่มีจำนวนกลุ่มตัวอย่าง 100, 150, 200, 250, และ 300 คน

กลุ่มตัวอย่าง	แบบทดสอบ 20 ข้อ				แบบทดสอบ 40 ข้อ			
	SE_b	$\overline{SE_b}$	SE_θ	$\overline{SE_\theta}$	SE_b	$\overline{SE_b}$	SE_θ	$\overline{SE_\theta}$
100	0.231ถึง0.288	0.2441	0.48ถึง1.09	0.6231	0.219ถึง0.316	0.2390	0.34ถึง1.09	0.4666
150	0.188ถึง0.245	0.2009	0.48ถึง1.10	0.6263	0.181ถึง0.242	0.1946	0.34ถึง1.09	0.4656
200	0.159ถึง0.207	0.1683	0.17ถึง1.09	0.6205	0.159ถึง0.194	0.1705	0.34ถึง1.09	0.4666
250	0.143ถึง0.178	0.1524	0.47ถึง1.07	0.6110	0.141ถึง0.187	0.1501	0.34ถึง1.08	0.4625
300	0.130ถึง0.163	0.1375	0.47ถึง1.08	0.6142	0.127ถึง0.172	0.1365	0.34ถึง1.08	0.4625

จากตาราง 4 ผลการวิเคราะห์แสดงให้เห็นว่า แบบทดสอบความเข้าใจในการอ่าน 20 ข้อของกลุ่มตัวอย่าง 100,150, 200, 250และ300 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบอยู่ระหว่าง 0.231 ถึง 0.288, 0.188 ถึง 0.245, 0.159 ถึง 0.207, 0.143 ถึง 0.178, 0.130 ถึง 0.163 และมีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากเท่ากับ 0.2441, 0.2009, 0.1683, 0.1504, 0.1375 ตามลำดับ แสดงว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากมีค่าน้อย ส่วนค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ อยู่ระหว่าง 0.48 ถึง 1.09, 0.48 ถึง 1.10, 0.47 ถึง 1.09, 0.47 ถึง 1.07, 0.47 ถึง 1.08 และมีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถเท่ากับ 0.6231, 0.6263, 0.62053, 0.6110, 0.6142 ตามลำดับ แสดงว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถมีความคลาดเคลื่อนน้อย สำหรับแบบทดสอบความเข้าใจในการอ่าน 40 ข้อของกลุ่มตัวอย่าง 100,150, 200, 250และ300 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบอยู่ระหว่าง 0.219 ถึง 0.316, 0.181 ถึง 0.242, 0.159 ถึง 0.194, 0.141 ถึง 0.187, 0.127 ถึง 0.172 และมีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากเท่ากับ 0.2390, 0.1946, 0.1705, 0.1501, 0.1365 ตามลำดับแสดงว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากมีค่าน้อย ส่วนค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ อยู่ระหว่าง 0.34 ถึง 1.09, 0.34 ถึง 1.09, 0.34 ถึง 1.09, 0.34 ถึง 1.08, 0.34 ถึง 1.08 และมีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถเท่ากับ 0.4667, 0.4656, 0.4666, 0.4625, 0.4625 ตามลำดับ แสดงว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถมีความคลาดเคลื่อนน้อย

2. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ระหว่างที่มีจำนวนข้อต่างกัน เมื่อกลุ่มตัวอย่างที่มีจำนวนเท่ากัน

ผู้วิจัยได้นำค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ระหว่างที่มีจำนวนข้อต่างกัน คือ 20 และ 40 ข้อ เมื่อกลุ่มตัวอย่างที่มีจำนวนเท่ากัน 100, 150, 200, 250 และ 300 คน มาเปรียบเทียบกับดังในตาราง 5 ดังนี้

ตาราง 5 เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างข้อสอบ 20 ข้อ และ 40 ข้อ เมื่อกลุ่มตัวอย่าง 100,150,200,250 และ300 คน

กลุ่มตัวอย่าง	แบบทดสอบ 20 ข้อ		แบบทดสอบ 40 ข้อ		t
	SE_b	SD	SE_b	SD	
100 คน	0.2441	1.19	0.2390	0.78	0.0180
150 คน	0.2009	1.23	0.1946	0.94	0.0200
200 คน	0.1683	1.10	0.1705	1.07	-0.007
250 คน	0.1504	1.14	0.1501	1.07	0.0001
300 คน	0.1375	1.12	0.1365	1.01	0.0025

จากตาราง 5 แสดงให้เห็นว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ระหว่าง 20 ข้อ และ 40 ข้อ เมื่อกลุ่มตัวอย่าง เท่ากันคือ 100 , 150, 200, 250 และ 300 คน ซึ่งมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบที่มีจำนวนข้อ 20 ข้อ และ 40 ข้อ เมื่อวิเคราะห์ระหว่างจำนวนกลุ่มตัวอย่างเท่ากันจะมีค่าใกล้เคียงกัน

3. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างข้อสอบที่มีจำนวนข้อต่างกัน เมื่อกลุ่มตัวอย่างมีจำนวนเท่ากัน

ผู้วิจัยได้นำค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างที่มีจำนวนข้อต่างกัน คือ 20 และ 40 ข้อ เมื่อกลุ่มตัวอย่างที่มีจำนวนเท่ากัน 100, 150, 200, 250 และ 300 คน มาเปรียบเทียบกับดังในตาราง 6 ดังนี้

ตาราง 6 เปรียบเทียบความคลาดเคลื่อนมาตรฐาน ในการประมาณค่าความสามารถของผู้สอบ ระหว่างข้อสอบ 20 ข้อ และ 40 ข้อ เมื่อกลุ่มตัวอย่าง 100,150,200,250 และ 300 คน

กลุ่มตัวอย่าง	แบบทดสอบ 20 ข้อ		แบบทดสอบ 40 ข้อ		t
	$\overline{SE}_0$	SD	$\overline{SE}_0$	SD	
100 คน	0.6231	1.19	0.4666	0.78	0.5336
150 คน	0.6263	1.23	0.4656	0.94	0.5140
200 คน	0.6205	1.10	0.4667	1.07	0.5154
250 คน	0.6111	1.14	0.4626	1.07	0.4854
300 คน	0.6142	1.12	0.4626	1.01	0.3903

จากตาราง 6 แสดงให้เห็นว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างข้อสอบ 20 ข้อ และ 40 ข้อ เมื่อกลุ่มตัวอย่าง เท่ากันคือจำนวน 100 , 150, 200, 250 และ 300 คนซึ่ง มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเมื่อข้อสอบที่มีจำนวนข้อ 20 ข้อ และ 40 ข้อวิเคราะห์ระหว่างจำนวนกลุ่มตัวอย่างเท่ากันจะมีค่าใกล้เคียงกัน

4. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อข้อสอบมีจำนวนเท่ากัน

4.1 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 20 ข้อ ระหว่างกลุ่มตัวอย่าง 100 คน,150 คน, 200 คน, 250 คน และ 300 คน

ผู้วิจัยได้หาความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากเฉลี่ย ของข้อสอบ 20 ข้อ จากกลุ่มตัวอย่างแต่ละกลุ่ม คือ 100 คน, 150 คน, 200 คน,250 คนและ 300 คน โดยการวิเคราะห์ความแปรปรวน ได้ผลดังตาราง 7 ดังนี้

ตาราง 7 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 20 ข้อ ระหว่างกลุ่มตัวอย่าง 100 คน,150 คน, 200 คน, 250 คน, 300 คน

แหล่งตัวแปร	SS	df	MS	F
ระหว่างกลุ่ม	0.1473	4	0.0368	12.6075 **
ภายในกลุ่ม	0.2775	95	0.0029	
รวมทั้งหมด	0.4248	99		

จากตาราง 7 แสดงว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าของข้อสอบ 20 ข้อ ระหว่างกลุ่มตัวอย่าง 100 คน, 150 คน, 200 คน, 250 คน และ 300 คน มีค่าแตกต่างกันอย่างน้อย 1 คู่ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ดังนั้นผู้วิจัยจึงทดสอบรายคู่ด้วยค่าสถิติการทดสอบ Tukey-Kramer Test ดังแสดงในตาราง 8

ตาราง 8 เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 20 ข้อ เป็นรายคู่

กลุ่มตัวอย่าง	100 คน	150 คน	200 คน	250 คน	300 คน
100 คน	-	8.7127**	16.2660**	20.8333**	24.5622**
150 คน		-	7.8057**	12.6566**	16.6710**
200 คน			-	4.9695**	9.0588**
250 คน				-	4.0345**
300 คน					-

$$q_{(0.01,2,95)} = 3.70$$

จากตาราง 8 แสดงว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 20 ข้อ ระหว่างกลุ่มตัวอย่าง 100 คน กับกลุ่ม 150 , 200 , 250, 300 คน กลุ่มตัวอย่าง 150 คน กับกลุ่ม 200, 250, 300 คน กลุ่มตัวอย่าง 200 คน กับกลุ่ม 250, 300 คน และกลุ่มตัวอย่าง 250 คน กับกลุ่ม 300 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 20 ข้อแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ ระดับ .01 ทุกคู่

4.2 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 40 ข้อ ระหว่างกลุ่มตัวอย่าง 100 คน, 150 คน, 200 คน, 250 คน และ 300 คน

ผู้วิจัยได้หาความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากเฉลี่ย ของข้อสอบ 40 ข้อ จากกลุ่มตัวอย่างแต่ละกลุ่ม คือ 100 คน, 150 คน, 200 คน, 250 คน และ 300 คน โดยการวิเคราะห์ความแปรปรวน ได้ผลดังตาราง 9 ดังนี้

ตาราง 9 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างกลุ่มตัวอย่าง 100 คน, 150 คน, 200 คน, 250 คน, 300 คน เมื่อข้อสอบ 40 ข้อ

แหล่งตัวแปร	SS	df	MS	F
ระหว่างกลุ่ม	0.2617	4	0.0654	53.7517**
ภายในกลุ่ม	0.2374	195	0.0012	
รวมทั้งหมด	0.4991	199	0.0666	

จากตาราง 9 แสดงว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าของข้อสอบ 40 ข้อ ระหว่างกลุ่มตัวอย่าง 100 คน, 150 คน, 200 คน, 250 คน และ 300 คน มีค่าแตกต่างกันอย่างน้อย 1 คู่ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ดังนั้นผู้วิจัยจึงทดสอบรายคู่ด้วยค่าสถิติการทดสอบ Tukey-Kramer Test ดังแสดงในตาราง 10

ตาราง 10 เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 40 ข้อ เป็นรายคู่

กลุ่มตัวอย่าง	100 คน	150 คน	200 คน	250 คน	300 คน
100 คน	-	13.8750**	22.8100**	30.7439**	36.7132**
150 คน		-	8.9664**	17.3024**	23.6953**
200 คน			-	8.7903**	15.5251**
250 คน				-	6.6276**
300 คน					-

$q_{(0.01,2,195)} = 3.64$

จากตาราง 10 แสดงว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 40 ข้อ ระหว่างกลุ่มตัวอย่าง 100 คน กับกลุ่ม 150 , 200 , 250 และ 300 คน กลุ่มตัวอย่าง 150 คน กับกลุ่ม 200, 250 และ 300 คน กลุ่มตัวอย่าง 200 คน กับกลุ่ม 250 และ 300 คน และกลุ่มตัวอย่าง 250 คน กับกลุ่ม 300 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 40 ข้อแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ ระดับ .01 ทุกคู่

5. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อข้อสอบมีจำนวนเท่ากัน

5.1 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างกลุ่มตัวอย่าง 100 คน, 150 คน, 200 คน, 250 คน, และ 300 คน เมื่อข้อสอบ 20 ข้อ

ผู้วิจัยได้หาความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเฉลี่ย ของข้อสอบ 20 ข้อ จากกลุ่มตัวอย่างแต่ละกลุ่ม คือ 100 คน,150 คน,200 คน,250 คนและ 300 คน โดยการวิเคราะห์ความแปรปรวน ได้ผลดังตาราง 11 ดังนี้

ตาราง 11 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่าง
กลุ่มตัวอย่าง 100 คน, 150 คน, 200 คน, 250 คน, และ 300 คน เมื่อข้อสอบ 20 ข้อ

แหล่งตัวแปร	SS	df	MS	F
ระหว่างกลุ่ม	0.0029	4	0.0007	0.0133
ภายในกลุ่ม	5.1173	95	0.0539	
รวมทั้งหมด	5.1202	99	0.0546	

จากตาราง 11 แสดงให้เห็นว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ของกลุ่มตัวอย่าง 100 คน 150 คน 200 คน 250 คน และ 300 คน ที่ใช้ข้อสอบ 20 ข้อ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบมีค่าใกล้เคียงกัน หรือไม่แปรปรวนระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน

5.2 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ
ระหว่างกลุ่มตัวอย่าง 100 คน, 150 คน, 200 คน, 250 คน, และ 300 คน เมื่อข้อสอบ 40 ข้อ

ผู้วิจัยได้หาความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเฉลี่ย ของข้อสอบ 40 ข้อ จากกลุ่มตัวอย่างแต่ละกลุ่ม คือ 100 คน, 150 คน, 200 คน, 250 คน และ 300 คน โดยการวิเคราะห์ความแปรปรวน ได้ผลดังตาราง 12 ดังนี้

ตาราง 12 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่าง
กลุ่มตัวอย่าง 100 คน, 150 คน, 200 คน, 250 คน, และ 300 คน เมื่อข้อสอบ 40 ข้อ

แหล่งตัวแปร	SS	df	MS	F
ระหว่างกลุ่ม	0.5787	4	0.1447	2.9470
ภายในกลุ่ม	7.6307	195	0.0491	
รวมทั้งหมด	8.2094	199	0.1938	

จากตาราง 12 แสดงให้เห็นว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ของกลุ่มตัวอย่าง 100 คน 150 คน 200 คน 250 คน และ 300 คน ที่ใช้ข้อสอบ 40 ข้อ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบมีค่าใกล้เคียงกัน หรือไม่แปรปรวนระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่าง ข้อสอบที่มีจำนวนข้อต่างกัน เมื่อมีกลุ่มตัวอย่างจำนวนเท่ากัน
2. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างกลุ่ม ตัวอย่างที่มีจำนวนต่างกัน เมื่อมีข้อสอบจำนวนข้อเท่ากัน
3. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างข้อสอบที่มีจำนวนข้อต่างกัน เมื่อมีกลุ่มตัวอย่างจำนวนเท่ากัน
4. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่าง กลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อมีข้อสอบจำนวนข้อเท่ากัน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2542 ของโรงเรียนใน สังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี จำนวน 22 โรงเรียน จำนวนนักเรียน 606 คน ซึ่งได้มาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีโรงเรียนเป็นชั้น (Strata) มีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit)

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นแบบทดสอบแบบเลือกตอบความเข้าใจในการอ่านภาษาไทย ที่ผู้วิจัยสร้างขึ้นเป็นชนิด 5 ตัวเลือก ที่มี จำนวนข้อต่างกัน 2 ฉบับคือ

1. แบบทดสอบเลือกตอบ จำนวน 40 ข้อ
2. แบบทดสอบเลือกตอบ จำนวน 20 ข้อ

วิธีดำเนินการเก็บรวบรวมข้อมูล

1. ติดต่อกับหัวหน้าสถานศึกษาที่ถูกเลือกเป็นกลุ่มตัวอย่าง เพื่อขอความร่วมมือในการตอบแบบทดสอบ โดยกำหนดวัน เวลาและสถานที่ในการสอบ
2. จัดเตรียมแบบทดสอบให้เพียงพอกับจำนวนนักเรียนที่สอบ
3. ดำเนินการเก็บรวบรวมข้อมูลโดยนำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่าง โดยดำเนินการสอบดังนี้
 - 3.1 ชี้แจงนักเรียนที่เป็นกลุ่มตัวอย่างให้ทราบวัตถุประสงค์ของการสอบและขอความร่วมมือในการสอบเพื่อให้ผลที่ได้เป็นจริงที่สุด
 - 3.2 นำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่าง

4. นำกระดาษคำตอบที่ได้มาตรวจให้คะแนน ผู้ตอบตอบเพียงตัวเลือกเดียว ตรวจให้คะแนนแบบ 0-1 หากคำตอบของผู้สอบตรงกับเฉลยให้ 1 คะแนน หากคำตอบไม่ตรงกับเฉลยหรือไม่ตอบให้ 0 คะแนน

5. นำคะแนนที่ได้มาวิเคราะห์หาค่าสถิติพื้นฐาน คำนวณหาค่าความยาก ค่าความสามารถของผู้สอบ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ด้วยโปรแกรม BICAL และทดสอบสมมติฐานการวิจัย

การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลดังนี้

1. ค่าสถิติพื้นฐาน

1.1 ค่าความยากของข้อสอบ และค่าความสามารถของผู้สอบ

1.2 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากและค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ

2. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ระหว่างที่มีจำนวนข้อต่างกัน เมื่อกลุ่มตัวอย่างที่มีจำนวนเท่ากัน

3. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างที่มีจำนวนข้อต่างกัน เมื่อกลุ่มตัวอย่างมีจำนวนเท่ากัน

4. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อข้อสอบมีจำนวนข้อเท่ากัน

5. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อข้อสอบมีจำนวนข้อเท่ากัน

สรุปผลการวิเคราะห์ข้อมูล

ผลการวิเคราะห์ข้อมูลสรุปได้ดังนี้

1. ค่าสถิติพื้นฐาน

1.1 ค่าความยากของข้อสอบ ค่าการวัดการกระจายคะแนน และค่าความสามารถของผู้สอบของ แบบทดสอบความเข้าใจในการอ่านที่มี 20 ข้อ และ 40 ข้อ ของกลุ่มตัวอย่าง 100, 150, 200, 250 และ 300 คน มีค่าความยากเฉลี่ย-0.008 ถึง 0.003 และ -0.001 ถึง 0.000 ตามลำดับแสดงว่าแต่ละฉบับมีค่าเข้าใกล้กับ 0 ซึ่งหมายความว่ามีความยากอยู่ระดับปานกลาง สำหรับค่าการวัดการกระจายคะแนนของกลุ่มตัวอย่างแต่ละกลุ่มมีค่า 1.10 ถึง 1.23 และ 0.78 ถึง 1.07 ตามลำดับแสดงว่า มีการกระจายคะแนนในแต่ละกลุ่มตัวอย่างใกล้เคียงกัน และค่าความสามารถของผู้สอบมีค่า ระหว่าง -3.37 ถึง 3.37 และ -4.23 ถึง 4.23 แสดงว่าความสามารถของผู้สอบมีค่าต่างกัน คือมีความสามารถในระดับน้อยถึงค่อนข้างมีความสามารถมาก

1.2 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากและค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของ แบบทดสอบความเข้าใจในการอ่าน 20 และ 40 ข้อ ของกลุ่มตัวอย่าง 100, 150, 200, 250 และ 300 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบมีค่าอยู่ระหว่าง 0.1375 ถึง 0.2441 และ 0.1365 ถึง 0.2390 ตามลำดับ แสดงว่ามีค่าความคลาดเคลื่อนมาตรฐานในการประมาณความยากของข้อสอบทั้งสองฉบับมีค่าใกล้เคียงกัน และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ มีค่าอยู่ระหว่าง 0.6110 ถึง 0.6263 และ 0.4625 ถึง 0.4666 แสดงว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของ ผู้สอบมีค่าใกล้เคียงกัน

2. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ระหว่าง 20 ข้อ และ 40 ข้อ เมื่อกลุ่มตัวอย่างเท่ากันคือ 100 , 150, 200, 250 และ 300 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบเมื่อวิเคราะห์ระหว่างจำนวนข้อสอบต่างกันจะมีค่าใกล้เคียงกัน

3. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างข้อสอบ 20 ข้อ และ 40 ข้อ เมื่อกลุ่มตัวอย่างมีจำนวนเท่ากันคือ 100, 150, 200, 250 และ 300 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างแบบทดสอบที่มีจำนวนข้อต่างกันเมื่อกลุ่มตัวอย่างมีจำนวนเท่ากันมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบมีค่าใกล้เคียงกัน

4. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 20 ข้อ และ 40 ข้อ ระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน คือกลุ่มตัวอย่าง 100 คน, 150 คน, 200 คน, 250 คน และ 300 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ ระดับ .01 ทุกคู่ แสดงว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างข้อสอบที่มีจำนวนข้อเท่ากัน เมื่อกลุ่มตัวอย่างที่มีจำนวนต่างกัน มีค่าแตกต่างกัน หรือ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบแปรปรวนไปตามกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากัน

5. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบของข้อสอบ 20 ข้อ และ 40 ข้อ ระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน คือ 100 คน, 150 คน, 200 คน, 250 คน, และ 300 คน เมื่อข้อสอบ 20 ข้อ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นคือ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบมีค่าใกล้เคียงกัน หรือไม่แปรปรวนระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน

อภิปรายผล

1. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ระหว่าง 20 ข้อ และ 40 ข้อ เมื่อกลุ่มตัวอย่าง เท่ากันคือ 100 , 150, 200, 250 และ 300 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่เป็นไปตามสมมติฐานข้อที่ 1 แสดงว่าขนาดความยาวของแบบทดสอบไม่มีผลต่อค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ ซึ่งสอดคล้องกับการศึกษาของ สุปัทม์ สุกมลสันต์ (2542 : 40) ที่พบว่า การประมาณค่าความยากของข้อสอบนั้นไม่ต้องคำนึงถึงความยาวของแบบทดสอบหรือขนาดของกลุ่มตัวอย่างแต่อย่างใด และยังสอดคล้องกับราสช์ (Rasch .1983: 1429-A) ที่ศึกษาพบว่าค่าพารามิเตอร์ที่ได้จากการวัดแต่ละกลุ่มมีการกระจายที่ไม่เปลี่ยนแปลงไปตามกลุ่มผู้สอบและแบบทดสอบ แสดงว่า ความยาวของข้อสอบไม่มีผลต่อค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ

2. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างข้อสอบที่มีจำนวน 20 ข้อ และ 40 ข้อ เมื่อกลุ่มตัวอย่าง เท่ากัน คือ 100, 150, 200, 250 และ 300 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่เป็นไปตามสมมติฐานข้อที่ 3 แสดงว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบมีค่าคงที่ไม่แปรผันไปตามความยาวของข้อสอบ ซึ่งเป็นไปตามคำกล่าวของราสช์(Rasch .1983: 1429-A) ที่ศึกษาพบว่าค่าพารามิเตอร์ที่ได้จากการวัดแต่ละกลุ่มมีการกระจายที่ไม่เปลี่ยนแปลงไปตามกลุ่มผู้สอบและแบบทดสอบ แสดงว่าความยาวของข้อสอบไม่มีผลต่อค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ

3. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อข้อสอบมีจำนวนเท่ากัน ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ 20 ข้อ และ 40 ข้อ ระหว่างกลุ่มตัวอย่าง 100 คน, 150 คน, 200 คน, 250 คน และ 300 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ ระดับ .01 ซึ่งสอดคล้องกับสมมติฐานข้อที่ 2 และสอดคล้องกับการศึกษาของบุญชิต รอดแก้ว (2533 : 75) ที่พบว่า ในการประมาณค่าความยากของแบบทดสอบกลุ่ม 1,052 คน กับกลุ่ม 2,104 คน และกลุ่ม 1,052 คน กับกลุ่ม 3,155 คน มีค่าความยากของข้อสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และจากการศึกษาของรุ่งฤดี ทรัพย์นิธิ (2539 : 89) พบว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบระหว่างกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากัน เมื่อประมาณค่าด้วยโมเดล 1 พารามิเตอร์ มีค่าแตกต่างกัน แสดงว่า จำนวนกลุ่มตัวอย่างต่างกันมีผลความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ

4. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกัน เมื่อข้อสอบมีจำนวนเท่ากัน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานข้อ 4 ซึ่งเป็นไปตามคำกล่าวของลอร์ด (Lord. 1980 : 35) ที่กล่าวว่าค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบไม่แปรผันไปตามกลุ่มผู้สอบ และผลการศึกษาครั้งนี้ยังได้ผลสอดคล้องกับการศึกษาของค็อก (Kock. 1980 :15032) ที่ศึกษาค่าพารามิเตอร์จากการใช้แบบทดสอบวัดเจตคติตามแนวคิดของลิเกิร์ตที่ยืนยันว่าค่าพารามิเตอร์ของข้อสอบมีคุณสมบัติไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง และยังสอดคล้องกับการศึกษาของสุพัทธ์ สุกมลสันต์ (2542 : 40) ที่พบว่า การประมาณค่าความยากของข้อสอบนั้นไม่ต้องคำนึงถึงความยาวของแบบทดสอบหรือขนาดของกลุ่มตัวอย่างแต่อย่างใด และยังสอดคล้องกับราสช์ (Rasch .1983: 1429-A) ที่ศึกษาพบว่าค่าพารามิเตอร์ที่ได้จากการวัดแต่ละกลุ่มมีการกระจายที่ไม่เปลี่ยนแปลงไปตามกลุ่มผู้สอบและแบบทดสอบ และยังสอดคล้องกับของสุพล นิลกลาง (2541:52)

ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบที่วิเคราะห์จากกลุ่มตัวอย่างที่มีจำนวนต่างกัน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ข้อเสนอแนะในการวิจัย

1. ข้อเสนอแนะจากผลการวิจัย

ข้อสอบที่ใช้วิเคราะห์โดยวิธีราชสโมเดลนั้น ใช้ข้อสอบเพียง 20 ข้อได้

2. ข้อเสนอแนะในการทำวิจัย

2.1 ควรมีการศึกษาในทำนองเดียวกันนี้ในเนื้อหาวิชาอื่น หรือในระดับชั้นที่แตกต่างกันไป

2.2 ควรมีการศึกษาในทำนองเดียวกัน โดยใช้ 2 พารามิเตอร์ 3 พารามิเตอร์ มาวิเคราะห์ข้อมูลและหาความสัมพันธ์กับผลการศึกษาในครั้งนี้

บรรณานุกรม

บรรณานุกรม

- โกวิท ประวาลพุกษ์. (2525 ,กันยายน-สิงหาคม). “ความเข้าใจพื้นฐานเกี่ยวกับโครงสร้างข้อสอบ,” ใน การวัดผลการศึกษา.
- ต่าย เชียงฉิน. (2534). การศึกษาเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบเทเลอรรูปปริมาตร ที่มีรูปแบบ จำนวนชั้น และวิธีการให้คะแนนที่แตกต่างกัน โดยใช้วิธีมอนติ คาร์โล. ปรินซ์นิพนธ์ กศ.ด. (การวิจัยและพัฒนาหลักสูตร). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- บุญชิต รอดแก้ว. (2533). การศึกษาความไม่แปรปรวนของค่าพารามิเตอร์ของข้อสอบที่วิเคราะห์ด้วยโมเดลโลจิสติกสามพารามิเตอร์กับกลุ่มตัวอย่างที่มีลักษณะต่างกัน. ปรินซ์นิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ปนัดดา วัชรวัฒน์. (2528). การศึกษาผลการวิเคราะห์ข้อสอบโดยใช้วิธีโลจิสติกโมเดลกับวิธีแบบเดิม (คลาสสิกอล) ในวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1. วิทยานิพนธ์ ศศ.ม.(การวิจัยการศึกษา). เชียงใหม่ : บัณฑิตวิทยาลัย มหาวิทยาลัยเชียงใหม่. ถ่ายเอกสาร.
- ผจงจิต อินทสุวรรณ. (2528).ทฤษฎีการตอบข้อคำถาม. ม.ป.ท.
- พิชญา เกตุสกุล. (2539). การเปรียบเทียบความเที่ยงตรงและความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบระหว่างแบบทดสอบแยกกลุ่มที่มีรูปแบบต่างกัน. ปรินซ์นิพนธ์ กศ.ม (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- รุ่งฤดี ทรัพย์นิธิ. (2540). การศึกษาเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ระหว่างการวิเคราะห์ด้วยโมเดลโลจิสติก แบบหนึ่ง สอง สาม พารามิเตอร์และขนาดกลุ่มตัวอย่างต่างกัน. ปรินซ์นิพนธ์ กศ.ม (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ล้วน สายยศ และอังคณา สายยศ. (2540). สถิติวิทยาทางการวิจัย. พิมพ์ครั้งที่ 3. กรุงเทพฯ : สุวีริยาสาส์น.
- วลีมาส แซ่อึ้ง. (2537). การศึกษาความคลาดเคลื่อนมาตรฐานและประสิทธิภาพสัมพัทธ์ของการประมาณค่าเจตคติต่อวิชาคณิตศาสตร์ระหว่างมาตรวัดแบบลิเคอร์ท์กับมาตรวัดแบบบอัสกูด. ปรินซ์นิพนธ์กศ.ม (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สงบ ลักษณะ. (2525,พฤษภาคม-สิงหาคม) “ความสามารถเกี่ยวกับการตอบข้อสอบ” ใน การวัดผลการศึกษา 4 (1): 47-54,88-94.
- สมพร บุญอ้อม. (2529). การศึกษาความคงที่ของค่าพารามิเตอร์ความยากในการวิเคราะห์ข้อกระทงด้วยวิธีซิมเพล. วิทยานิพนธ์ ค.ม. (การวิจัยทางการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- สุพล นิลกลาง. (2541). การเปรียบเทียบคุณลักษณะของแบบทดสอบเลือกตอบระหว่างกลุ่มตัวอย่างที่มีจำนวนต่างกันที่วิเคราะห์ด้วยวิธีซิมเพล. ปรินซ์นิพนธ์ กศ.ม (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

- สุพัฒน์ สุขมลสันต์. (2530). *การวิเคราะห์ข้อสอบแนวใหม่ด้วยเครื่องคอมพิวเตอร์*. กรุงเทพฯ: สถาบัน
ภาษาจุฬาลงกรณ์มหาวิทยาลัย.
- สำเร็จ บุญเรืองรัตน์. (2529, มกราคม-เมษายน). “การพัฒนาทฤษฎีลาเทนท์เทรทเพื่อวิเคราะห์ข้อสอบ”
ใน *วารสารการวัดผลการศึกษา*. 7 (2): 44-68 .
- อวยพร วิบูลย์กาญจน์. (2526). *การเปรียบเทียบผลการวิเคราะห์แบบสอบอุปมาอุปไมยด้วยคลาสสิกอล
โมเดลกับราล์ฟโมเดล*. วิทยานิพนธ์ ค.ม. (การวิจัยทางการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย
จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- ไอลดา คล้ายสำริด. (2538). *การเปรียบเทียบความคลาดเคลื่อนในการประมาณค่าประสิทธิภาพสัมพัทธ์ใน
การประมาณค่าและความสามารถของผู้สอบในแบบทดสอบวิชาคณิตศาสตร์ที่ใช้วิธีต่างกัน*. ปรินญา
นิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.
ถ่ายเอกสาร.
- Andre De Champlain and Mare E. Gessaroli. (1998,May). “Assessing the Dimensionality of Item
Response Matrices With Small Sample Sizes and Short Test Lengths,” *Applied
Measurement in education*. 11(3): 199, 231-253.
- Baker,R.B. (1977,January). “Advances in item analyses,” *Review of educational Research*. 47(1) :
151-178 .
- Birnbaum, Allan. (1968,May). “Some latent Trait Model and Their Use in Inferring in Examinee’s
Ability,” In F.M.Novick.” *Statistical Theories of Mental Test Scores*. 4(2):22
- Cook,Linda L. Daniel Eignor and Hessay L,Taft. (1988,Summer). “A comparative Study of The
Effects of Recency of Instruction on The Stability of IRT and Conventional Item Parameter
Estimates,” *Journal of Educational Measurement*. 25(1) : 31-45.
- Duglass,I.B. (1981,Winter). “A Comparison of Item Characteristic Curve Model for Classroom
Examination System,” *Dissertation Abstracts International*. 41 (9): 4000-A..
- Hambleton, Ronald K. (1989). “Principles and Selected Applications of Item Response Theory,”
Educational Measurement. 3 rd Ed. New York : American Council on Education and
Macmillan Publishing Company A Division of Macmillan, Inc.
- Hambleton,R.K. and H.Swaminathan. (1985). *Item Response Theory*. Boston : Kluwer - Nijhoff
Publishing.
- Hambleton,R.K. and L. Cook. (1977,Summer) “Latent Trait Model and their Use in the Analysis of
Educational Test Data,” *Journal Educational Measurement*. 4 (2): 75-96 .
- Hutten. (1982, May) “The Fit of Empirical Data to Two Latent Trait Models,” *Dissertation Abstracts
International*. 42 (11): 4779-A.
- Koch,W.R. (1983, Winter) “Likert Scaling Using the Grade Response Latent Trait Model,” *Applied
Psychological Measurement*. 7 (1): 15-32.
- Lord,F.M. (1980). *Application of Item Response Theory to Practical Testing Problems*. New Jersey :
Lawrence Erlbaum Associates Publishers.

- Lord, F.M. and M.R. Novick. (1968). *Statistical Theory of Mental Test Scores*. Massachusetts : Addison-Wesley Publishing Company Inc.
- Lord, Fedderic M. and M.L. Stocking. (1988) "Item Response Theory," in *Educational Research Methodology' and Measurement: An International Handbook*. P.269-272. Editedd by John p. Keeves : Pergamon Press Plc.
- Rasch, G. (1960) *PROBABILISTIC models for Some Intellingence and Attainment tests*. Copenhagen Denmark. : Diedoggiogiske Institute.
- Tinsley and Dawis. (1975, Summer). "An In Investingation of the Rasch single Logistic Model. Sample Free Item and Test Calobration," *Educational and Psychological Measurement*. 35(1) : 325-339.
- Warm, T.A. (1978). "A Primer of Item Response Theory," *Technical Report 941278*. Oklahoma : U.S. Coast Guard Institute.
- Weiss, D.J. (1974) "Strategies of Adaptive Ability Measurement," *Research Report 74-5*. Minneapolis : University of Minnesota, Psychometric Methods Program, Department of Psychology.
- Wright, B.D. and N. Panchapakesan, (1969, July). "A Procedure for Sample Free Item Anaiysis," *Educational and Psychological Measurement*. 29(7):23-48.

ภาคผนวก ก.

ค่าความยากและค่าอำนาจจำแนกของข้อสอบ

ตาราง 13 แสดงค่าความยากและค่าอำนาจจำแนกของแบบทดสอบเลือกตอบความเข้าใจในการอ่าน
จำนวน 60 ข้อ ในการทดลองสอบครั้งที่ 1

ข้อ	ตัวเลือก	P	r	Δ	ข้อ	ตัวเลือก	P	r	Δ
1	ก	0.189	0.101	16.56	8**	ก	0.024	0.025	20.97
	ข	0.037	0.114	20.18		ข	0.068	0.139	19.02
	ค*	0.243	0.114	15.82		ค*	0.676	0.418	11.14
	ง	0.044	0.076	19.87		ง	0.216	0.228	16.18
	จ	0.473	-0.190	13.31		จ	0.010	0.000	22.32
2**	ก	0.223	0.203	16.09	9**	ก*	0.551	0.380	12.45
	ข	0.162	0.076	16.98		ข	0.041	0.089	20.02
	ค	0.024	0.051	20.97		ค	0.128	0.025	17.57
	ง	0.010	0.013	22.32		ง	0.051	0.101	19.59
	จ*	0.574	0.215	12.21		จ	0.216	0.139	16.18
3**	ก	0.213	0.241	16.22	10**	ก	0.247	0.076	15.78
	ข	0.034	0.051	20.35		ข	0.037	0.076	20.18
	ค	0.041	0.063	20.02		ค*	0.270	0.291	15.49
	ง	0.044	0.025	19.87		ง	0.074	0.000	18.81
	จ*	0.655	0.418	11.36		จ	0.365	0.114	14.42
4**	ก	0.014	0.025	21.88	11	ก*	0.051	-0.089	19.59
	ข	0.024	0.025	20.97		ข	0.010	0.038	22.32
	ค*	0.689	0.291	10.99		ค	0.024	0.076	20.97
	ง	0.203	0.114	16.37		ง	0.534	-0.025	12.62
	จ	0.061	0.114	19.23		จ	0.368	-0.177	14.39
5	ก	0.014	0.025	21.88	12**	ก	0.098	0.165	18.21
	ข*	0.730	0.076	10.51		ข*	0.605	0.329	11.90
	ค	0.169	-0.038	16.87		ค	0.091	0.063	18.37
	ง	0.057	0.038	19.34		ง	0.095	0.051	18.29
	จ	0.017	0.000	21.53		จ	0.095	0.038	18.29
6**	ก	0.030	0.101	20.54	13**	ก	0.027	0.025	20.74
	ข	0.054	0.076	16.46		ข	0.257	0.190	15.65
	ค	0.220	0.170	16.13		ค	0.118	0.038	17.77
	ง*	0.517	0.367	12.79		ง*	0.334	0.291	14.75
	จ	0.172	0.025	16.82		จ	0.250	0.063	15.74
7	ก	0.142	-0.076	17.79	14	ก	0.034	0.089	20.35
	ข	0.243	0.076	15.82		ข*	0.318	0.114	14.94
	ค	0.128	0.013	17.57		ค	0.321	0.063	14.90
	ง*	0.402	0.000	14.03		ง	0.220	-0.101	16.13
	จ	0.078	0.013	18.72		จ	0.101	0.038	18.14

ตาราง 13 (ต่อ)

ข้อ	ตัวเลือก	P	r	Δ	ข้อ	ตัวเลือก	P	r	Δ
15**	ก	0.071	0.127	18.91	22**	ก	0.078	0.051	18.72
	ข	0.057	0.165	19.34		ข	0.459	0.013	13.45
	ค	0.024	0.089	20.97		ค	0.07	0.253	18.81
	ง	0.037	0.063	20.18		ง	0.166	0.139	16.93
	จ*	0.799	0.468	9.54		จ*	0.209	0.228	16.27
16**	ก	0.034	0.051	20.35	23	ก	0.071	0.215	18.91
	ข	0.172	0.063	16.82		ข	0.034	0.114	20.35
	ค	0.071	0.114	18.91		ค*	0.429	-0.241	13.75
	ง*	0.402	0.380	14.03		ง	0.274	0.089	15.45
	จ	0.314	0.127	14.97		จ	0.182	-0.038	16.66
17	ก	0.020	0.076	21.24	24	ก	0.125	0.152	17.64
	ข	0.064	0.114	19.12		ข	0.189	0.000	16.56
	ค	0.044	0.127	19.87		ค*	0.267	0.076	15.53
	ง	0.037	0.036	20.18		ง	0.054	0.089	19.46
	จ*	0.824	0.418	9.23		จ	0.348	-0.203	14.60
18	ก	0.108	0.101	17.98	25**	ก*	0.774	0.380	9.96
	ข*	0.351	0.316	14.57		ข	0.051	0.063	19.59
	ค	0.068	0.114	19.02		ค	0.037	0.101	20.18
	ง	0.095	0.203	18.29		ง	0.034	0.063	20.35
	จ	0.365	-0.139	14.42		จ	0.091	0.127	18.37
19**	ก*	0.780	0.494	9.87	26**	ก	0.088	0.165	18.45
	ข	0.027	0.051	20.74		ข*	0.764	0.443	10.09
	ค	0.064	0.177	19.12		ค	0.101	0.152	18.14
	ง	0.051	0.051	19.59		ง	0.014	0.038	21.88
	จ	0.064	0.165	19.12		จ	0.017	0.051	21.53
20	ก	0.044	0.127	19.87	27**	ก	0.037	0.089	20.18
	ข	0.304	0.025	15.09		ข*	0.764	0.532	10.09
	ค	0.213	-0.089	16.22		ค	0.024	0.076	20.97
	ง	0.199	-0.013	16.42		ง	0.078	0.241	18.72
	จ*	0.223	0.101	16.09		จ	0.081	0.076	18.63
21**	ก	0.074	0.177	18.81	28**	ก	0.199	0.127	16.42
	ข*	0.799	0.354	8.61		ข*	0.392	0.278	14.14
	ค	0.027	0.051	20.74		ค	0.193	0.013	16.51
	ง	0.007	0.013	22.92		ง	0.068	0.063	19.02
	จ	0.020	0.076	21.24		จ	0.135	0.025	17.45

ตาราง 13 (ต่อ)

ข้อ	ตัวเลือก	P	r	Δ	ข้อ	ตัวเลือก	P	r	Δ
29	ก	0.030	0.114	20.54	36**	ก	0.101	0.051	18.14
	ข	0.152	0.013	17.15		ข*	0.625	0.443	11.69
	ค	0.514	-0.025	12.82		ค	0.139	0.228	17.39
	ง*	0.236	0.165	15.91		ง	0.061	0.114	19.23
	จ	0.057	0.025	19.34		จ	0.057	0.013	19.34
30**	ก	0.084	0.203	18.54	37**	ก*	0.706	0.544	10.79
	ข	0.111	0.177	17.91		ข	0.081	0.101	18.63
	ค	0.051	0.076	19.59		ค	0.068	0.177	19.02
	ง*	0.676	0.532	11.14		ง	0.054	0.114	19.46
	จ	0.057	0.013	19.34		จ	0.071	0.076	18.91
31**	ก*	0.618	0.709	11.76	38**	ก*	0.618	0.519	11.76
	ข	0.064	0.177	19.12		ข	0.041	0.101	20.02
	ค	0.196	0.278	16.46		ค	0.057	0.165	19.34
	ง	0.064	0.139	19.12		ง	0.182	0.127	16.66
	จ	0.041	0.063	20.02		จ	0.088	0.076	18.45
32	ก*	0.334	0.342	14.75	39	ก	0.078	0.089	18.72
	ข	0.199	0.152	16.42		ข*	0.399	0.127	14.07
	ค	0.095	0.228	18.29		ค	0.409	-0.228	13.96
	ง	0.111	0.152	17.91		ง	0.064	0.101	19.12
	จ	0.247	-0.241	15.78		จ	0.034	0.127	20.53
33	ก*	0.821	0.519	9.29	40**	ก*	0.574	0.418	12.21
	ข	0.054	0.139	19.46		ข	0.101	0.203	18.14
	ค	0.037	0.101	20.18		ค	0.155	0.051	17.09
	ง	0.037	0.101	20.18		ง	0.122	0.076	17.71
	จ	0.020	0.063	21.24		จ	0.034	0.038	20.35
34**	ก*	0.794	0.468	9.68	41**	ก	0.105	0.228	18.06
	ข	0.020	0.100	21.24		ข*	0.632	0.620	11.61
	ค	0.071	0.203	18.96		ค	0.044	0.089	19.87
	ง	0.044	0.114	19.87		ง	0.068	0.076	19.02
	จ	0.057	0.101	19.34		จ	0.135	0.168	17.45
35**	ก*	0.500	0.443	13.00	42**	ก	0.209	0.190	16.27
	ข	0.041	0.101	20.02		ข	0.088	0.215	18.45
	ค	0.213	0.241	16.22		ค	0.081	0.190	18.63
	ง	0.189	0.038	16.56		ง	0.030	0.025	20.54
	จ	0.044	0.089	19.87		จ*	0.578	0.671	12.18

ตาราง 13 (ต่อ)

ข้อ	ตัวเลือก	P	r	Δ	ข้อ	ตัวเลือก	P	r	Δ
43**	ก	0.176	0.051	16.79	50**	ก	0.159	0.191	17.04
	ข*	0.274	0.329	15.45		ข	0.128	0.253	17.57
	ค	0.108	0.190	17.98		ค	0.051	0.127	19.59
	ง	0.125	0.190	17.64		ง*	0.209	0.316	16.24
	จ	0.297	0.165	15.17		จ	0.436	0.076	13.69
44**	ก	0.395	0.101	14.10	51**	ก	0.095	0.114	18.29
	ข*	0.331	0.291	14.79		ข	0.091	0.152	18.37
	ค	0.186	0.076	16.61		ค*	0.470	0.608	13.35
	ง	0.044	0.101	19.87		ง	0.233	0.215	15.95
	จ	0.017	0.063	21.53		จ	0.091	0.051	18.37
45	ก	0.051	0.114	19.59	52**	ก	0.179	0.063	16.71
	ข	0.240	0.139	15.87		ข	0.101	0.101	18.14
	ค	0.068	0.203	18.02		ค	0.098	0.291	18.21
	ง*	0.199	0.392	16.42		ง	0.216	0.127	16.18
	จ	0.426	-0.101	13.79		จ*	0.375	0.443	14.31
46**	ก	0.047	0.038	19.73	53**	ก	0.236	0.127	15.91
	ข*	0.733	0.519	10.47		ข*	0.217	0.481	12.79
	ค	0.068	0.177	19.02		ค	0.071	0.190	18.81
	ง	0.098	0.203	18.21		ง	0.030	0.051	20.54
	จ	0.037	0.051	20.18		จ	0.111	0.013	17.91
47**	ก	0.054	0.114	19.46	54**	ก	0.054	0.114	19.46
	ข	0.139	0.127	17.39		ข*	0.520	0.342	12.76
	ค*	0.422	0.506	13.82		ค	0.098	0.215	18.21
	ง	0.084	0.127	18.54		ง	0.064	0.114	19.12
	จ	0.287	0.087	15.29		จ	0.243	-0.177	15.82
48	ก	0.412	-0.291	13.93	55	ก*	0.666	0.544	11.25
	ข	0.105	0.089	18.06		ข	0.068	0.101	19.02
	ค	0.216	0.038	16.18		ค	0.088	0.127	18.45
	ง	0.118	0.215	17.77		ง	0.111	0.165	17.91
	จ*	0.132	0.101	17.51		จ	0.051	0.089	19.59
49**	ก	0.122	0.089	17.71	56	ก	0.088	0.152	18.45
	ข	0.166	0.076	16.93		ข	0.088	0.228	18.45
	ค	0.389	0.127	14.17		ค	0.074	0.177	18.81
	ง*	0.250	0.392	15.74		ง	0.034	0.000	20.35
	จ	0.061	0.051	19.23		จ*	0.693	0.646	10.95

ตาราง 13 (ต่อ)

ข้อ	ตัวเลือก	P	r	Δ	ข้อ	ตัวเลือก	P	r	Δ
57	ก	0.115	0.101	17.84	59**	ก*	0.524	0.468	12.72
	ข	0.098	0.152	18.21		ข	0.091	0.228	18.37
	ค	0.084	0.139	18.54		ค	0.081	0.203	18.63
	ง	0.078	0.165	18.72		ง	0.091	0.013	18.37
	จ*	0.595	0.646	12.00		จ	0.193	0.013	16.51
58	ก	0.189	0.025	16.56	60**	ก	0.098	0.177	18.21
	ข	0.257	0.127	15.65		ข	0.078	0.127	18.72
	ค*	0.213	0.139	16.22		ค*	0.642	0.646	11.51
	ง	0.152	0.013	17.15		ง	0.047	0.086	19.73
	จ	0.166	0.025	16.93		จ	0.118	0.190	17.77

* คือตัวเลือกถูก
** คือข้อที่คัดเลือกไว้

ภาคผนวก ข

ค่าความยาก ค่าความสามารถ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก
และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ

ตาราง 14 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 20 ข้อ
กลุ่มตัวอย่าง 100 คน

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
1	-0.059	-3.33	0.232	1.09
2	-0.387	-2.48	0.237	0.79
3	-0.919	-1.96	0.252	0.67
4	-0.331	-1.57	0.235	0.59
5	0.262	-1.24	0.231	0.55
6	0.804	-0.96	0.237	0.52
7	-0.675	-0.70	0.244	0.50
8	0.102	-0.46	0.231	0.49
9	-1.999	-0.23	0.315	0.48
10	1.091	0.00	0.244	0.48
11	-0.675	0.23	0.244	0.48
12	-1.637	0.46	0.288	0.49
13	-0.331	0.70	0.235	0.50
14	-0.005	0.96	0.231	0.52
15	0.748	1.24	0.235	0.55
16	0.860	1.57	0.238	0.59
17	0.368	1.96	0.231	0.67
18	0.422	2.48	0.231	0.79
19	1.151	3.33	0.245	1.09
20	1.212	*	0.247	*
รวม			4.883	11.84

ตาราง 15 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 20 ข้อ
กลุ่มตัวอย่าง 150 คน

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
1	-0.171	-3.37	0.191	1.10
2	-0.747	-2.52	0.203	0.80
3	-0.830	-1.99	0.205	0.67
4	-0.135	-1.59	0.190	0.60
5	0.541	-1.26	0.189	0.55
6	0.831	-0.97	0.192	0.52
7	-0.788	-0.71	0.204	0.50
8	-0.319	-0.46	0.193	0.49
9	-1.717	-0.23	0.245	0.48
10	1.375	0.00	0.204	0.48
11	-0.969	0.23	0.209	0.48
12	-1.658	0.46	0.241	0.49
13	-0.208	0.71	0.191	0.50
14	-0.135	0.97	0.190	0.52
15	0.757	1.26	0.191	0.55
16	1.056	1.59	0.196	0.60
17	0.116	1.99	0.188	0.67
18	0.399	2.52	0.188	0.80
19	0.868	3.37	0.192	1.10
20	1.726	*	0.216	*
รวม			4.018	11.9

ตาราง 16 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 20 ข้อ
กลุ่มตัวอย่าง 200 คน

ข้อที่	b	θ	SE_b	SE_θ
1	-0.005	-3.30	0.159	1.09
2	-0.661	-2.46	0.167	0.79
3	-0.832	-1.95	0.171	0.66
4	0.325	-1.55	0.160	0.59
5	0.376	-1.23	0.160	0.55
6	0.635	-0.95	0.162	0.52
7	-0.444	-0.69	0.163	0.50
8	-0.183	-0.45	0.160	0.48
9	-1.517	-0.23	0.194	0.48
10	1.076	0.00	0.171	0.47
11	-0.861	0.23	0.171	0.48
12	-1.797	0.45	0.207	0.48
13	-0.339	0.69	0.162	0.50
14	-0.209	0.95	0.160	0.52
15	0.504	1.23	0.161	0.55
16	0.768	1.55	0.164	0.59
17	0.350	1.95	0.160	0.66
18	0.247	2.46	0.160	0.79
19	0.661	3.30	0.163	1.09
20	1.723	*	0.192	*
รวม			3.367	11.79

ตาราง 17 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 20 ข้อ
กลุ่มตัวอย่าง 250 คน

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
1	-0.174	-3.22	0.146	1.07
2	-0.460	-2.40	0.151	0.78
3	-0.694	-1.90	0.156	0.65
4	0.119	-1.52	0.144	0.58
5	0.426	-1.20	0.143	0.54
6	0.589	-0.93	0.143	0.51
7	-0.670	-0.68	0.155	0.49
8	-0.138	-0.44	0.147	0.48
9	-1.380	-0.22	0.178	0.47
10	1.096	0.00	0.148	0.47
11	-0.718	0.22	0.156	0.47
12	-1.380	0.44	0.178	0.48
13	-0.089	0.68	0.145	0.49
14	-0.260	0.93	0.147	0.51
15	0.569	1.20	0.143	0.54
16	0.528	1.52	0.143	0.58
17	0.242	1.90	0.143	0.65
18	0.365	2.40	0.143	0.78
19	0.692	3.22	0.144	1.07
20	1.440	*	0.155	*
รวม			3.008	11.61

ตาราง 18 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 20 ข้อ
กลุ่มตัวอย่าง 300 คน

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
1	-0.301	-3.26	0.134	1.08
2	-0.541	-2.43	0.138	0.78
3	-0.541	-1.92	0.138	0.66
4	0.083	-1.53	0.131	0.59
5	0.557	-1.22	0.130	0.54
6	0.693	-0.94	0.131	0.51
7	-0.737	-0.68	0.142	0.49
8	-0.337	-0.45	0.135	0.48
9	-1.425	-0.22	0.163	0.47
10	1.136	0.00	0.136	0.47
11	-0.861	0.22	0.145	0.47
12	-1.507	0.45	0.167	0.48
13	-0.229	0.68	0.133	0.49
14	-0.141	0.94	0.132	0.51
15	0.625	1.22	0.131	0.54
16	0.676	1.53	0.131	0.59
17	0.320	1.92	0.130	0.66
18	0.337	2.43	0.130	0.78
19	0.710	3.26	0.131	1.08
20	1.486	*	0.143	*
รวม			2.751	11.67

ตาราง 19 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 40 ข้อ
กลุ่มตัวอย่าง 100 คน

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
1	-0.468	-4.23	0.220	1.09
2	-1.291	-3.40	0.239	0.78
3	-0.913	-2.90	0.227	0.64
4	-0.419	-2.53	0.219	0.57
5	0.311	-2.24	0.226	0.51
6	1.712	-2.00	0.299	0.48
7	-0.913	-1.79	0.227	0.45
8	-0.275	-1.60	0.219	0.42
9	0.978	-1.43	0.250	0.41
10	-2.263	-1.27	0.299	0.39
11	0.683	-1.12	0.237	0.38
12	-1.070	-0.98	0.231	0.37
13	-1.592	-0.84	0.253	0.36
14	1.386	-0.71	0.274	0.36
15	-1.291	-0.59	0.239	0.35
16	-0.371	-0.47	0.219	0.35
17	-0.275	-0.35	0.219	0.34
18	-1.070	-0.23	0.231	0.34
19	0.014	-0.12	0.221	0.34
20	1.173	0.00	0.260	0.34
21	-1.124	0.12	0.233	0.34
22	1.312	0.23	0.269	0.34
23	-0.712	0.35	0.223	0.34
24	-0.228	0.47	0.219	0.35
25	0.111	0.59	0.222	0.35
26	-0.228	0.71	0.219	0.36
27	0.111	0.84	0.222	0.36
28	0.210	0.98	0.223	0.37
29	1.173	1.12	0.260	0.38
30	0.740	1.27	0.239	0.39
31	-0.228	1.43	0.219	0.41
32	0.628	1.60	0.235	0.42

ตาราง 19 (ต่อ)

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
33	0.466	1.79	0.230	0.45
34	0.856	2.00	0.244	0.48
35	-0.419	2.24	0.219	0.51
36	1.625	2.53	0.291	0.57
37	1.900	2.90	0.316	0.64
38	0.311	3.40	0.226	0.78
39	0.062	4.23	0.221	1.09
40	-0.613	*	0.221	*

ตาราง 20 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 40 ข้อ
กลุ่มตัวอย่าง 150 คน

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
1	-1.224	-4.21	0.201	1.09
2	-1.617	-3.38	0.218	0.78
3	-0.845	-2.89	0.190	0.64
4	-0.365	-2.52	0.182	0.56
5	0.093	-2.23	0.181	0.51
6	1.295	-1.99	0.209	0.47
7	-0.917	-1.78	0.191	0.45
8	-0.365	-1.59	0.182	0.42
9	0.777	-1.42	0.191	0.41
10	-1.869	-1.26	0.232	0.39
11	0.927	-1.11	0.196	0.38
12	-0.954	-0.97	0.192	0.37
13	-1.307	-0.84	0.204	0.36
14	1.624	-0.71	0.226	0.36
15	-1.224	-0.59	0.201	0.35
16	-0.465	-0.47	0.183	0.35
17	0.159	-0.35	0.181	0.34
18	-1.067	-0.23	0.196	0.34
19	-0.201	-0.11	0.181	0.34
20	0.814	0.00	0.192	0.34
21	-1.029	0.11	0.195	0.34
22	1.125	0.23	0.203	0.34
23	-0.600	0.35	0.185	0.34
24	-0.299	0.47	0.181	0.35
25	0.324	0.59	0.183	0.35
26	-0.005	0.71	0.181	0.36
27	0.291	0.84	0.183	0.36
28	0.258	0.97	0.182	0.37
29	1.624	1.11	0.226	0.38
30	0.741	1.26	0.191	0.39
31	-0.299	1.42	0.181	0.41
32	0.563	1.59	0.187	0.42

ตาราง 20 (ต่อ)

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
33	0.392	1.78	0.184	0.45
34	1.166	1.99	0.204	0.47
35	-0.634	2.23	0.185	0.51
36	1.295	2.52	0.209	0.56
37	1.897	2.89	0.242	0.64
38	0.528	3.38	0.186	0.78
39	0.060	4.21	0.181	1.09
40	-0.668	*	0.186	*

ตาราง 21 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 40 ข้อ
กลุ่มตัวอย่าง 200 คน

ข้อที่	b	θ	SE_b	SE_θ
1	-0.943	-4.23	0.169	1.09
2	-1.338	-3.40	0.181	0.78
3	-0.666	-2.90	0.164	0.64
4	-0.277	-2.54	0.159	0.57
5	0.125	-2.25	0.159	0.51
6	1.191	-2.00	0.177	0.48
7	-0.914	-1.79	0.168	0.45
8	-0.252	-1.60	0.159	0.42
9	1.068	-1.43	0.174	0.41
10	-2.154	-1.27	0.219	0.39
11	0.864	-1.12	0.168	0.38
12	-1.119	-0.98	0.174	0.37
13	-1.542	-0.84	0.188	0.36
14	1.491	-0.71	0.188	0.36
15	-1.438	-0.59	0.184	0.35
16	-0.508	-0.47	0.161	0.35
17	0.303	-0.35	0.160	0.34
18	-1.404	-0.23	0.183	0.34
19	-0.277	-0.12	0.159	0.34
20	1.098	0.00	0.175	0.34
21	-0.971	0.12	0.170	0.34
22	0.752	0.23	0.166	0.34
23	-0.613	0.35	0.163	0.34
24	-0.101	0.47	0.159	0.35
25	0.125	0.59	0.159	0.35
26	-0.101	0.71	0.159	0.36
27	0.252	0.84	0.159	0.36
28	0.277	0.98	0.160	0.37
29	1.638	1.12	0.194	0.38
30	0.697	1.27	0.165	0.39
31	-0.482	1.43	0.161	0.41
32	0.643	1.60	0.164	0.42

ตาราง 20 (ต่อ)

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
33	0.431	1.79	0.161	0.45
34	1.320	2.00	0.182	0.48
35	-0.328	2.25	0.160	0.51
36	1.129	2.54	0.175	0.57
37	2.010	2.90	0.213	0.64
38	0.670	3.40	0.165	0.78
39	-0.252	4.23	0.159	1.09
40	-0.405	*	0.160	*

ตาราง 22 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 40 ข้อ
กลุ่มตัวอย่าง 250 คน

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
1	-0.836	-4.14	0.147	1.08
2	-1.351	-3.33	0.160	0.77
3	-0.729	-2.84	0.146	0.64
4	-0.460	-2.48	0.143	0.56
5	-0.000	-2.20	0.141	0.51
6	1.323	-1.96	0.164	0.47
7	-0.969	-1.75	0.150	0.44
8	-0.562	-1.57	0.144	0.42
9	0.591	-1.40	0.146	0.40
10	-1.351	-1.24	0.160	0.39
11	0.677	-1.10	0.147	0.38
12	-1.014	-0.96	0.151	0.37
13	-1.403	-0.83	0.161	0.36
14	1.377	-0.70	0.166	0.35
15	-1.351	-0.58	0.160	0.35
16	-0.379	-0.46	0.142	0.34
17	0.159	-0.34	0.142	0.34
18	-1.276	-0.23	0.158	0.34
19	-0.339	-0.11	0.142	0.34
20	0.947	0.00	0.153	0.34
21	-1.083	0.11	0.153	0.34
22	1.043	0.23	0.156	0.34
23	-0.439	0.40	0.142	0.34
24	-0.179	0.46	0.141	0.34
25	0.280	0.58	0.142	0.35
26	-0.119	0.70	0.141	0.35
27	0.300	0.83	0.143	0.36
28	0.341	0.96	0.143	0.37
29	1.433	1.10	0.168	0.38
30	0.878	1.24	0.151	0.39
31	-0.159	1.40	0.141	0.40
32	0.548	1.57	0.145	0.42

ตาราง 22 (ต่อ)

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
33	0.300	1.75	0.143	0.44
34	1.377	1.96	0.166	0.47
35	-0.259	2.20	0.141	0.51
36	0.971	2.48	0.154	0.56
37	1.869	2.84	0.187	0.64
38	0.382	3.33	0.143	0.77
39	-0.219	4.14	0.141	1.08
40	-0.319	*	0.142	*

ตาราง 23 ค่าความยากของแบบทดสอบและค่าความสามารถของผู้สอบของแบบทดสอบ 40 ข้อ
กลุ่มตัวอย่าง 300 คน

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
1	-0.668	-4.18	0.131	1.08
2	-1.244	-3.36	0.143	0.77
3	-0.755	-2.86	0.133	0.77
4	-0.300	-2.50	0.128	0.64
5	0.074	-2.22	0.27	0.56
6	1.345	-1.98	0.149	0.51
7	-0.953	-1.77	0.136	0.47
8	-0.448	-1.58	0.129	0.44
9	0.841	-1.41	0.136	0.42
10	-1.871	-1.25	0.164	0.40
11	0.750	-1.11	0.134	0.39
12	-1.086	-0.97	0.139	0.38
13	-1.435	-0.83	0.148	0.37
14	1.506	-0.71	0.154	0.36
15	-1.435	-0.58	0.148	0.35
16	-0.398	-0.46	0.129	0.35
17	0.090	-0.34	0.128	0.34
18	-1.184	-0.23	0.141	0.34
19	-0.382	-0.11	0.128	0.34
20	0.993	0.00	0.139	0.34
21	-1.204	0.11	0.142	0.34
22	0.954	0.23	0.138	0.34
23	-0.482	0.34	0.129	0.34
24	-0.202	0.46	0.128	0.34
25	0.237	0.58	0.128	0.35
26	-0.153	0.71	0.127	0.35
27	0.237	0.83	0.128	0.36
28	0.270	0.97	0.128	0.37
29	1.435	1.11	0.152	0.38
30	0.768	1.25	0.135	0.39
31	-0.283	1.41	0.128	0.40
32	0.643	1.58	0.133	0.42

ตาราง 23 (ต่อ)

ข้อที่	b	θ	$\overline{SE}_b$	$\overline{SE}_\theta$
33	0.320	1.77	0.129	0.44
34	1.236	1.98	0.146	0.47
35	-0.349	2.22	0.128	0.51
36	1.153	2.50	0.143	0.56
37	1.929	2.86	0.172	0.64
38	0.488	3.36	0.130	0.77
39	-0.072	4.18	0.127	1.08
40	-0.366	*	0.128	*

ประวัติย่อผู้วิจัย

ประวัติย่อของผู้วิจัย

ชื่อ นางสาวเกศมณี พยัคฆ์	
เกิดวันที่ 19 เดือน มีนาคม	พุทธศักราช 2516
สถานที่เกิด	ปราจีนบุรี
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 95/1 หมู่ 1 ถนนปราจีนตะคาม ตำบลโนนหอม อำเภอเมือง จังหวัดปราจีนบุรี
ตำแหน่งหน้าที่การงานปัจจุบัน	อาจารย์ 1 ระดับ 4
สถานที่ทำงานปัจจุบัน	โรงเรียนบ้านวังดินสอ อำเภอกบินทร์บุรี สำนักงานการประถมศึกษาจังหวัดปราจีนบุรี
ประวัติการศึกษา	
พ.ศ. 2534	ชั้นมัธยมศึกษาตอนปลาย จากโรงเรียนปราจีนกัลยาณี
พ.ศ. 2538	กศ.บ. (วิชาเอกการประถมศึกษา, วิชาโทการวัดผลการศึกษา) จากมหาวิทยาลัยบูรพา
พ.ศ. 2543	กศ.ม. (วิชาเอกการวัดผลการศึกษา) จากมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร