

371.26013

พ 281 ก

ร. 2

การศึกษาประสิทธิภาพสัมพัทธ์ และคุณภาพของแบบทดสอบแบบเลือกตอบ
ที่มีจำนวนข้อ และจำนวนตัวเลือกต่างกันด้วยโมเดลโลจิสติกสามพารามิเตอร์

23 ส.ค. 2539

ปริญญาานิพนธ์

ของ

พรศรี จิมแก้ว

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

ตุลาคม 2538

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

197032

การศึกษาประสิทธิภาพสัมพัทธ์และคุณภาพของแบบทดสอบแบบเลือกตอบ
ที่มีจำนวนข้อ และจำนวนตัวเลือกต่างกัน ด้วยโมเดลโลจิสติก
สามพารามิเตอร์

บทคัดย่อ
ของ
พรศรี นิยมแก้ว

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
ตุลาคม 2538

การวิจัยครั้งนี้ มีจุดมุ่งหมายเพื่อ ศึกษาค่าประสิทธิภาพสัมพัทธ์ และคุณภาพของ แบบทดสอบแบบเลือกตอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกัน โดยศึกษากับกลุ่มตัวอย่าง นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2537 ของโรงเรียนสังกัดสำนัก งานคณะกรรมการการ ประถมศึกษาแห่งชาติ จังหวัดสมุทรสาคร จำนวน 3,016 คน ซึ่งได้มาจากการสุ่มแบบหลายชั้น ตอน เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ที่ผู้วิจัย สร้างขึ้นเอง ซึ่งมี 3 ฉบับคือ แบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก การประมาณค่าความยากของข้อสอบ ค่าอำนาจ จำแนก ค่าสัมประสิทธิ์การเดา และค่าความสามารถของผู้สอบใช้โปรแกรมไบล๊อค ตามทฤษฎีการ คอบสนองข้อสอบ แบบโลจิสติกโมเดล ด้วยวิธีการประมาณค่าความเป็นไปได้สูงสุดแบบไม่มี เงื่อนไข

ผลการวิจัยสรุปได้ดังนี้

1. ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ ระหว่างแบบทดสอบฉบับ ที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือกมีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มี ความสามารถระดับปานกลางถึงระดับสูงมาก และมีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถ ตั้งแต่ว่าระดับต่ำลงมา ส่วนค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถระหว่างแบบ ทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก และฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือกเทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถ ระดับต่ำลงมาและระดับสูงขึ้นไป แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง
2. แบบทดสอบแบบเลือกตอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกัน มีค่าความยาก ของข้อสอบและค่าอำนาจจำแนก แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แสดงว่าการลดจำนวน ข้อและจำนวนตัวเลือก โดยวิธีการตัดตัวลงที่มีค่าอำนาจจำแนกต่ำสุดออกด้วยเทคนิค 27 % และ ตารางสำเร็จรูปของจุดเฉฟานตามแนวทฤษฎีการทดสอบแบบคลาสสิคอลล ไม่ส่งผลให้ค่าความยาก ของข้อสอบและค่าอำนาจจำแนกของข้อสอบแตกต่างกัน ส่วนค่าสัมประสิทธิ์การเดาของข้อสอบ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยแบบทดสอบฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าสัมประสิทธิ์การเดาของข้อสอบต่ำที่สุด และจะเพิ่มขึ้นเรื่อย ๆ เมื่อจำนวนข้อเพิ่มขึ้นและ จำนวนตัวเลือกลดลงเป็น 44 ข้อ 4 ตัวเลือก และ 59 ข้อ 3 ตัวเลือก ตามลำดับ และการศึกษา ฟังก์ชันอินฟอर्मชันของแบบทดสอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกันพบว่า แบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีฟังก์ชันอินฟอर्मชันของแบบทดสอบสูงสุด ซึ่งเหมาะสำหรับวัด

นักเรียนที่มีระดับความสามารถประมาณ 1.60 ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีฟังก์ชันอินฟอร์เมชันของแบบทดสอบรองลงมา ซึ่งเหมาะสำหรับวัดนักเรียนที่มีระดับความสามารถ 1.40 ฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีฟังก์ชันอินฟอร์เมชันของแบบทดสอบต่ำสุด ซึ่งเหมาะสำหรับวัดนักเรียนที่มีระดับความสามารถประมาณ 1.20 แสดงว่า แบบทดสอบที่มีจำนวนข้อเพิ่มขึ้นจะให้ฟังก์ชันอินฟอร์เมชันของแบบทดสอบเพิ่มขึ้นด้วย ซึ่งฟังก์ชันอินฟอร์เมชันนี้เป็นความเชื่อมั่นในการวัดของแบบทดสอบแต่ละฉบับ ณ ระดับความสามารถที่เหมาะสมที่สุดของแบบทดสอบฉบับนั้น

A STUDY OF THE RELATION EFFICIENCY AND THE QUALITY OF MULTIPLE -
CHOICE TESTS WITH DIFFERENT NUMBERS OF ITEMS AND OPTIONS BY
THREE - PARAMETER LOGISTIC MODEL

AN ABSTRACT
BY
PORN SRI CHIM KOEW

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

October 1995

The purpose of this research was to study relative efficiency and the quality of multiple - choice tests with different numbers of items and options. A sample of Prathomsuksa 6 students of the academic year 1994 in Samusakorn was selected by multistage random sampling technique. The research instruments used were mathematics achievement tests constructed by the researcher. There test formats were the first test 59 items 3 choices, the second test 44 items 4 choices and the third test 35 items 5 choices tests. Unconditional Maximum Likelihood Procedure was used to estimate items difficulty, item discriminating, pseudo chance level and examinus' ability for Logistic Model in Item Response Theory by the BILOG computer program.

The results of the research were Summarized as follows.

1. The relative efficiency of ability estimation between the first test 59 items 3 choices and the second test 44 items 4 choices tests had morn than 1 at the moderate, high and very high ability levels but less than 1 at the very low and low ability levels. For the relative efficiency of ability estimation between the first test 59 items 3 choices and the third test 35 items 5 choices, the second test 44 items 4 choices and the third test 35 items 5 choices tests had more than 1 at the very low, low, high and very high ability levels but less than 1 at the moderate.

2. The multiple - choice tests with different numbers of items and options was no significant difference between the test for difficulty and discrimination. The study indicate that reduction in the number of items and choices, by omitting the items and choices which had the lowest value of discrimination through the use of 27 present technique and the Chung Teh Fan's Table references according the the Classical Test Theory had no effect on the value of difficulty and that of the discrimination. Pseudo-chance level differences were found significant at the .05 level of confidence. The test which had the third test 35 items 5 choices had the lowest pseudo-chance level. The pseudo-chance level increased as the number of items increased and choice decreased was 44 items 4 choices and 59 items 3 choices. The test in which there were 59 items 3 choices had the highest value of test information function which was suitable for students that had ability scale for about 1.60. The test in which there were 44 items 4 choices had the value of test information function less than the test information function of the one with 59 items 3 choices. It was suitable for

students who had ability scale for about 1.40. The test in which there were 35 items 5 choices had the lowest value of the test information function and was suitable for students who had ability scale for about 1.20. The results showed that a great number of item increased the value of the test information function which was the reliability of test at the most suitable ability scale of that test.

คณะกรรมการควบคุมและคณะกรรมการสอบ ได้พิจารณาปฏิญานีพนธ์ฉบับนี้แล้ว
เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร ปฏิญญาการศึกษามหาบัณฑิต วิชาเอก
การวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

คณะกรรมการควบคุม

.....
..... ประธาน
(อาจารย์ ดร.วัน สังข์สะอาด)

.....
..... กรรมการ
(อาจารย์ ดร.ชูศักดิ์ ชัมภลลิขิต)

คณะกรรมการสอบ

.....
..... ประธาน
(อาจารย์ ดร.วัน สังข์สะอาด)

.....
..... กรรมการ
(อาจารย์ ดร.ชูศักดิ์ ชัมภลลิขิต)

.....
..... กรรมการที่แต่งตั้งเพิ่มเติม
(อาจารย์ ยุพา มานะจิตต์)

บัณฑิตวิทยาลัยอนุมัติให้รับปฏิญานีพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตาม
หลักสูตรปฏิญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัย
ศรีนครินทรวิโรฒ

.....
..... คณบดีบัณฑิตวิทยาลัย
(ดร.ศิริยุภา พูลสุวรรณ)

วันที่ 11 เดือน ตุลาคม พ.ศ. 2538

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จได้ด้วยความกรุณาอย่างยิ่งจาก อาจารย์ดร.วัน
สังข์สะอาด ที่เสียสละเวลาอันมีค่าในการให้คำแนะนำ ชี้แนะแนวทางและแก้ไขข้อบกพร่องต่าง ๆ
จนงานวิจัยนี้สำเร็จลุล่วงไปด้วยดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณ อาจารย์ดร.ชูศักดิ์ ชัมภลลิขิต ที่กรุณาให้คำปรึกษาแนะนำใน
ส่วนของการวิเคราะห์ข้อมูลด้วยความเอาใจใส่เป็นอย่างดี

ขอกราบขอบพระคุณอาจารย์ยุพา มานะจิตต์ ที่ได้กรุณาให้ข้อเสนอแนะเพิ่มเติม
เพื่อให้ปริญญานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ผู้วิจัยขอขอบพระคุณ คุณสุรศักดิ์ จินตรัตน์ คุณฉวีวรรณ เอื้อวณิชกุล และคุณรุ่งฤดี
ทรัพย์นิธิ ที่กรุณาให้ความอนุเคราะห์ในการวิเคราะห์ข้อมูลด้วยเครื่องคอมพิวเตอร์

ผู้วิจัยขอขอบพระคุณ คณะครูโรงเรียนในกลุ่มตัวอย่าง และขอขอบใจนักเรียนทุกคนที่
ได้ให้ความร่วมมือในการเก็บข้อมูลเป็นอย่างดี

ขอขอบคุณ คุณสาธิตา เมธนาวิน คุณวีระยุทธ ศิริรัตน์ ตลอดจนเจ้านี่ ๆ เพื่อน ๆ น้อง ๆ
ทุกคนที่เป็นกำลังใจและให้ความช่วยเหลือในการทำปริญญานิพนธ์ฉบับนี้

ผู้วิจัยขอขอบคุณผู้มีคุณค่าของงานวิจัยนี้ และความดีทั้งมวลให้แก่คุณสามารภ
จิมแก้ว และเด็กชายน้ามนต์ จิมแก้ว และขอโน้มรำลึกถึงสิ่งศักดิ์สิทธิ์ทั้งหลาย พระคุณของพ่อ
แม่ ครู อาจารย์ ทุกคนที่เป็นกำลังใจแก่ผู้วิจัยตลอดมา

พรศรี จิมแก้ว

สารบัญ

บทที่	หน้า
1 บทนำ	
ภูมิหลัง	1
ความมุ่งหมายของการศึกษาค้นคว้า	3
ความสำคัญของการศึกษาค้นคว้า	4
ขอบเขตของการศึกษาค้นคว้า	4
คำนิยามศัพท์เฉพาะ	5
2 เอกสารและงานวิจัยที่เกี่ยวข้อง	8
ความเป็นมาของทฤษฎีการตอบข้อคำถาม	8
หลักการของทฤษฎีการตอบข้อคำถาม	9
ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ	17
งานวิจัยที่เกี่ยวข้องภายในประเทศ	23
งานวิจัยที่เกี่ยวข้องต่างประเทศ	28
สมมติฐานในการศึกษาค้นคว้า	30
3 วิธีดำเนินการศึกษาค้นคว้า	31
ประชากร	31
กลุ่มตัวอย่าง	31
เครื่องมือที่ใช้ในการศึกษาค้นคว้า	34
วิธีดำเนินการสร้างแบบทดสอบ	34
วิธีดำเนินการเก็บรวบรวมข้อมูล	37
สถิติที่ใช้ในการวิเคราะห์ข้อมูล	40
4 ผลการวิเคราะห์ข้อมูล	43

บทที่	หน้า
5 สรุป อภิปรายผล และข้อเสนอแนะ	57
ความมุ่งหมายของการศึกษาค้นคว้า	57
กลุ่มตัวอย่าง	57
เครื่องมือที่ใช้ในการศึกษาค้นคว้า	57
วิธีดำเนินการเก็บรวบรวมข้อมูล	58
สรุปผลการวิเคราะห์ข้อมูล	58
อภิปรายผล	61
ข้อเสนอแนะ	65
บรรณานุกรม	66
ภาคผนวก	70
ประวัติย่อของผู้วิจัย	111

บัญชีตาราง

ตาราง		หน้า
1	โรงเรียน อำเภอ และจำนวนนักเรียนที่เป็นกลุ่มตัวอย่างของการวิจัย	32
2	ค่าสถิติพื้นฐานของความยาก อำนาจจำแนก และสัมประสิทธิ์ การเดาของ แบบทดสอบทั้ง 3 ฉบับ.....	44
3	ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบ ทดสอบที่แตกต่างกัน 3 ฉบับ ที่ระดับความสามารถ ตั้งแต่ -4.00 ถึง 4.00 จำนวน 81 ระดับความสามารถ	48
4	การทดสอบความแตกต่างของค่าความยากของแบบทดสอบทั้ง 3 ฉบับ.....	54
5	การทดสอบความแตกต่างของค่าอำนาจจำแนกของแบบทดสอบทั้ง 3 ฉบับ.....	54
6	การทดสอบความแตกต่างของค่าสัมประสิทธิ์การเดาของแบบทดสอบทั้ง 3 ฉบับ.....	55
7	การทดสอบความแตกต่างรายคู่ด้วยการเปรียบเทียบพหุคูณ	55
8	ค่าไอเกน (มากกว่า 1) เปอร์เซนต์ของความแปรปรวน และเปอร์เซนต์ของความ แปรปรวนสะสมจากการนำแบบทดสอบ จำนวน 3 ฉบับมาวิเคราะห์องค์ประกอบ	72
9	ค่าพารามิเตอร์ของแบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก	75
10	ค่าพารามิเตอร์ของแบบทดสอบ ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก	77
11	ค่าพารามิเตอร์ของแบบทดสอบ ฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก	79
12	ค่าฟังก์ชันอินฟอर्मชันของแบบทดสอบ ระหว่างฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ฉบับ ที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก ที่ระดับความ สามารถตั้งแต่ -4.00 ถึง 4.00 จำนวน 81ระดับความสามารถ	82
13	ค่าประสิทธิภาพสัมพัทธ์ ของการประมาณค่าความสามารถของผู้สอบระหว่าง แบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และ ฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก ที่ระดับความสามารถตั้งแต่ -4.00 ถึง4.00 จำนวน 81 ระดับความสามารถ	87
14	ความน่าจะเป็นในการตอบแบบทดสอบได้ถูก ของแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก จำนวน 81 ระดับความสามารถ	92
15	ค่าความยากค่าอำนาจจำแนกและค่าความยากมาตรฐานของเครื่องมือ	95

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 โค้งอินฟอร์เมชันของข้อสอบ 5 ข้อ และโค้งอินฟอร์เมชันของแบบทดสอบ	16
2 โค้งอินฟอร์เมชันของแบบทดสอบ ที่มีประสิทธิภาพสำหรับการวัด ระดับความ สามารถต่ำ ($\theta = 2.3$)	19
3 โค้งอินฟอร์เมชันของแบบทดสอบ ที่มีประสิทธิภาพสำหรับการวัด ทั้งระดับความ สามารถสูงและต่ำ	19
4 โค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ของภาพประกอบ 3 เทียบกับ แบบทดสอบของภาพประกอบ 2	20
5 ลำดับขั้นการสร้างเครื่องมือที่ใช้ในการวิจัย	35
6 ค่าไอเกนขององค์ประกอบของแบบทดสอบ ทั้ง 3 ฉบับ	39
7 โค้งลักษณะแบบทดสอบ และโค้งอินฟอร์เมชันของแบบทดสอบ ทั้ง 3 ฉบับ	46
8 การเปรียบเทียบโค้งอินฟอร์เมชันของแบบทดสอบทั้ง 3 ฉบับ	47
9 โค้งประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบระหว่าง แบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก	51
10 โค้งประสิทธิภาพสัมพัทธ์ ของการประมาณค่าความสามารถของผู้สอบ ระหว่าง แบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก	52
11 โค้งประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ ระหว่าง แบบทดสอบฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก เทียบกับ ฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก.	53

บทที่ 1

บทนำ

ภูมิหลัง

การวัดและประเมินผลทางการศึกษาเป็นสิ่งจำเป็นควบคู่ไปกับการเรียนการสอน เพราะการวัดผลและประเมินผลจะทำให้ครู-อาจารย์ หรือผู้ที่เกี่ยวข้องกับ การศึกษาทราบว่าผลการเรียนการสอนที่ผ่านมานั้น ประสบผลสำเร็จเป็นไปตามจุดมุ่งหมายของการเรียนการสอนหรือไม่เพียงใด ผลจากการวัดจะอยู่ในรูปของคะแนน ซึ่งจะนำไปใช้ในการอธิบายหรือทำนายความสามารถที่แท้จริงของผู้ที่ถูกวัด ซึ่งจะถูกต้องเชื่อถือได้ และมีประสิทธิภาพเพียงใด ย่อมขึ้นอยู่กับคุณภาพของแบบทดสอบที่ใช้เป็นเครื่องมือในการวัดนั้น (อวยพร วิบูลย์. 2526 : 7)

สำหรับคุณภาพของแบบทดสอบที่ดีนั้น นักวัดผลและนักการศึกษาได้กล่าวว่า มี 3 ประการ คือความเที่ยงตรง ความเชื่อมั่น และสามารถนำไปใช้ได้ (Lyman. 1963 : 25; Chase. 1974 : 56) มีบางท่านกล่าวเน้นว่า คุณสมบัติที่มีความสำคัญที่สุดของเครื่องมือวัดผลการศึกษา ได้แก่ ความเชื่อมั่น และความเที่ยงตรงของแบบทดสอบ (Ebel. 1965 : 60; บุญเจิด ภิญโญอนันตพงษ์. 2525 : 128; อนันต์ ศรีโสภา. 2525 : 59) และแบบทดสอบที่เป็นที่นิยมใช้มากที่สุดคือ แบบทดสอบปรนัย ซึ่งเป็นแบบทดสอบที่มีคำถามเฉพาะเจาะจง สามารถระบุได้ค่อนข้างชัดเจนว่าในการตอบข้อสอบแต่ละข้อนั้น ผู้ตอบจะต้องใช้ความสามารถด้านใด การตอบแต่ละข้อใช้เวลาไม่มากนักจึงใช้ข้อสอบได้มากข้อได้ และทำให้สามารถถามได้ครอบคลุมเนื้อหาที่ต้องการวัดได้มาก นอกจากนั้นการตรวจให้คะแนนมีความเป็นปรนัยสูง ทำได้รวดเร็ว สามารถรายงานผลการสอบได้ทันกับความต้องการใช้ แม้จะมีผู้เข้าสอบเป็นจำนวนมากก็ตาม

แบบทดสอบปรนัยมีหลายชนิดที่นิยมใช้คือแบบเลือกตอบ ทั้งนี้เพราะเป็นแบบทดสอบที่มีความสามารถในการจำแนกเด็กเป็นกลุ่มเก่ง และกลุ่มอ่อนได้ดี แม้จะสร้างยากแต่ก็มีความเชื่อมั่นสูงกว่าแบบทดสอบปรนัยชนิดต่าง ๆ ทั้งหมด สำหรับความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบปรนัยชนิดเลือกตอบนั้น ขึ้นอยู่กับความยาวของแบบทดสอบ ความยากง่ายของแบบทดสอบ และจำนวนตัวเลือก แบบทดสอบแบบเลือกตอบแต่ละข้อควรมี 4 ตัวเลือกหรือ 5 ตัวเลือก เพราะยังมีตัวเลือกมากโอกาสในการเดาคำตอบที่ถูกต้องยังมีน้อยลง คือถ้ามีตัวเลือก 5 ตัวเลือก โอกาสในการเดาคำตอบที่ถูกต้องมี 1 ใน 5 (Gronlund. 1982 : 26) ถ้ามี 4 ตัวเลือกก็เป็น 1 ใน 4 ถ้ามี 3 ตัวเลือก ก็เป็น 1 ใน 3 ในกรณี 3 ตัวเลือก ถ้าบังเอิญมีตัวลวงบกพร่องหนึ่งตัว ก็กลายเป็น

แบบถูกผิด แบบทดสอบที่มีจำนวนข้อมาก หรือจำนวนตัวเลือกมากนั้น จะสามารถวัดพฤติกรรม ได้มากพอเพียง และจะได้คะแนนที่เกิดจากการเดาได้ถูกน้อยลง ความแปรปรวนในการวัดจึงมาก ค่าความเชื่อมั่นสูงขึ้น ความยากง่ายของแบบทดสอบที่มีเกณฑ์ง่ายหรือยากเกินไปจะทำให้ค่าความ เชื่อมั่นต่ำเพราะการกระจายของคะแนนมีน้อย ถ้าข้อสอบง่ายกลุ่มของคะแนนจะอยู่ช่วงบนและถ้า ข้อสอบยาก กลุ่มของคะแนนจะอยู่ช่วงล่างของสเกล ความแปรปรวนในการวัดน้อย ค่าความเชื่อ มั่นจึงต่ำ ได้มีผู้ศึกษาเกี่ยวกับ จำนวนข้อและจำนวนตัวเลือกของแบบทดสอบ ที่จะทำให้แบบ ทศสอบมีคุณภาพสูงสุด เช่น เคนดอลล์ (Kendall. 1964 : 789-798) เรมเมอร์ และ เอดคินส์ (Remmers and Adkins. 1952 : 385- 390) ซิมเมอร์แมน และฮัมฟรีย์ (Zimmermand and Humphrey. 1953 : 460- 461) โดยเปรียบเทียบแบบทดสอบที่มีจำนวนข้อเท่ากันแต่มีจำนวนตัว เลือกแตกต่างกัน และเวลาที่ใช้ในการทำแบบทดสอบเท่ากัน ปรากฏว่าแบบทดสอบที่มีจำนวน ตัวเลือกน้อย จะมีค่าความเชื่อมั่นต่ำกว่าแบบทดสอบที่มีจำนวนตัวเลือกมาก ทั้งนี้เพราะแบบ ทศสอบที่มีจำนวนตัวเลือกน้อย เด็กจะมีเวลาในการพิจารณาคำตอบมาก คะแนนที่ได้จึงไม่แตก ต่างกัน ความแปรปรวนในการวัดจึงมีน้อย เป็นผลให้ค่าความเชื่อมั่นต่ำ ทเวทสกี (Tversky. 1964 : 386-391) จึงได้เสนอแนะว่า ข้อสอบที่จะนำมาเปรียบเทียบคุณภาพได้คือนั้น ควรมีขนาด เท่ากันคือ มีผลคูณระหว่างจำนวนข้อและจำนวนตัวเลือกเท่ากัน เช่น ข้อสอบ 30 ข้อ 2 ตัวเลือก ควรเปรียบเทียบกับข้อสอบ 20 ข้อ 3 ตัวเลือก เป็นต้น

การศึกษาวิจัยในเรื่องนี้ของไทย ธรรม หาญตระกูล (2519 : 1-43) ได้ศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบแบบเลือกตอบ ที่มีจำนวนข้อเท่ากัน แต่จำนวนตัวเลือกไม่ เท่ากัน พบว่า แบบทดสอบ 4 ตัวเลือกมีค่าความเชื่อมั่นสูงกว่า แบบทดสอบที่มี 2, 3 และ 5 ตัว เลือก ส่วน ชนะ ทานะวงศ์ (2521 : 1-41) ได้ศึกษาพบว่า แบบทดสอบ 2 และ 3 ตัวเลือก ให้ค่า ความเชื่อมั่นสูงกว่า แบบทดสอบ 4 และ 5 ตัวเลือก และ วิภาดา วัฒนกุลกิตติ์ (2529 : 1- 53) พบว่า การลดจำนวนตัวเลือก โดยการตัดตัวลงที่มีค่าอำนาจจำแนกต่ำสุดออก จะทำให้ค่าความ เชื่อมั่นของแบบทดสอบทุกฉบับแตกต่างกันอย่างไม่มีนัยสำคัญและค่าความยากของแบบทดสอบ 3 ตัวเลือก มีค่ามากกว่า 4 และ 5 ตัวเลือก

จากการศึกษาเอกสารและงานวิจัยคุณภาพของแบบทดสอบแบบเลือกตอบเกี่ยวกับ จำนวนตัวเลือก และจำนวนข้อสอบที่ไม่เท่ากัน พบว่า การวิจัยได้มีข้อขัดแย้งกันเกี่ยวกับค่าความ เชื่อมั่น ค่าความยาก และค่าอำนาจจำแนก ซึ่งเป็นคุณลักษณะที่สำคัญของการศึกษาคุณภาพ ของแบบทดสอบ และการศึกษาวิจัยเหล่านี้ ได้อาศัยแนวทฤษฎีการทดสอบ แบบคลาสสิคอล (Classical Test Theory) ซึ่งในปัจจุบันการทดสอบตามแนวทฤษฎีคลาสสิคอลได้ถูกมองว่า มีจุด อ่อนหลายประการ โดยเฉพาะเรื่องค่าพารามิเตอร์ที่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง ดังนั้นไม่

สามารถที่จะสรุปคุณภาพของแบบทดสอบที่ชัดเจนและสมบูรณ์ได้

ในช่วงปี ค.ศ.1970 และ 1980 ทฤษฎีการตอบข้อคำถาม (Item Response Theory) ได้รับความสนใจจากนักวัดผลทางจิตวิทยาเป็นอย่างมาก เนื่องจากสามารถแก้จุดอ่อนของทฤษฎีคลาสสิกอลได้ โดยค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก ค่าอำนาจจำแนกและค่าสัมประสิทธิ์การเดา จะเป็นค่าที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ไม่ว่าจะนำไปสอบกับผู้ใด และเมื่อทราบลักษณะการตอบข้อสอบแต่ละข้อของผู้สอบคนใด เราสามารถใช้ข้อสอบข้อใดก็ได้หรือฉบับใดก็ได้ ที่วัดสิ่งเดียวกันในการคำนวณหาความสามารถของบุคคลนั้นได้ การวิจัยที่เกี่ยวกับการหาคุณภาพของแบบทดสอบที่มีจำนวนตัวเลือกต่างกัน ในทฤษฎีการตอบข้อคำถามนั้น ชาดิซาช กูกิตติโมตรี (2533 : บทคัดย่อ) พบว่า แบบทดสอบแบบเลือกตอบที่มีจำนวนตัวเลือกต่างกัน มีค่าความยากของข้อสอบ และค่าอำนาจจำแนกของข้อสอบไม่ต่างกัน แบบทดสอบชนิด 5 ตัวเลือก มีค่าสัมประสิทธิ์การเดาของข้อสอบต่ำที่สุดและมีค่าฟังก์ชันอินฟอเมชันสูงสุด

จากเหตุผลดังกล่าวข้างต้น ทำให้ผู้วิจัยมีความสนใจที่จะทำการศึกษาเกี่ยวกับจำนวนข้อสอบ และจำนวนตัวเลือกของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ โดยใช้ทฤษฎีการตอบข้อคำถามเป็นแนวทางการศึกษา โดยศึกษาค่าประสิทธิภาพสัมพัทธ์และคุณภาพของแบบทดสอบเลือกตอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกัน 3 แบบคือ แบบ 60 ข้อ 3 ตัวเลือก แบบ 45 ข้อ 4 ตัวเลือก แบบ 36 ข้อ 5 ตัวเลือก ว่าจะแตกต่างกันหรือไม่ ซึ่งผลการวิจัยครั้งนี้จะเป็นประโยชน์ต่อครูได้ใช้เป็นแนวทางปฏิบัติเกี่ยวกับการสร้างข้อสอบ เพราะแบบทดสอบที่ใช้วัดผลสัมฤทธิ์ทางการเรียนนั้น ส่วนมากเป็นแบบทดสอบที่ครูสร้างขึ้นเอง ดังนั้นในการสร้างแบบทดสอบจึงควรรู้ว่าควรมีจำนวนข้อและจำนวนตัวเลือกเท่าไรที่จะทำให้แบบทดสอบมีคุณภาพมากที่สุด และแบบทดสอบมีจำนวนข้อและจำนวนตัวเลือกเท่าใด ที่จะเหมาะสมกับความสามารถของเด็กที่มีระดับความสามารถต่างกันมากที่สุด ทั้งนี้เพื่อเป็นประโยชน์ต่อการสร้างแบบทดสอบแบบเลือกตอบเพื่อวัดผลสัมฤทธิ์ของการเรียนในโรงเรียนอย่างมีคุณภาพต่อไป

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อศึกษาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบแบบเลือกตอบที่มีจำนวนข้อ และจำนวนตัวเลือกต่างกัน
2. เพื่อศึกษาค่าความยากของข้อสอบแบบเลือกตอบที่มีจำนวนข้อ และจำนวนตัวเลือกต่างกัน
3. เพื่อศึกษาค่าอำนาจจำแนกของข้อสอบแบบเลือกตอบที่มีจำนวนข้อ และ จำนวนตัว

เลือกต่างกัน

4. เพื่อศึกษาค่าสัมประสิทธิ์การแจกของข้อสอบแบบเลือกตอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกัน

5. เพื่อศึกษาฟังก์ชันอินฟอร์เมชันของแบบทดสอบแบบเลือกตอบที่มีจำนวนข้อ และจำนวนตัวเลือกต่างกัน

ความสำคัญของการศึกษาค้นคว้า

ผลจากการศึกษาค้นคว้าครั้งนี้ จะทำให้ทราบว่า แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบที่มีจำนวนข้อ และจำนวนตัวเลือกต่างกัน 3 แบบ แบบใดมีคุณภาพอย่างไร และเป็นแนวทางในการนำทฤษฎีการตอบข้อคำถามไปทดสอบในวิชาอื่น ๆ นอกจากวิชาคณิตศาสตร์ต่อไป

ขอบเขตของการศึกษาค้นคว้า

1. ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษา ปีที่ 6 ภาคเรียนที่ 1 ปีการศึกษา 2537 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ ในจังหวัดสมุทรสาคร จำนวน 112 โรงเรียน นักเรียน 5,246 คน

2. กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษา ปีที่ 6 ภาคเรียนที่ 1 ปีการศึกษา 2537 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ ในจังหวัดสมุทรสาคร จำนวน 55 โรงเรียน นักเรียน 3,016 คน

3. ตัวแปรที่ใช้ในการศึกษาค้นคว้า

3.1 ตัวแปรอิสระ คือ รูปแบบของข้อสอบในแบบทดสอบ ที่มีขนาด 180ตัวเลือกเท่ากัน มี 3 รูปแบบ คือ

แบบที่ 1. จำนวนข้อสอบ 60 ข้อ 3 ตัวเลือก

แบบที่ 2 จำนวนข้อสอบ 45 ข้อ 4 ตัวเลือก

แบบที่ 3 จำนวนข้อสอบ 36 ข้อ 5 ตัวเลือก

3.2 ตัวแปรตาม คือ

3.2.1 ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ

3.2.2 ค่าความยากของข้อสอบ

3.2.3 ค่าอำนาจจำแนกของข้อสอบ

3.2.4 ค่าสัมประสิทธิ์การเดาของข้อสอบ

3.2.5 ฟังก์ชันอินฟอร์เมชันของแบบทดสอบ

4. เนื้อหาที่ใช้ในการสร้างเครื่องมือ เป็นเนื้อหาวิชาคณิตศาสตร์ ประถมศึกษาปีที่ 6 เรื่อง ตัวประกอบของจำนวนนับ และเศษส่วนซึ่งสร้างตามหลักสูตรประถมศึกษาพุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ.2533)

ค่านิยามศัพท์เฉพาะ

1. แบบทดสอบแบบเลือกตอบ หมายถึง แบบทดสอบที่ประกอบด้วยตัวคำถามหรือคอนำ กับตัวเลือกหรือคำตอบ ในแต่ละข้อมีจำนวนตัวเลือกตั้งแต่ 3 ตัวขึ้นไป และมีตัวเลือกที่ถูกต้องเพียงตัวเดียว

2. แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน หมายถึง แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่อง ตัวประกอบของ จำนวนนับ และเศษส่วนที่ผู้วิจัยสร้างขึ้นเอง ตามหลักสูตรประถมศึกษา พุทธศักราช2521 (ฉบับปรับปรุง พ.ศ.2533)

3. ทฤษฎีการตอบข้อคำถาม (Item Response Theory : IRT) หมายถึง ทฤษฎีที่อธิบายถึงความสัมพันธ์ระหว่างความสามารถที่แท้จริง ที่มีอยู่ในตัวผู้สอบกับการตอบข้อคำถาม ซึ่งความสามารถที่แท้จริง คือ ลักษณะที่วัดได้ด้วยแบบทดสอบ

4. แบบทดสอบที่มีมิติเดียว (Unidimension Test) หมายถึง แบบทดสอบที่วัดความสามารถด้านใดด้านหนึ่งเพียงด้านเดียว หรือคุณลักษณะเดียวของผู้สอบ หรือแบบทดสอบที่ข้อคำถามแต่ละข้อมีความเป็นเอกพันธ์กัน ในการศึกษาครั้งนี้ ทดสอบความมีมิติเดียวโดยการวิเคราะห์องค์ประกอบ

5. โค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) หมายถึง เส้นภาพที่เกิดจากความสัมพันธ์ของค่าความสามารถที่แท้จริง กับค่าความน่าจะเป็นในการตอบข้อสอบรายข้อได้ถูกต้อง

6. ค่าพารามิเตอร์ของข้อสอบ (Item parameter) หมายถึง ค่าอำนาจจำแนกของข้อสอบ (a) ค่าความยากของข้อสอบ (b) และค่าสัมประสิทธิ์การเดาของข้อสอบ (c) การวิจัยครั้งนี้หาโดยใช้ทฤษฎีการตอบข้อคำถามซึ่งวิเคราะห์ด้วยโปรแกรมสำเร็จรูปไบล็อก (Bilog)

7. ค่าประสิทธิภาพสัมพัทธ์ (Relative Efficiency : RE) ของการประมาณค่าความสามารถ หมายถึง คำนวณถึงประสิทธิภาพของแบบทดสอบ ซึ่งคำนวณได้จากการหาอัตราส่วนระหว่างค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบแต่ละฉบับที่นำมาเปรียบเทียบกับ ณ ระดับ

ความสามารถเดียวกัน โดยใช้สูตรดังนี้ (Lord and Novick. 1968 : 472)

$$RE(\theta, X_1, X_2) = I(\theta, X_1) / I(\theta, X_2)$$

เมื่อ $RE(\theta, X_1, X_2)$ แทนค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ
ระหว่างแบบทดสอบรูปแบบที่ 1 (X_1) เทียบกับแบบทดสอบ
รูปแบบที่ 2 (X_2)

$I(\theta, X_1)$ แทนค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบรูปแบบ ที่ 1 (X_1)
ณ ระดับความสามารถ θ

$I(\theta, X_2)$ แทนค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบรูปแบบ ที่ 2 (X_2)
ณ ระดับความสามารถ θ

สำหรับค่าประสิทธิภาพสัมพัทธ์ที่คำนวณได้ นำมาแปลผลดังนี้ (warm. 1978 : 76)

เมื่อ $RE(\theta, X_1, X_2) > 1$ แสดงว่า แบบทดสอบรูปแบบที่ 1 (X_1) มีประสิทธิภาพสูงกว่า
แบบทดสอบ รูปแบบที่ 2 (X_2) ณ ระดับความสามารถ θ

$RE(\theta, X_1, X_2) < 1$ แสดงว่า แบบทดสอบรูปแบบที่ 1 (X_1) มีประสิทธิภาพต่ำกว่า
แบบทดสอบ รูปแบบที่ 2 (X_2) ณ ระดับความสามารถ θ

$RE(\theta, X_1, X_2) = 1$ แสดงว่า แบบทดสอบรูปแบบที่ 1 (X_1) มีประสิทธิภาพพอ ๆ กับ
แบบทดสอบรูปแบบที่ 2 (X_2) ณ ระดับความสามารถ θ

8. ค่าความยากของข้อสอบ (Item Difficulty : b) หมายถึง ค่าที่แสดงถึงระดับความ
สามารถที่จุดโค้งลักษณะข้อสอบมีความชันมากที่สุด โดยทั่วไปมีค่าอยู่ระหว่าง -2 ถึง 2 ค่า -2
แสดงว่าข้อสอบข้อนั้นง่ายมาก ค่า 2 แสดงว่าข้อสอบนั้นยากมาก

9. ค่าอำนาจจำแนกของข้อสอบ (Item Discriminating : a) หมายถึง ค่าที่เป็นสัดส่วน
โดยตรงกับความชันของโค้งลักษณะข้อสอบ ณ จุดเปลี่ยนโค้งหรือที่จุด $\theta = b$, โดยทั่วไปมีค่าอยู่
ระหว่าง 0 ถึง 2 ค่า 0 แสดงว่า ข้อสอบไม่มีอำนาจจำแนก ค่า 2 แสดงว่า ข้อสอบมีค่าอำนาจ
จำแนกสูง

10. ค่าสัมประสิทธิ์การเดาของข้อสอบ (Pseudo Chance Level : c) หมายถึง ความน่า
จะเป็นที่ผู้สอบสามารถตอบข้อสอบนั้นถูกโดยไม่มีความรู้ มีค่าจาก 0 ถึง 1 ในทางปฏิบัติจะคัด
เลือกข้อสอบที่มีค่าการเดาน้อยกว่า 0.30

11. ฟังก์ชันอินฟอร์เมชันของข้อสอบ (Item Information Function) หมายถึง ฟังก์ชันที่
แสดงถึงรายละเอียดของข้อสอบ ซึ่งเป็นค่าที่เป็นสัดส่วนกับกำลังสองของช่วงแห่งความเชื่อมั่น

แอสิมโทติก (Asymptotic Confidence Interval) สำหรับการประมาณค่าความสามารถของผู้สอบ จากผลการตอบของเขา การวิจัยครั้งนี้ ประมาณค่าโดยใช้สูตรของเบียร์นบอม (ผจงจิต อินทสุวรรณ. 2528 : 97, 27- 29 ; อ้างอิงมาจาก Birnbaum. 1968)

12. ฟังก์ชันอินฟอร์เมชันของแบบทดสอบ (Test Information Function) หมายถึง ฟังก์ชันที่แสดงรายละเอียดของแบบทดสอบ ซึ่งเป็นผลรวมของค่าที่เป็นสัดส่วน กับกำลังสองของ ช่วงแห่งความเชื่อมั่นแอสิมโทติก การวิจัยครั้งนี้ ประมาณค่าโดยใช้สูตรของ เบียร์นบอม (ผจงจิต อินทสุวรรณ, 2528 : 27-29 ; อ้างอิงมาจาก Birnbaum. 1968)

13. คุณภาพของแบบทดสอบแบบเลือกตอบ หมายถึง คุณลักษณะหรือพารามิเตอร์ ประจำข้อของข้อสอบ คือ ค่าความยาก (b) ค่าอำนาจจำแนก (a) ค่าสัมประสิทธิ์การเดา (c) และ ค่าไค้แสดงลักษณะของข้อสอบ (ICC)

14. ค่าความสามารถของผู้สอบ (Examinee's Ability = θ) หมายถึงระดับความสามารถ ของผู้สอบแต่ละคน ซึ่งวัดในหน่วย θ ที่มีค่าเฉลี่ยเท่ากับ 0 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 1 สำหรับค่าความสามารถมีค่าเป็นไปได้ ตั้งแต่ -2 ถึง +2 แต่ในการศึกษาครั้งนี้ ผู้วิจัยใช้ค่าความ สามารถในช่วง -4.00 ถึง 4.00 โดยแบ่งระดับความสามารถออกเป็น 5 ระดับ ดังนี้

14.1 ระดับต่ำมาก $\theta < -3.00$

14.2 ระดับต่ำ $-3.00 \leq \theta < -1.00$

14.3 ระดับปานกลาง $-1.00 \leq \theta \leq 1.00$

14.4 ระดับสูง $1.00 < \theta \leq 3.00$

14.5 ระดับสูงมาก $\theta > 3.00$

15. ขนาดของโรงเรียน แบ่งเป็น 3 ขนาด ตามเกณฑ์ของสำนักงานคณะกรรมการ ประถมศึกษาแห่งชาติ ดังนี้

15.1 โรงเรียนขนาดใหญ่ จำนวนนักเรียน ตั้งแต่ 901 คนขึ้นไป

15.2 โรงเรียนขนาดกลาง จำนวนนักเรียน ตั้งแต่ 301-900 คน

15.3 โรงเรียนขนาดเล็ก จำนวนนักเรียนน้อยกว่าหรือเท่ากับ 300 คน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้ศึกษาเนื้อหาต่าง ๆ ดังนี้

1. ความเป็นมาของทฤษฎีการตอบข้อคำถาม
2. หลักการของทฤษฎีการตอบข้อคำถาม
 - 2.1 ข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม
 - 2.2 พารามิเตอร์ในทฤษฎีการตอบข้อคำถาม
 - 2.3 โมเดลในทฤษฎีการตอบข้อคำถาม
 - 2.4 คุณภาพของแบบทดสอบของทฤษฎีการตอบข้อคำถาม
3. ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบในทฤษฎีการตอบข้อคำถาม
4. งานวิจัยที่เกี่ยวข้องภายในประเทศ
5. งานวิจัยที่เกี่ยวข้องต่างประเทศ

1. ความเป็นมาของทฤษฎีการตอบข้อคำถาม

ทฤษฎีการตอบข้อคำถาม (Item Response Theory) เป็นทฤษฎีที่ศึกษาเกี่ยวกับความสัมพันธ์ในการตอบข้อสอบกับคุณลักษณะภายในของผู้สอบ เป้าหมายหลักของทฤษฎีการตอบข้อคำถาม คือ ต้องการหาคุณภาพรายข้อของแบบทดสอบและการประมาณความสามารถของผู้สอบ ซึ่งต้องอาศัยความสัมพันธ์กันระหว่างโมเดลและข้อมูล (Baker. 1987 : 121 ; citing. Yen. 1981)

ผู้ริเริ่มให้แนวคิดและหลักการเบื้องต้นของทฤษฎีการตอบข้อคำถาม คือ เฟร็ดริกสันและเลอเลย์ (Lord and Novick. 1968:369: citing Ferguson and Lawley : 1942, 1943) โดยในปี ค.ศ.1952 ลอร์ด (Frederic M.Lord) ได้เสนอทฤษฎีในรูปโค้งลักษณะข้อสอบ (Item Characteristic Curve Theory) ซึ่งกล่าวว่าโอกาสในการตอบข้อสอบถูกหรือไม่นั้น ขึ้นอยู่กับความยากของข้อสอบกับระดับความสามารถของผู้สอบ (Traite or Ability) ซึ่งกำหนดด้วยฟังก์ชันทางคณิตศาสตร์ที่เรียกว่า เส้นโค้งลักษณะข้อสอบ (Item Characteristic Function or Item Characteristic Curve : ICC) โค้งลักษณะของข้อสอบแต่ละข้อมีลักษณะเป็นโค้งปกติสะสม หรือ

เรียกว่า โมเดลโค้ง ปกติสะสม (Normal Ogive Model) เป็นโมเดลที่กล่าวถึง พารามิเตอร์ 2 ตัว คือ ค่าความยาก และค่าอำนาจจำแนก ซึ่งมีวิธีการคำนวณที่ค่อนข้างยุ่งยาก จึงไม่เป็นที่นิยม

ในปี ค.ศ.1960 ราสช์ (Georg Rasch) ได้ ศึกษาและเสนอแนวคิดในรูปของพารามิเตอร์ เพียงตัวเดียว คือ ค่าความยาก (b) ที่เรียกว่า ราสช์โมเดล (Rasch Model) และในปี ค.ศ.1965 ลอร์ด ได้กลับมาสนใจทฤษฎีนี้ใหม่ (Warm. 1979 : 19)

ในปี ค.ศ.1968 เบิร์นบอม (Birnbaum) ได้เสนอแนวคิดของทฤษฎีในรูปของโลจิสติก โมเดล (Logistic Model) ที่ใช้พารามิเตอร์ 2 ตัว คือ ค่าความยาก (b) และค่าอำนาจจำแนก (a) ซึ่งเป็นโมเดลที่ง่ายกว่าโมเดลของลอร์ดจึงทำให้เป็นที่นิยมใช้กันอย่างแพร่หลายและมีการพัฒนาขึ้นเรื่อย ๆ จนสามารถดัดแปลงใช้กับโมเดลพารามิเตอร์ตัวเดียว โมเดลพารามิเตอร์สามตัว (Warm. 1979 : 19-21) แม็คโดนัลด์ และคนอื่น ๆ (McDonald. 1967; Barton and Lord. 1981) ได้พัฒนาโลจิสติกโมเดลใช้กับพารามิเตอร์สี่ตัว (Hambleton and Swaminathan. 48-49)

2. หลักการของทฤษฎีการตอบข้อคำถาม

ลอร์ด และโนวิก (Lord and Novick. 1968 : 258) กล่าวว่า ทฤษฎีการตอบข้อคำถาม เป็นทฤษฎีที่อธิบายหรือทำนายความสัมพันธ์ระหว่างลักษณะหรือความสามารถที่มีอยู่ในตัวบุคคล กับพฤติกรรมการตอบสนองต่อข้อสอบของบุคคลนั้น โดยทฤษฎีนี้เชื่อว่าพฤติกรรมการตอบสนองต่อข้อสอบของบุคคลจะถูกกำหนด โดยลักษณะหรือความสามารถที่อยู่ภายใน ซึ่งไม่สามารถสังเกตได้

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

ครอคเกอร์ และแอลจินา (Crocker and Algina. 1986 : 339) กล่าวว่าทฤษฎีการตอบข้อคำถามได้พัฒนาขึ้น เพื่ออธิบายคุณลักษณะภายในของบุคคลในการตอบข้อสอบ และหัวใจสำคัญของทฤษฎีนี้ คือ การใช้โมเดลทางคณิตศาสตร์ เพื่อค้นหาระดับความแตกต่างของความสามารถของผู้สอบ

สงบ ลักษณะ (2525 : 49) กล่าวว่า ทฤษฎีการตอบข้อคำถามเป็นการมุ่งหาความสัมพันธ์ระหว่างความสามารถที่แท้จริงกับการตอบข้อสอบ ซึ่งความสามารถที่แท้จริง คือ คุณลักษณะที่วัดด้วยแบบทดสอบ เช่น ความรู้ ความเข้าใจทักษะในด้านต่าง ๆ ส่วนการตอบข้อสอบ

หมายถึง ข้อสอบที่ให้คะแนนแบบตอบผิดให้ 0 คะแนน ตอบถูกให้ 1 คะแนน นอกจากนี้ สังเกตลักษณะ ยังกล่าวว่าถ้าข้อตกลงเบื้องต้นของการสอบตรงตามที่ทฤษฎีต้องการ และเราสามารถคำนวณค่าพารามิเตอร์ได้เหมาะสมกับโมเดลแล้วจะมีผลคือ คือ

1). ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนก ค่าความยากและค่าการเดา จะเป็นค่าที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบไม่ว่าจะนำไปสอบกับผู้ใด

2). เมื่อทราบลักษณะการตอบข้อสอบแต่ละข้อของผู้สอบคนใด เราสามารถคำนวณหาความสามารถที่แท้จริงของบุคคลนั้นได้ ค่าความสามารถที่แท้จริงนี้มีความสัมพันธ์โดยตรงกับคะแนนจริง เรียกลักษณะเช่นนี้ว่า ความเป็นอิสระของข้อสอบ

สำเร็จ บุญเรืองรัตน์ (2525 : 2) กล่าวถึง ใจความสำคัญของทฤษฎีการตอบข้อคำถามว่าเป็นผลการสอบจากแบบทดสอบใด ๆ ซึ่งสามารถพยากรณ์ หรือทำนายได้จากความสามารถของคน ๆ นั้น

อุทุมพร จามรมาน (2529 : 94) กล่าวว่าเป็นการหาความสัมพันธ์ระหว่างโอกาสที่จะตอบคำถามได้ถูกต้องกับระดับของความสามารถของผู้ตอบจึงมุ่งที่การประมาณ ผลการตอบข้อคำถามกับความสามารถว่ามีความสัมพันธ์กันอย่างไร

จากหลักการดังกล่าวจะเห็นได้ว่า ทฤษฎีการตอบข้อคำถามนี้เป็นทฤษฎีที่ใช้อธิบายความสัมพันธ์ระหว่างปริมาณของการกระทำหรือคะแนนของผู้สอบ กับปริมาณ ความสามารถของผู้สอบนั่นเอง ซึ่งสามารถเขียนในรูปสมการได้ดังนี้ (สำเร็จ บุญเรืองรัตน์, 2525 : 2)

$$P = f(\theta)$$

เมื่อ P แทน ผลการสอบ (Performance)

θ แทน ความสามารถ (Ability)

f แทน ฟังก์ชัน (Function)

2.1 ข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม

1). แบบทดสอบที่มีมิติเดียว (Unidimension Test) หมายความว่าข้อสอบแต่ละข้อในแบบทดสอบจะต้องวัดความสามารถหรือคุณลักษณะเดียวกัน หรือมีความเป็นเอกพันธ์กัน

2). ข้อสอบแต่ละข้อจะต้องเป็นอิสระจากกัน (Local Independence) หมายความว่า การตอบข้อสอบข้อใดข้อหนึ่งถูกจะไม่มีผลต่อการตอบข้อสอบข้ออื่น ๆ

3). โอกาสที่ผู้สอบจะตอบข้อสอบข้อนั้นถูกขึ้นอยู่กับโค้งลักษณะข้อสอบ (Item Characteristic Curve ; ICC) ของแต่ละโมเดลที่ใช้ ไม่ขึ้นกับการแจกแจงความสามารถของกลุ่มตัวอย่าง

นอกจากนี้ ลอร์ด (Lord. 1980 : 9) ได้กล่าวว่า เมื่อแบบทดสอบใดมีคุณลักษณะของ Unidimension. แล้วจะมีคุณสมบัติ Local Independence โดยอัตโนมัติไม่จำเป็นต้องกล่าวแยกกันก็ได้

2.2 พารามิเตอร์ในทฤษฎีการตอบข้อคำถาม

ตามทฤษฎีนี้แบ่งพารามิเตอร์เป็น 2 ชนิด คือ

1). พารามิเตอร์ของข้อสอบ (Item Parameter) ได้แก่

1.1) ค่าความยาก (b) มีค่าตั้งแต่ $-\alpha$ ถึง $+\alpha$ แต่ในทางปฏิบัติจะอยู่ระหว่าง -2 ถึง 2 ค่า -2 แสดงว่าข้อสอบนั้นง่ายมาก และค่า 2 แสดงว่าข้อสอบนั้นยากมาก

1.2) ค่าอำนาจจำแนก (a) มีค่าตั้งแต่ $-\alpha$ ถึง $+\alpha$ แต่ในทางปฏิบัติมีค่าตั้งแต่ 0 ถึง 2 เพราะค่า a ที่เป็นลบแสดงว่าข้อสอบไม่ดี ใช้ไม่ได้ ต้องตัดทิ้ง ค่า 0 แสดงว่าข้อสอบไม่มีค่าอำนาจจำแนก ค่า +2 แสดงว่าข้อสอบมี ค่าอำนาจจำแนกสูง ในการคัดเลือกข้อสอบ ข้อสอบที่คัดไว้จะมีค่า a ตั้งแต่ 0.3 ขึ้นไป

1.3) ค่าสัมประสิทธิ์การเดา (c) เป็นค่าที่แสดงถึงโอกาสการตอบข้อสอบถูก โดยไม่มีความรู้ในเรื่องนั้น ๆ มีค่าจาก 0 ถึง 2 จะคัดเลือกเอาข้อสอบที่มีค่า c ต่ำกว่า 0.3 ลงมา

2). พารามิเตอร์ของผู้สอบ (Examinees Parameter) ได้แก่ ระดับความสามารถของผู้สอบ (θ) มีค่าอยู่ระหว่าง -3 ถึง +3 ค่า -3 แสดงว่ามีความสามารถต่ำ และค่า +3 แสดงว่ามีความสามารถสูง

2.3 โมเดลในทฤษฎีการตอบข้อคำถาม

ทฤษฎีการตอบข้อคำถามได้พัฒนารูปแบบต่าง ๆ ทำให้เกิดโมเดลเฉพาะขึ้น โดยแต่ละโมเดลจะแตกต่างกันตามฟังก์ชันทางคณิตศาสตร์ และจำนวนพารามิเตอร์ที่ใช้อธิบายโค้งลักษณะข้อคำถามของแต่ละโมเดล โมเดลที่กล่าวถึงกันมากมีอยู่ 2 โมเดล คือ

2.3.1 นอร์มัล โอจีฟ โมเดล (Normal Ogive Model) เป็นโมเดลที่ใช้อธิบายโค้งลักษณะข้อสอบ ด้วยพารามิเตอร์ 2 ตัว คือ ค่าความยาก และค่าอำนาจจำแนกของข้อสอบ ซึ่งเขียนเป็นสมการได้ คือ (Hambleton and Swaminathan. 1985 : 35)

$$P_i(\theta) = \int_{-\alpha}^{a_i(\theta-b_i)} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz \quad ; i = 1, 2, 3, \dots, n$$

$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบคนหนึ่ง ซึ่งมีระดับความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

$$\frac{e^{\frac{-z^2}{2}}}{\sqrt{2\pi}} \quad \text{คือ Normal Density Function}$$

b_i คือ ค่าความยากของข้อสอบข้อที่ i

a_i คือ ค่าอำนาจจำแนกของข้อสอบที่ i

θ คือ ระดับความสามารถของผู้สอบ

2.3.2 โลจิสติก โมเดล (Logistic Model) เป็นโมเดลที่มีลักษณะใกล้เคียงกับนอร์มัล โอโจฟ โมเดล แต่สะดวกในการประมาณค่ามากกว่า และเป็นโมเดลที่ได้พัฒนาให้เหมาะสมกับสถานการณ์การทดสอบ เช่น ในกรณีที่ตอบข้อสอบได้ด้วยการเดาปลายค่าสุดของโค้ง ICC ไม่เป็นศูนย์ แสดงว่าการทำข้อสอบย่อมมีการเดา และในกรณีผู้สอบทำข้อสอบด้วยความสะเพร่า โมเดลนี้ จะเพิ่มค่าความรอบคอบขึ้นอีกหนึ่งค่า ซึ่งค่าความรอบคอบของข้อสอบแต่ละข้อ มีค่าน้อยกว่า 1 โมเดลโลจิสติกแบ่งเป็น 4 โมเดลย่อย ๆ ตามพารามิเตอร์ ดังนี้ (Hambleton and Swaminathan. 1985 : 34-49)

2.3.2.1 โมเดลที่ใช้พารามิเตอร์ตัวเดียว (One - Parameter Logistic Model) หรือ ราสช์โมเดล (Rasch Model) โมเดลนี้ถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันและไม่มี การเดา เขียนเป็นสมการได้ดังนี้

$$P_i(\theta) = \frac{e^{D\bar{a}(\theta - b_i)}}{1 + e^{D\bar{a}(\theta - b_i)}} \quad ; i = 1, 2, 3, \dots, n$$

หรือ
$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D\bar{a}(\theta - b_i)}} \quad ; i = 1, 2, 3, \dots, n$$

เมื่อ $\bar{a}$ เป็นค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ

$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบคนหนึ่ง ซึ่งมีระดับความสามารถ θ จะตอบข้อสอบ

ข้อที่ i ได้ถูกต้อง

b_i คือ ระดับความยากของข้อสอบข้อที่ i

θ คือ ระดับความสามารถของผู้สอบ

D คือ ค่าคงที่มีค่าเท่ากับ 1.702

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818

2.3.2.2 โมเดลที่ใช้พารามิเตอร์ 2 ตัว (Two - Parameter Logistic Model) เบิร์นบอม (Birnbaum) ได้เสนอโมเดลนี้ ในปี ค.ศ.1957 โดยใช้พารามิเตอร์ 2 ตัว คือ ค่าความยากและอำนาจจำแนก เขียนเป็นสมการได้ดังนี้

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad ; i = 1, 2, 3, \dots, n$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบคนหนึ่ง ซึ่งมีระดับความสามารถ θ จะตอบ

ข้อสอบข้อที่ i ได้ถูกต้อง

b_i คือ ค่าความยากที่แสดงระดับความสามารถที่แท้จริงที่จุดโค้ง ชั้นที่สุด

หรือในกรณีที่ไม่มีค่า b_i เท่ากับ θ ณ จุดความน่าจะเป็น 0.5

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i ที่เป็นสัดส่วนกับความชันของ P_i

(θ) ณ ตำแหน่ง $\theta = b_i$

D คือ Scaling Factor มีค่าเท่ากับ 1.7

e คือ ค่าคงที่มีค่าเท่ากับ 2.718271

2.3.2.3 โมเดลที่ใช้พารามิเตอร์ 3 ตัว (Three- Parameter Logistic Model) เป็นโมเดลที่พัฒนามาจากโมเดลที่ใช้พารามิเตอร์ 2 ตัว เพื่อให้เหมาะสมกับแบบทดสอบที่มีอิทธิพลของการเดาเข้ามาเกี่ยวข้องโดยเพิ่มพารามิเตอร์อีกตัวหนึ่ง คือ ค่าการเดา เขียนเป็นสมการได้ดังนี้

$$P_i(\theta) = c_i + (1 + c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad ; i = 1, 2, 3, \dots, n$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบคนหนึ่ง ซึ่งมีระดับความสามารถ θ จะตอบ

ข้อสอบข้อที่ i ได้ถูกต้อง

C_i คือ ค่าความน่าจะเป็นของผู้สอบที่มีความสามารถต่ำมาก มีโอกาสจะทำข้อสอบข้อนี้ได้ถูกต้อง หรือค่าการเดา

b_i คือ ค่าความยากที่แสดงระดับความสามารถที่แท้จริงที่จุดโค้งชันที่สุด คือ θ

ณ จุดความน่าจะเป็น $(1+C_i) / 2)$

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i ที่เป็นสัดส่วนกับความชันของ P_i

(θ) ณตำแหน่ง $\theta = b_i$

D คือ ค่าคงที่มีค่าเท่ากับ 1.702

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818

2.3.2.4 โมเดลที่ใช้พารามิเตอร์ 4 ตัว (Four-Parameter Logistic Model) โมเดลนี้เสนอโดยแม็คโดแนลด์ (McDonald) ในปี ค.ศ.1967 ต่อมา ลอร์ด (Lord) และ บาร์ตัน (Barton) ได้พัฒนาในปีค.ศ.1981 โดยเพิ่มพารามิเตอร์ ค่าความรอบคอบในโมเดลที่ใช้พารามิเตอร์ 3 ตัว เป็นค่าแสดงให้เห็นถึงในกรณีที่ผู้สอบที่มีความสามารถสูงมักตอบข้อสอบผิด อาจจะเป็นเพราะความไม่รอบคอบทำให้ผู้สอบนั้นไม่เลือกคำตอบที่ถูก จะเห็นได้จากปลายสูงสุดของโค้ง ICC ต่ำกว่า 1 เขียนเป็นสมการได้ดังนี้ (Hambleton and Swaminathan. 1985 : 48)

$$p_i(\theta) = c_i + (r_i + c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad ; i = 1, 2, 3, \dots, n$$

เมื่อ r_i คือ ค่าความรอบคอบในการตอบข้อสอบข้อที่ i ซึ่งมีค่าน้อยกว่า 1

2.4 คุณภาพของแบบทดสอบของทฤษฎีการตอบข้อคำถาม

ในทฤษฎีการตอบข้อคำถาม การวิเคราะห์คุณภาพของแบบทดสอบหรือพิจารณาความเชื่อถือได้ของการประมาณค่าความสามารถที่แท้จริงนั้นเป็นการ พิจารณาจากฟังก์ชันอินฟอร์เมชันของแบบทดสอบ (Test Information Function) ซึ่งได้จากผลรวมของฟังก์ชันอินฟอร์เมชันของข้อสอบ (Item Information Function) โดยที่ค่าฟังก์ชันอินฟอร์เมชันของข้อสอบ หรือแบบทดสอบเป็นดัชนีผสม (Composite index) ที่สร้างจากดัชนีบอกคุณลักษณะของข้อสอบหลายลักษณะ เช่น ค่าอำนาจจำแนก ค่าความยาก เป็นต้น รวมเป็นดัชนีเพียงตัวเดียวเพื่อชี้ถึงคุณภาพของข้อสอบหรือแบบทดสอบได้ จะเห็นว่าในทฤษฎี IRT จะใช้ค่าฟังก์ชันอินฟอร์เมชัน ของแบบทดสอบ แทน

การหาค่าความเชื่อมั่นของคะแนนและความคลาดเคลื่อนมาตรฐานของการวัด ตามแนวคิดของ
ทฤษฎีคลาสสิกอล (Hambleton. 1977 : 64)

สำหรับในโมเดลโลจิสติก 3 พารามิเตอร์ จะสามารถหาค่าฟังก์ชันอินฟอร์เมชันของข้อ
สอบ ในแต่ละระดับความสามารถ (θ) ได้จากสูตร (Lord. 1980 : 72-74)

$$I(\theta, u_i) = \frac{(1.7a_i)^2(1 - c_i)}{[c_i + e^{1.7a_i(\theta - b_i)}][1 + e^{-1.7a_i(\theta - b_i)}]^2}$$

และสามารถกำหนดโค้งฟังก์ชันอินฟอร์เมชันของข้อสอบได้จากสมการ

$$I(\theta, u_i) = \frac{(P'_i)^2}{P_i Q_i}$$

$I(\theta, u_i)$ คือ ฟังก์ชันอินฟอร์เมชันของข้อสอบข้อที่ i

P'_i คือ ความชันของ ICC ที่ระดับความสามารถ θ

P_i คือ ความน่าจะเป็นที่ผู้สอบที่มีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้
ถูกต้อง

Q_i คือ $1 - P_i$

และสามารถหาค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบได้จากผลรวมของฟังก์ชันอินฟอร์เมชัน
ของข้อสอบทั้งหมดในแบบทดสอบ เขียนเป็นสมการได้ดังนี้

$$I(\theta) = \sum_{i=1}^n I(\theta, u_i)$$

เมื่อ $I(\theta)$ คือ ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ

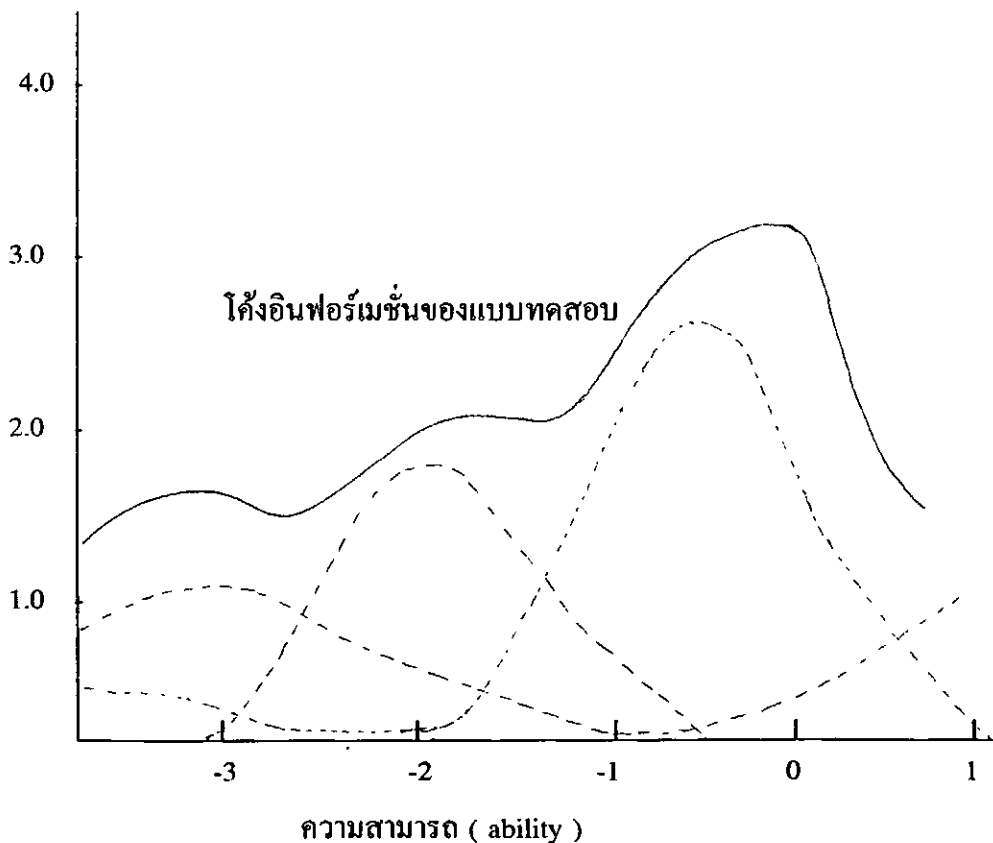
จากสมการ จะเห็นว่า ค่าฟังก์ชันอินฟอร์เมชันแต่ละข้อ ขึ้นอยู่กับความชันของโค้ง
ลักษณะข้อสอบ และความแปรปรวนของการตอบข้อสอบถูกต้องของแต่ละข้อ ในแต่ละระดับความ
สามารถ θ และยิ่งความชันของ ICC มีค่ามาก ประกอบกับค่าความแปรปรวนมีค่าน้อย ๆ โค้ง
ฟังก์ชันอินฟอร์เมชันของข้อสอบ ที่ระดับความสามารถนั้นจะยิ่งสูงขึ้น สำหรับโค้งฟังก์ชันอิน
ฟอร์เมชันของข้อสอบที่มีค่าสูงสุด ณ ระดับความสามารถใดก็จะจำแนกระดับความสามารถของผู้

สอบได้ดี ณ ระดับความสามารถนั้น (Hambleton. 1977 : 66 อ้างถึงใน ประดิษฐ์ เรื่องตระกูล, 2529 : 23-24)

สำหรับค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบฉบับหนึ่งจะมีได้หลายค่าแตกต่างกันไปตามระดับ θ และ ณ ระดับความสามารถใดที่ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบที่มีค่าสูง แสดงว่า แบบทดสอบมีคุณภาพดีมากในการประมาณค่าความสามารถที่ระดับนั้น แอมเบิลตัน และสวามินาธาน (Hambleton and Swaminathan. 1985 : 107) กล่าวว่าเมื่อแบบทดสอบมีข้อสอบจำนวนมากขึ้น ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบก็จะเพิ่มมากขึ้น

ถ้าเรามีกลุ่มข้อสอบที่ทราบโค้งอินฟอร์เมชัน เราสามารถสร้างแบบทดสอบให้มีโค้งอินฟอร์เมชันของแบบทดสอบ ณ ระดับหนึ่งของความสามารถที่เราต้องการได้ เช่น เพื่อคัดเลือกนักเรียนให้ได้รับทุนก็ต้องใช้ข้อสอบที่มีประสิทธิภาพสูงสุดที่ ระดับความสามารถสูง ๆ นั่นก็คือให้มีโค้งอินฟอร์เมชันของแบบทดสอบสูง ณ ระดับความสามารถสูง ๆ นั้นเอง (ดังภาพประกอบ 1)

อินฟอร์เมชัน (information)



ภาพประกอบ 1 โค้งอินฟอร์เมชันของข้อสอบ 5 ข้อ และโค้งอินฟอร์เมชันของแบบทดสอบ

จากภาพประกอบ 1 พบว่าข้อสอบข้อที่ 1 จะเป็นข้อสอบที่ไม่ค่อยดี เพราะค่าอินฟอร์เมชันต่ำมาก ข้อ 2 ใช้ได้กับผู้สอบอ่อนมาก ๆ ข้อ 3 จะใช้ได้ดีกับผู้สอบค่อนข้างอ่อน ข้อ 4 ใช้ได้ดีและให้ข้อมูลที่ดีกับผู้สอบที่มีความสามารถระดับกลาง ๆ และข้อ 5 จะให้ข้อมูลที่ดีกับผู้ที่มีความสามารถสูง ๆ และเมื่อรวมข้อสอบทั้ง 5 ข้อเป็นแบบทดสอบแล้วจะเป็นแบบทดสอบที่ให้ข้อมูลของผู้สอบได้ดีในระดับความสามารถกลาง ๆ

วาร์ม (Warm. 1978 : 77) กล่าวว่า ในการสอบทุกครั้งจะมีความคลาดเคลื่อนในการวัดเกิดขึ้น ที่เรียกว่า ความคลาดเคลื่อนมาตรฐานในการประมาณ ค่าความสามารถของผู้สอบ

[Standard Error of the ability estimates : $SE(\theta)$] ความสัมพันธ์ระหว่างฟังก์ชันอินฟอร์เมชันของแบบทดสอบและความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ แสดงด้วยสมการ (Warm. 1978 : 77)

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}}$$

เมื่อ $SE(\theta)$ คือ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ณ ที่ความสามารถ θ

ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบจะแปรเปลี่ยนไปตามระดับความสามารถและแปรผกผัน (Inverses) กับค่าความคลาดเคลื่อนมาตรฐาน ในการประมาณค่าความสามารถ (Warm. 1979 : 77, Hambleton and Swaminathan. 1985 : 104-105)

3. ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบในทฤษฎีการตอบข้อคำถาม

จากแนวคิดของทฤษฎีการตอบข้อคำถามที่มีจุดมุ่งหมายในการสอบ คือ เมื่อใช้แบบทดสอบวัดความสามารถของแต่ละบุคคลแล้วต้องการให้ผลการสอบนั้นไปประมาณค่าความสามารถที่มีอยู่ในตัวบุคคลที่ไม่สามารถสังเกตได้หรือวัดได้โดยตรง ดังนั้น แบบทดสอบที่มีคุณภาพที่ดีควรเป็นแบบทดสอบที่ให้ผลการสอบสามารถนำไปใช้ประมาณค่าความสามารถที่ต้องการวัดได้ถูกต้องแม่นยำที่สุด ดังนั้นที่ใช้ประมาณค่าความสามารถที่ต้องการวัดได้ถูกต้องแม่นยำที่สุดจึงเป็นการพิจารณาคุณภาพของแบบทดสอบ คือ ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ โดยกำหนดว่า ณ ระดับความสามารถที่ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบมีค่าสูง แสดงว่าแบบทดสอบมีคุณภาพที่ดีในการประมาณค่าความสามารถที่ระดับนั้น และนอกจากใช้พิจารณา

คุณภาพของแบบทดสอบแล้ว ยังสามารถนำมาใช้เปรียบเทียบคุณภาพของแบบทดสอบได้อีกด้วย ดังนี้ (Hambleton and Swaminathan. 1985 : 101)

ถ้าเรามีแบบทดสอบหลายฉบับที่วัดความสามารถอย่างเดียวกันแล้ว เราสามารถเปรียบเทียบคุณภาพของแบบทดสอบเหล่านั้นได้ โดยการนำค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบของแต่ละแบบทดสอบที่ต้องการเปรียบเทียบ 2 ฉบับ ณ ระดับความสามารถเดียวกันมาหาค่าอัตราส่วนกัน ค่าอัตราส่วนของฟังก์ชันอินฟอร์เมชันของแบบทดสอบ 2 ฉบับที่นำมาเปรียบเทียบกัน ณ ระดับความสามารถเดียวกัน คือค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบซึ่งนิยามดังนี้

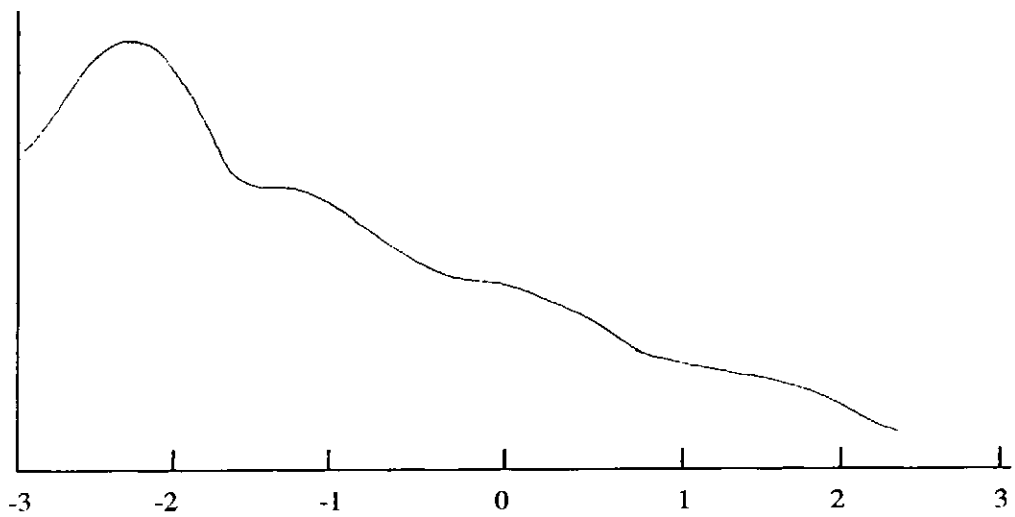
ค่าประสิทธิภาพสัมพัทธ์ของคะแนนแบบทดสอบ y กับคะแนนแบบทดสอบ x คืออัตราส่วนของค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบทั้ง 2 ฉบับ ดังสมการ

$$RE(y, x) = \frac{I(\theta, y)}{I(\theta, x)}$$

โดยที่คะแนน x และ y เป็นคะแนนจากแบบทดสอบที่แตกต่างกัน 2 ฉบับ ณ ระดับความสามารถเดียวกัน หรือ x และ y เป็นผลจากวิธีการให้คะแนนที่แตกต่างกัน 2 วิธี ของแบบทดสอบฉบับเดียวกัน ตามนิยามของค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบนั้นค่าความสามารถใน $I(\theta, y)$ ต้องเป็นค่าความสามารถเดียวกันใน $I(\theta, x)$ และค่าประสิทธิภาพสัมพัทธ์ของคะแนนแบบทดสอบ 2 ฉบับ มีค่าแปรเปลี่ยนไปตามระดับความสามารถด้วย (Lord. 1980 : 83)

การแปลความหมายของค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ พิจารณาดังนี้ระดับความสามารถใด ถ้าค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่ามากกว่า 1 แสดงว่าแบบทดสอบ y มีคุณภาพสูงกว่าแบบทดสอบ x ที่ระดับความสามารถนั้น ถ้าค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่าน้อยกว่า 1 แสดงว่าแบบทดสอบ y มีคุณภาพต่ำกว่าแบบทดสอบ x ที่ระดับความสามารถนั้น และถ้าประสิทธิภาพสัมพัทธ์ของแบบทดสอบมีค่าเท่ากับ 1 แสดงว่า แบบทดสอบ y และแบบทดสอบ x มีคุณภาพเท่ากัน ที่ระดับความสามารถนั้น (Warm. 1978 : 76) พิจารณาจากภาพดังนี้

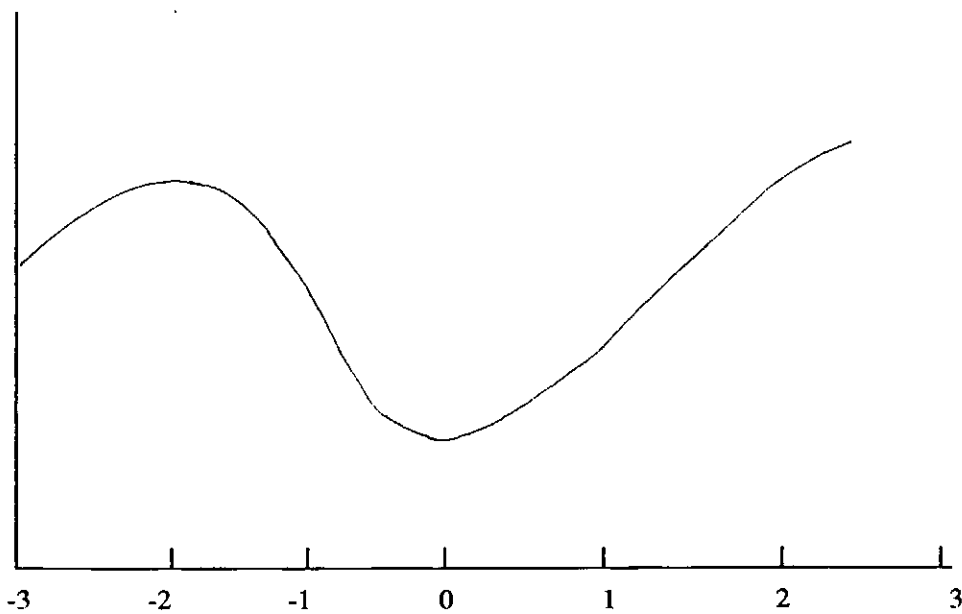
อินฟอร์เมชันของแบบทดสอบ



ระดับความสามารถ

ภาพประกอบ 2 โค้งอินฟอร์เมชันของแบบทดสอบ ที่มีประสิทธิภาพสำหรับการวัดระดับ
ความสามารถต่ำ ($\theta = -2.3$)

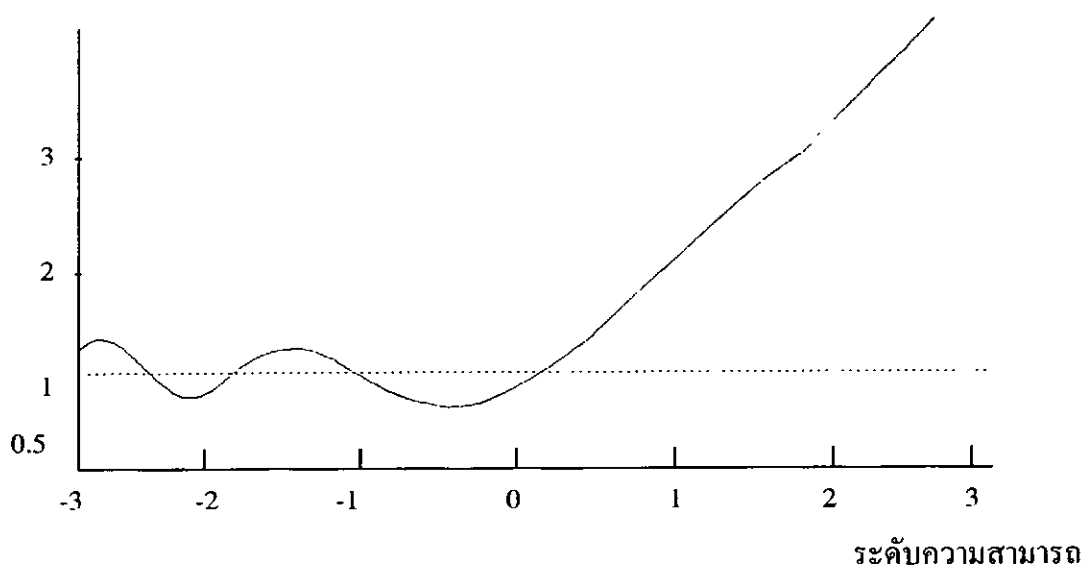
อินฟอร์เมชันของแบบทดสอบ



ระดับความสามารถ

ภาพประกอบ 3 โค้งอินฟอร์เมชันของแบบทดสอบ ที่มีประสิทธิภาพสำหรับการวัดทั้งระดับ
ความสามารถสูงและต่ำ

ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ



ภาพประกอบ 4 โค้งประสิทธิภาพสัมพัทธ์ของแบบทดสอบของภาพประกอบ 3 เมื่อเปรียบเทียบกับแบบทดสอบของภาพประกอบ 2

จากภาพประกอบ 4 เป็นการแสดงถึงการเปรียบเทียบคุณภาพของแบบทดสอบ 2 ฉบับ โดยการนำเอาค่าฟังก์ชันอินฟอर्मชันของแบบทดสอบภาพประกอบ 3 และภาพประกอบ 2 มาหาค่าอัตราส่วนกัน ณ ระดับความสามารถเดียวกัน ของทุก ๆ ระดับความสามารถตั้งแต่ -3 ถึง 3 ซึ่งสรุปได้ว่าในกลุ่มนักเรียนระดับความสามารถสูง (ระดับความสามารถมากกว่า 1 ถึง 3) แบบทดสอบของภาพประกอบ 3 มีคุณภาพสูงกว่าแบบทดสอบภาพประกอบ 2 ในกลุ่มนักเรียนระดับความสามารถปานกลาง (ระดับความสามารถตั้งแต่ -1 ถึง 1) แบบทดสอบของภาพประกอบ 3 มีคุณภาพต่ำกว่าแบบทดสอบของภาพประกอบ 2 และในกลุ่มนักเรียนระดับความสามารถต่ำ (ระดับความสามารถน้อยกว่า -1 ถึง -3) มีทั้งแบบทดสอบของภาพประกอบ 3 มีคุณภาพสูงกว่าแบบทดสอบของภาพประกอบ 2 และแบบทดสอบของภาพประกอบ 3 มีคุณภาพต่ำกว่าแบบทดสอบของภาพประกอบ 2

ลอร์ด (Lord, 1980 : 96-112) ได้ศึกษาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบแบบคำศัพท์ระดับ 6 จำนวน 7 ฉบับด้วยกัน โดยศึกษาจำนวนตัวเลือกที่เหมาะสมของข้อสอบที่มีความยาวและค่าการเดาข้อสอบที่แตกต่างกัน โดยได้เปรียบเทียบกับ Metropolitan Reading Tests (1970), Intermediate Level, Form F, Word Analysis Subtest (MAT) กับแบบทดสอบอีก 6 ฉบับ ดังนี้

1. Sequential Tests Education Progress Series (1969), Level 4, Form A, Reading Subtest (STEP)

2. California Achievement Tests (1970), Level 4, Form A, Reading Vocabulary (CAT)

3. Iowa Test of Basic Skills (1970), Level 12, Form 5, Vocabulary (ITBS)

4. Stanford Reading Tests (1964), Intermediate, Form W, Word Meaning (Stanford)

5. Comprehensive Tests of Basic Skills (1968), Level 3, Form Q, Reading Vocabulary (CTBS)

6. SRA Achievement Series (1971), Green Edition, Form E, Vocabulary (SRA)
ผลการศึกษารูปได้ดังนี้

1. STEP มีคุณภาพสูงกว่า MAT ในกลุ่มนักเรียนที่มีระดับความสามารถต่ำ (ต่ำกว่าเปอร์เซ็นต์ไทล์ที่ 20) ส่วนระดับความสามารถอื่น ๆ STEP มีคุณภาพต่ำกว่า

2. CAT มีคุณภาพสูงกว่า MAT ในกลุ่มนักเรียนที่มีระดับความสามารถสูงมาก (มากกว่าเปอร์เซ็นต์ไทล์ที่ 95) ส่วนระดับความสามารถอื่น ๆ CAT มีคุณภาพต่ำกว่า

3. ITBS มีคุณภาพสูงกว่า MAT ในกลุ่มนักเรียนที่มีระดับความสามารถสูงมาก (มากกว่าเปอร์เซ็นต์ไทล์ที่ 90) ส่วนระดับความสามารถอื่น ๆ ITBS มีคุณภาพต่ำกว่า

4. Stanford มีคุณภาพสูงกว่า MAT ในกลุ่มนักเรียนที่มีระดับความสามารถปานกลางและสูง (มากกว่าเปอร์เซ็นต์ไทล์ที่ 50) และมีคุณภาพต่ำกว่า MAT ในกลุ่มนักเรียนที่มีระดับความสามารถปานกลางและต่ำ (น้อยกว่าเปอร์เซ็นต์ไทล์ที่ 50)

5. CTBS มีคุณภาพสูงกว่า MAT ในกลุ่มนักเรียนที่มีความสามารถสูง (มากกว่าเปอร์เซ็นต์ไทล์ที่ 80) มีคุณภาพใกล้เคียงกันในกลุ่มนักเรียนที่มีระดับความสามารถปานกลางและค่อนข้างสูง (มากกว่าเปอร์เซ็นต์ไทล์ที่ 50 ถึงเปอร์เซ็นต์ไทล์ที่ 80) และมีคุณภาพต่ำกว่า MAT ในกลุ่มนักเรียนที่มีระดับความสามารถปานกลางและต่ำ (น้อยกว่าเปอร์เซ็นต์ไทล์ที่ 50)

6. SRA มีคุณภาพต่ำกว่า MAT ในกลุ่มนักเรียนที่มีระดับความสามารถต่ำ (น้อยกว่าเปอร์เซ็นต์ไทล์ที่ 30) ส่วนระดับความสามารถอื่น ๆ SRA มีคุณภาพสูงกว่า

การศึกษาจำนวนตัวเลือกที่เหมาะสมของข้อสอบโดยเปรียบเทียบคุณภาพของแบบทดสอบที่มีความยาวของแบบทดสอบและค่าการเดาของข้อสอบที่แตกต่างกันซึ่งได้ศึกษาแบบทดสอบความถนัดทางการเรียนด้านภาษา (TSA 13) คือ The College Board Scholastic Aptitude Test, Form TSA 13 เป็นแบบทดสอบแบบเลือกตอบที่มี 5 ตัวเลือก จำนวน 90 ข้อ

ค่าพารามิเตอร์ของข้อสอบประมาณค่าตามโมเดลโลจิสติกที่มีพารามิเตอร์ 3 ตัว ค่ามัธยฐานของค่าการเดาของข้อสอบมีค่าประมาณ .15 และได้สร้างแบบทดสอบสมมุติ (Hypothetical Test) ขึ้นจากแบบทดสอบนี้อีก 4 ฉบับ โดยการแทนค่าการเดาของข้อสอบใหม่ในแต่ละข้อของแบบทดสอบสมมุติแต่ละฉบับตามลำดับดังนี้ .20, .25, .33 และ .50 ส่วนค่าอำนาจจำแนกของข้อสอบและค่าความยากของข้อสอบยังใช้ค่าเดิมใน TSA 13 ซึ่งค่าการเดาของข้อสอบใหม่ในแบบทดสอบสมมุติแต่ละฉบับ มีความสัมพันธ์กับจำนวน ตัวเลือกของข้อสอบตามลำดับดังนี้ 5 ตัวเลือก 4 ตัวเลือก 3 ตัวเลือก และ 2 ตัวเลือก และเนื่องจากจำนวนของข้อสอบในแบบทดสอบ สมมุติแต่ละฉบับเป็นสัดส่วนโดยตรงกับค่าการเดาของข้อสอบ ดังนั้นจำนวนของข้อสอบในแบบทดสอบสมมุติ 4 ฉบับมีจำนวนแตกต่างกันตามลำดับดังนี้ 120 ข้อ 150 ข้อ 200 ข้อ และ 300 ข้อ แล้วเปรียบเทียบคุณภาพของแบบทดสอบโดยการหาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบของ TSA 13 ที่มีค่ามัธยฐานของค่าการเดาของข้อสอบประมาณ .15 แบบทดสอบแบบสมมุติที่มีค่าการเดาของข้อสอบเป็น .20, .25 และ .50 เปรียบเทียบกับแบบทดสอบสมมุติที่มีค่าการเดาของข้อสอบเป็น .33 ผลการศึกษาพบว่า ที่ระดับความสามารถสูง แบบทดสอบที่มีคุณภาพสูงที่สุด คือ แบบทดสอบที่มีค่าการเดาของข้อสอบเป็น .50 รองลงมาเป็น .33, .25, .20 และ TSA 13 ตามลำดับ และที่ระดับความสามารถต่ำ แบบทดสอบที่มีคุณภาพสูงที่สุดคือ TSA 13 รองลงมาเป็นแบบทดสอบที่มีค่าการเดาของข้อสอบเป็น .20, .25, .33 และ .50 นั่นคือ ผลกระทบจากการลดจำนวนตัวเลือกของข้อสอบโดยที่เป็นสัดส่วนโดยตรงกับความยาวของแบบทดสอบ ทำให้คุณภาพของแบบทดสอบเพิ่มขึ้นสำหรับกลุ่มผู้สอบที่มีระดับความสามารถสูง และลดลงสำหรับกลุ่มผู้สอบที่มีระดับความสามารถต่ำ

จากแบบทดสอบสมมุติ 4 ฉบับ ลอร์ด ได้ศึกษาเพิ่มเติมโดยพิจารณาถึงค่าความยากของข้อสอบในแบบทดสอบด้วย ซึ่งจากค่ามัธยฐานของค่าความสามารถของข้อสอบของ TSA 13 มีค่าประมาณ .5 ดังนั้นค่าความยากของข้อสอบของแบบทดสอบสมมุติทั้ง 4 ฉบับ ซึ่งหาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบที่มีค่าการเดาของข้อสอบ .50 (โดยลดค่าความยากของข้อสอบทุกข้อเท่ากับ 0.1 เพื่อให้ยากขึ้นเล็กน้อย) แบบทดสอบที่มีค่าการเดาของข้อสอบ .25 (โดยการเพิ่มค่าความยากของข้อสอบทุกข้อเท่ากับ 0.1 เพื่อให้ยากขึ้นเล็กน้อย) และแบบทดสอบที่มีค่าการเดาของข้อสอบ .20 (โดยการเพิ่มค่าความยากของข้อสอบทุกข้อเท่ากับ 0.2 เพื่อให้ยากขึ้นมาก) เปรียบเทียบกับแบบทดสอบที่มีค่าการเดาของข้อสอบ .33 (โดยมีค่าความยากของข้อสอบทุกข้อเท่ากับ 0.5) ผลการศึกษาพบว่าแบบทดสอบที่มีค่าการเดาของข้อสอบ .33 และมีค่าความยากของข้อสอบทุกข้อเท่ากับ 0.5 มีคุณภาพสูงกว่าแบบทดสอบที่มีค่าการเดาของข้อสอบ .50 และมีความยากของข้อสอบทุกข้อเท่ากับ 0.4 แบบทดสอบที่มีค่าการเดาของข้อสอบ .25 และมีค่าความยากของ

ข้อสอบทุกข้อเท่ากับ 0.6 และแบบทดสอบที่มีค่าการเดาของข้อสอบ .20 และมีค่าความยากของข้อสอบทุกข้อเท่ากับ 0.7

งานวิจัยที่เกี่ยวข้องภายในประเทศ

ชนะ ทาวงศ์ (2521 : 25-27) ได้ศึกษาเปรียบเทียบคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนของแบบทดสอบปรนัยชนิดเลือกตอบที่มีจำนวนข้อและตัวเลือกแตกต่างกันในวิชาวิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 4 โดยใช้แบบทดสอบวัดความเข้าใจทางวิทยาศาสตร์ วัดทักษะทางวิทยาศาสตร์ และวัดปฏิบัติการทางวิทยาศาสตร์ ในแบบทดสอบหนึ่ง ๆ สร้างเป็น 4 ฉบับย่อย คือ แบบทดสอบชนิด 30 ข้อ 2 ตัวเลือก, 20 ข้อ 3 ตัวเลือก, 15 ข้อ 4 ตัวเลือก และ 12 ข้อ 5 ตัวเลือก ผลการศึกษาปรากฏว่าแบบทดสอบที่มี 30 ข้อ 2 ตัวเลือก มีแนวโน้มให้ค่าความเชื่อมั่นสูงกว่าแบบทดสอบลักษณะอื่น สำหรับผลการเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบเป็นดังนี้

แบบทดสอบวัดความเข้าใจทางวิทยาศาสตร์ แบบทดสอบที่มี 30 ข้อ 2 ตัวเลือก มีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มี 15 ข้อ 4 ตัวเลือก และ 12 ข้อ 5 ตัวเลือก อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.5 แบบทดสอบที่มี 20 ข้อ 3 ตัวเลือก มีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มี 15 ข้อ 4 ตัวเลือก อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และแบบทดสอบที่มี 15 ข้อ 4 ตัวเลือก มีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มี 12 ข้อ 5 ตัวเลือก อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แบบทดสอบวัดทักษะทางวิทยาศาสตร์ แบบทดสอบที่มี 30 ข้อ 2 ตัวเลือก, 20 ข้อ 3 ตัวเลือก, 15 ข้อ 4 ตัวเลือก และ 12 ข้อ 5 ตัวเลือก มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

แบบทดสอบวัดปฏิบัติการทางวิทยาศาสตร์ แบบทดสอบที่มี 30 ข้อ 2 ตัวเลือก มีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มี 20 ข้อ 3 ตัวเลือก 15 ข้อ 4 ตัวเลือก และ 12 ข้อ 5 ตัวเลือก อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แบบทดสอบที่มี 20 ข้อ 3 ตัวเลือก มีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มี 15 ข้อ 4 ตัวเลือก อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แบบทดสอบที่มี 15 ข้อ 4 ตัวเลือก มีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มี 12 ข้อ 5 ตัวเลือก อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

บุญเรือน จรัสวิล (2521 : 40-41) ได้ศึกษาการเพิ่มข้อสอบที่มีผลต่อค่าความเชื่อมั่นของแบบทดสอบ โดยเพิ่มจำนวนข้อสอบในแบบทดสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์จากเดิม 30 ข้อ เป็น 60 ข้อ ข้อสอบที่เพิ่มในแต่ละฉบับมีช่วงความยากต่างกัน คือ .40 -.60, .30 -.70 และ .20 -.80 ผลการศึกษาปรากฏว่าการเพิ่มจำนวนข้อสอบด้วยข้อสอบที่มีระดับความยากต่าง

กัน ไม่ทำให้ค่าความเชื่อมั่นแตกต่างกัน เมื่อนำค่าความเชื่อมั่นของแบบทดสอบที่เพิ่มความยาวนี้ไปเปรียบเทียบกับค่าความเชื่อมั่นของแบบทดสอบฉบับเดิม พบว่าแบบทดสอบที่เพิ่มจำนวนข้อในช่วงระดับความยาก .30 -.70 และ .20 -.80 มีค่าความเชื่อมั่นสูงกว่าค่าความเชื่อมั่นเดิมอย่างมีนัยสำคัญทางสถิติที่ .01 และ .05 ตามลำดับ ส่วนแบบทดสอบที่เพิ่มจำนวนข้อในช่วงระดับความยาก .40 -.60 ค่าความเชื่อมั่นสูงกว่าค่าความเชื่อมั่นเดิมอย่างไม่มีนัยสำคัญทางสถิติ

วิภาดา วัฒนกุลกิตต์ (2529 : 1-53) ได้ศึกษาเปรียบเทียบค่าความยาก ค่าความเชื่อมั่น และสัดส่วนการเดาของแบบทดสอบแบบเลือกตอบที่มีตัวเลือกต่างกัน โดยใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ จำนวน 32 ข้อ ซึ่งผู้วิจัยได้สร้างขึ้นเอง 3 ลักษณะ คือ แบบ 3 ตัวเลือก 4 ตัวเลือก และ 5 ตัวเลือก โดยมีข้อคำถามและจำนวนข้อเหมือนกัน แต่มีจำนวนตัวเลือกไม่เท่ากันมีค่าอำนาจจำแนกเฉลี่ยของตัวลงในรายข้อไม่แตกต่างกัน นำไปทดสอบกับนักเรียน ชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2529 โรงเรียนสังกัดกรมสามัญศึกษา ในจังหวัดขอนแก่น ซึ่งเลือกมาโดยวิธีการสุ่มแบบแบ่งชั้น จำนวน 24 ห้องเรียน รวม 851 คน ปรากฏว่าค่าความยากของแบบทดสอบแบบเลือกตอบ 3 ตัวเลือก มีค่ามากกว่า 4 และ 5 ตัวเลือก ในกลุ่มนักเรียนที่ต่างระดับความสามารถกัน แต่ค่าความเชื่อมั่น ของแบบทดสอบแบบเลือกตอบทั้ง 3 ฉบับ แตกต่างกันอย่างไม่มีนัยสำคัญ และกลุ่มนักเรียนที่มีความสามารถทางการเรียนต่างกัน มีลักษณะตอบแบบทดสอบแบบเลือกตอบในลักษณะเดียวกัน ด้วยการเดาไม่แตกต่างกัน

ธีระศักดิ์ อรรถนันทน์ (2530 : 1-71) ได้ศึกษาเปรียบเทียบคุณภาพของแบบทดสอบชนิดเลือกตอบที่มีจำนวนตัวเลือกไม่เท่ากัน โดยใช้ทฤษฎีการตอบข้อคำถาม โดยใช้แบบทดสอบวัดความถนัดทางการใช้เหตุผลเชิงภาษาแบบอุปมาอุปไมยแบบเลือกตอบ 5 ตัวเลือก จำนวน 32 ข้อ ของ วินูลย์ บุญสุวรรณ และแยกเป็น 3 ฉบับ โดยในแต่ละฉบับจะมีข้อคำถามเหมือนกันและจำนวนข้อเท่ากัน แต่จำนวนตัวเลือกในแต่ละข้อในแบบทดสอบแต่ละฉบับไม่เท่ากัน โดยมี 5 ตัวเลือก 4 ตัวเลือก และ 3 ตัวเลือก ในการลดหรือตัดตัวลงในวิธีการสุ่มตัวลงที่ละข้อออก นำแบบทดสอบไปทดสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในจังหวัดเพชรบุรี จำนวน 2,960 คน การวิเคราะห์ข้อมูลใช้โปรแกรมโลจิสต์ 5 (Logist 5) ปรากฏว่าแบบทดสอบ 5 ตัวเลือก มีคุณภาพสูงสุดในกลุ่มที่มีความสามารถสูง รองลงมา คือแบบทดสอบ 4 ตัวเลือก และ 3 ตัวเลือก แบบทดสอบ 4 ตัวเลือก มีคุณภาพสูงสุดในกลุ่มที่มีความสามารถต่ำ รองลงมาคือ แบบทดสอบ 5 ตัวเลือก 3 ตัวเลือกแบบทดสอบ 3 ตัวเลือก มีคุณภาพสูงสุดในกลุ่มที่มีความสามารถปานกลาง รองลงมา คือ แบบทดสอบ 5 ตัวเลือก 4 ตัวเลือก

จักรกฤษณ์ สำราญใจ (2530 : 121-125) ได้ศึกษาประสิทธิภาพเชิงสัมพัทธ์ของข้อสอบเลือกตอบชนิดตัดสินใจคำตอบทุกตัวเลือกเทียบกับข้อสอบเลือกตอบชนิดแบบฉบับในแบบทดสอบ

วัดผลสัมฤทธิ์ โดยใช้ข้อสอบวัดผลสัมฤทธิ์ทางการเรียนในรายวิชาศึกษา 251 การประเมินผลและการสร้างแบบทดสอบตามหลักสูตรประกาศนียบัตรวิชาชีพการศึกษาระดับสูง และหลักสูตรปริญญาครุศาสตรบัณฑิตของสภาการฝึกหัดครู พุทธศักราช 2524 แบบทดสอบมี 3 ฉบับ คือ แบบทดสอบเลือกตอบ ชนิดแบบฉบับ (T.MC) แบบทดสอบเลือกตอบชนิดตัดสินคำตอบทุกตัวเลือกที่ตรวจสอบให้คะแนนแก่ความรู้ที่ถูกต้องสมบูรณ์ (MTF.A) และแบบทดสอบเลือกตอบชนิดตัดสินคำตอบทุกตัวเลือกที่ตรวจสอบให้คะแนนแก่ความรู้บางส่วน (MTF.P) นำไปทดสอบกับนักศึกษาวิทยาลัยครูที่เรียนวิชาศึกษา 251 จำนวน 13 วิทยาลัย จำนวน 1,672 คน ผลการวิจัยพบว่า

1. ข้อสอบในแบบทดสอบ MTF.P ส่วนใหญ่มีประสิทธิภาพสูงกว่าข้อสอบในแบบทดสอบ T.MC ที่ระดับความสามารถสูงมากลงมา ($\theta < 3.50$) ส่วนที่ระดับความสามารถสูงมากขึ้นไป ส่วนใหญ่ข้อสอบในแบบทดสอบ T.MC มีประสิทธิภาพสูงกว่าข้อสอบในแบบทดสอบ MTF.P เมื่อพิจารณาทั้งฉบับพบว่าแบบทดสอบ T.MC มีประสิทธิภาพสูงกว่าแบบทดสอบ MTF.P ที่ระดับความสามารถสูงมากขึ้นไป ($\theta > 3.00$) แบบทดสอบ MTF.P มีประสิทธิภาพสูงกว่าแบบทดสอบ T.MC ที่ระดับความสามารถสูงลงมาในเกือบทุกกรณี ยกเว้นกรณีที่ข้อสอบ MTF.P ประมาณค่าพารามิเตอร์ของตัวเลือกได้เพียง 2 ตัวเลือก แบบทดสอบ T.MC จะมีประสิทธิภาพสูงกว่าแบบทดสอบ MTF.P ที่ระดับความสามารถต่ำขึ้นไป ($\theta > -1.75$)

2. ข้อสอบในแบบทดสอบ MTF.P ส่วนใหญ่มีประสิทธิภาพสูงกว่าข้อสอบในแบบทดสอบ MTF.A ที่ระดับความสามารถสูงมากลงมา ($\theta < 2.75$) ส่วนที่ระดับความสามารถสูงมากขึ้นไปนั้น ส่วนใหญ่ข้อสอบในแบบทดสอบ MTF.A มีประสิทธิภาพสูงกว่าข้อสอบในแบบทดสอบ MTF.P เมื่อพิจารณาทั้งฉบับพบว่าแบบทดสอบ MTF.A มีประสิทธิภาพสูงกว่าแบบทดสอบ MTF.P ที่ระดับความสามารถสูงขึ้นไป ($\theta > 2.00$) ส่วนแบบทดสอบ MTF.P มีประสิทธิภาพสูงกว่าแบบทดสอบ MTF.A ที่ระดับความสามารถสูงลงมาในเกือบทุกกรณี

3. ข้อสอบในแบบทดสอบ MTF.A ส่วนใหญ่มีประสิทธิภาพสูงกว่าข้อสอบในแบบทดสอบ T.MC ที่ระดับความสามารถปานกลางค่อนข้างสูงขึ้นไป ($\theta > 0.75$) ส่วนที่ระดับความสามารถค่อนข้างสูงลงมา ข้อสอบในแบบทดสอบ T.MC มีประสิทธิภาพสูงกว่าข้อสอบในแบบทดสอบ MTF.A เมื่อพิจารณาทั้งฉบับ พบว่า แบบทดสอบ MTF.A ส่วนใหญ่มีประสิทธิภาพสูงกว่าแบบทดสอบ T.MC ที่ระดับความสามารถปานกลางค่อนข้างสูงขึ้นไป ($\theta > 1.00$) ส่วนที่ระดับความสามารถต่ำกว่านี้แบบทดสอบ T.MC ส่วนใหญ่มีประสิทธิภาพสูงกว่าแบบทดสอบ MTF.A

4. ข้อสอบในแบบทดสอบ MTF.P มีประสิทธิภาพสูงที่สุดที่ระดับความสามารถสูงมากลงมา ($\theta < 2.25$) ส่วนที่ระดับความสามารถสูงมากขึ้นไป ข้อสอบในแบบทดสอบ MTF.A มี

ประสิทธิภาพสูงสุด และเมื่อเปรียบเทียบประสิทธิภาพของแบบทดสอบทั้ง 3 แบบพร้อมกันสรุปได้ว่า แบบทดสอบ MTF.P มีประสิทธิภาพสูงกว่าแบบทดสอบ T.MC และแบบทดสอบ MTF.A ที่ระดับความสามารถสูงลงมาส่วนที่ระดับความสามารถสูงขึ้นไปนั้น แบบทดสอบ MTF.A จะมีประสิทธิภาพสูงกว่าแบบทดสอบ MTF.P และแบบทดสอบ T.MC

จิราพร เนียมสุวรรณ (2532 : 85-86) ได้เปรียบเทียบประสิทธิภาพสัมพัทธ์ของแบบทดสอบแบบเลือกตอบที่มีรูปแบบตัวเลือกแตกต่างกัน 3 แบบ คือ แบบตัวเลือกแบบธรรมดา แบบตัวเลือกแบบปลายเปิด "ไม่มีข้อใดถูก" และแบบตัวเลือกผสมระหว่างแบบธรรมดา กับแบบปลายเปิด โดยใช้แบบทดสอบคณิตศาสตร์ ชนิด 5 ตัวเลือก จำนวน 30 ข้อ นำแบบทดสอบทั้ง 3 ฉบับ ไปทดสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 2 สังกัดกรมสามัญศึกษา ในจังหวัดนนทบุรี จำนวน 4,995 คน ผลการวิจัยพบว่า แบบทดสอบแบบตัวเลือกปลายเปิด มีคุณภาพสูงสุดในกลุ่มระดับความสามารถ สูง รองลงไป คือ แบบผสม และแบบธรรมดา แบบทดสอบแบบตัวเลือกแบบธรรมดามีคุณภาพสูงสุดในกลุ่มระดับความสามารถปานกลาง รองลงไป คือ แบบผสม และแบบปลายเปิด ตามลำดับ แบบทดสอบแบบเลือกตอบธรรมดา คุณภาพสูงสุดในกลุ่มระดับความสามารถต่ำ รองลงไป คือ แบบปลายเปิด และแบบผสม ตามลำดับ

ชาติชาย กู้กิตติไมตรี (2533 : 60) ได้ศึกษาคุณภาพของแบบทดสอบแบบเลือกตอบที่มีจำนวนตัวเลือกต่างกัน โดยการใช้ทฤษฎีการตอบข้อคำถาม โดยใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ แบบเลือกตอบ 5 ตัวเลือกจำนวน 40 ข้อ และแยกเป็น 3 ฉบับ โดยในแต่ละฉบับจะมีข้อคำถามเหมือนกันและจำนวนข้อเท่ากัน แต่จำนวนตัวเลือกในแต่ละข้อในแบบทดสอบแต่ละฉบับไม่เท่ากัน โดยมี 5 ตัวเลือก 4 ตัวเลือก และ 3 ตัวเลือก ในการทดสอบหรือตัดตัวลงใช้วิธีการสุ่มตัวลงที่มีค่าอำนาจจำแนกต่ำสุดออกทีละข้อ นำแบบทดสอบทั้ง 3 ฉบับ ไปทดสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในกรุงเทพมหานคร จำนวน 3,880 คน ปรากฏว่า ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบ 3 ฉบับ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าสัมประสิทธิ์การเดาแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยพบว่าแบบทดสอบที่มีจำนวน 5 ตัวเลือกมีค่าสัมประสิทธิ์การเดาต่ำที่สุด และค่าสัมประสิทธิ์การเดาจะเพิ่มขึ้นเมื่อลดจำนวนตัวเลือกลงเหลือ 4 ตัวเลือกและ 3 ตัวเลือก ตามลำดับ สำหรับค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบพบว่าแบบทดสอบที่มีจำนวนตัวเลือก 5 ตัวเลือก มีค่าสูงสุด

วราภรณ์ พูลผลอำนวย (2533 : 131-135) ได้ศึกษาค่าฟังก์ชันอินฟอร์เมชันของข้อสอบและแบบทดสอบของแบบทดสอบที่มีตัวเลือกต่างกัน 5 แบบ คือ แบบตัวเลือกดัก แบบตัวเลือกเนียน แบบตัวเลือกปลายเปิด แบบตัวเลือกแปลก และแบบตัวเลือกให้เดิม โดยใช้แบบทดสอบวัดผลสัมฤทธิ์เชิงวินิจฉัยวิชาคณิตศาสตร์ เรื่อง สมการ ชนิด 4 ตัวเลือก 60 ข้อ นำแบบ

ทดสอบทั้ง 5 ฉบับ ไปทดสอบกับนักเรียน ชั้นมัธยมศึกษาปีที่ 1 สังกัดกรมสามัญศึกษา ใน จังหวัดนครราชสีมา จำนวน 1,965 คน ผลการวิจัยพบว่า

1. ข้อสอบในแบบทดสอบแบบตัวเลือกคัก ส่วนใหญ่ให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่า ในแบบตัวเลือกเฉียด ที่ระดับความสามารถปานกลางขึ้นไป ($\theta > 0$) ส่วนที่ระดับความสามารถ ปานกลางลงไปในั้น ข้อสอบในแบบทดสอบแบบตัวเลือกเฉียดส่วนใหญ่ให้ค่าฟังก์ชันอินฟอร์ เมชันสูงกว่าในแบบตัวเลือกคัก เมื่อพิจารณาทั้งฉบับพบว่า แบบทดสอบแบบตัวเลือกคักให้ค่า ฟังก์ชันอินฟอร์เมชันสูงกว่าแบบทดสอบแบบตัวเลือกเฉียดที่ระดับความสามารถปานกลางขึ้นไป ($\theta > 0$) ส่วนที่ระดับความสามารถปานกลางลงไปในั้น แบบทดสอบแบบตัวเลือกเฉียดให้ค่า ฟังก์ชันอินฟอร์เมชันสูงกว่า แบบทดสอบแบบตัวเลือกคัก

2. ข้อสอบในแบบทดสอบแบบตัวเลือกคัก ส่วนใหญ่ให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่า ในแบบตัวเลือกปลายเปิด ที่ระดับความสามารถปานกลางค่อนข้างสูงลงไป ($\theta < .75$) ส่วนที่ ระดับความสามารถปานกลางค่อนข้างสูงขึ้นไปนั้น ข้อสอบในแบบทดสอบแบบตัวเลือกปลายเปิด ส่วนใหญ่ ให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าในแบบตัวเลือกคัก เมื่อพิจารณาทั้งฉบับ พบว่า แบบทดสอบแบบตัวเลือกคักให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าแบบทดสอบแบบตัวเลือกปลายเปิด ที่ระดับความสามารถปานกลางค่อนข้างสูงลงไป ($\theta < 1$) และที่ระดับสูงมากขึ้นไป ($\theta > 2.75$) ส่วนที่ระดับความสามารถสูงขึ้นไปถึงสูงมาก ($1 \leq \theta < 3$) นั้นแบบตัวเลือกปลายเปิดให้ค่าฟังก์ชัน อินฟอร์เมชันสูงกว่าแบบตัวเลือกคัก

3. ข้อสอบในแบบทดสอบแบบตัวเลือกคักส่วนใหญ่ ให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่า ในแบบตัวเลือกแปลก ที่ระดับความสามารถปานกลางค่อนข้างสูงลงไป ($\theta < 1$) ส่วนที่ระดับ ความสามารถปานกลางค่อนข้างสูงขึ้นไปนั้น ข้อสอบในแบบทดสอบแบบตัวเลือกแปลกส่วน ใหญ่ให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าแบบตัวเลือกคัก เมื่อพิจารณาทั้งฉบับพบว่า แบบทดสอบ แบบตัวเลือกคักให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าในแบบตัวเลือกแปลกที่ระดับความสามารถปาน กลาง ($-0.75 < \theta \leq 1.00$) ระดับต่ำลงไป ($\theta < -1.50$) และที่ระดับสูงมากขึ้นไป ($\theta < 3.25$) ส่วนที่ระดับความสามารถปานกลางลงไปถึงต่ำ ($-1.75 < \theta < -0.50$) และระดับสูงขึ้นไปถึงสูงมาก ($1 < \theta < 3.50$) นั้น แบบทดสอบแบบตัวเลือกแปลก ให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่า แบบ ทดสอบแบบตัวเลือกคัก

4. ข้อสอบในแบบทดสอบแบบตัวเลือกคักส่วนใหญ่ ให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่า ในแบบตัวเลือกให้เดิมที่ระดับความสามารถปานกลางค่อนข้างสูงลงไปถึงระดับต่ำ ($-1.75 < \theta < 0.75$) ส่วนที่ความสามารถระดับต่ำลงไปและระดับปานกลางค่อนข้างสูงขึ้นไปนั้น ข้อสอบใน แบบทดสอบแบบตัวเลือก ให้เดิมส่วนใหญ่ให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าในแบบตัวเลือกคัก

เมื่อพิจารณาทั้งฉบับ พบว่าแบบทดสอบแบบตัวเลือกคัก ให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าแบบทดสอบแบบตัวเลือกให้เติมที่ระดับความสามารถปานกลางลงไป ($\theta < 0$) ส่วนที่ระดับความสามารถปานกลางขึ้นไปนั้น แบบทดสอบแบบตัวเลือกให้เติมให้ค่าฟังก์ชันอินฟอร์เมชันสูงกว่าแบบทดสอบแบบตัวเลือกคัก

งานวิจัยที่เกี่ยวข้องต่างประเทศ

วิลเลียมส์ และอีเบล (Williams and Ebel. 1957 : 63-65) ได้ศึกษาจำนวนตัวเลือกของแบบทดสอบคำศัพท์ โดยวิธีการตัดตัวลงที่มีค่าอำนาจ จำแนกน้อยที่สุดของตัวเลือกของแบบทดสอบ 4 ตัวเลือก ให้เหลือ 3 ตัวเลือก และ 2 ตัวเลือก นำแบบทดสอบทั้ง 3 ชุดไปทดสอบกับเด็กในเวลาที่กำหนด ปรากฏว่าแบบทดสอบ 2 ตัวเลือก ให้ค่าความเชื่อมั่นสูงกว่าแบบทดสอบ 3 ตัวเลือก และ 4 ตัวเลือก

ทเวอร์สกี (Tversky. 1964 : 386-391) ได้ศึกษาแบบทดสอบที่มีจำนวนข้อและจำนวนตัวเลือกที่ไม่เท่ากัน เนื่องจากแบบทดสอบที่มีจำนวนตัวเลือกน้อยกว่านักเรียนจะมีเวลาในการตอบมากกว่าแบบทดสอบที่มีจำนวนตัวเลือกมาก ทำให้มีผลต่อการเปรียบเทียบคุณภาพของแบบทดสอบที่มีจำนวนตัวเลือกไม่เท่ากัน เพื่อเป็นการแก้ไขในเรื่องนี้ ทเวอร์สกีได้ใช้แบบทดสอบที่มีขนาดเท่ากันคือ การให้ผลคูณระหว่างจำนวนข้อกับจำนวนตัวเลือกเท่ากัน ในการศึกษาครั้งนี้ให้ขนาดของแบบทดสอบเท่ากับ 12 แล้วเปรียบเทียบแบบทดสอบที่มี 6 ข้อ 2 ตัวเลือก 4 ข้อ 3 ตัวเลือก, 3 ข้อ 4 ตัวเลือก และ 1 ข้อ 12 ตัวเลือก ปรากฏว่า แบบทดสอบที่มีความสามารถในการจำแนก ได้แก่แบบทดสอบที่มี 4 ข้อ 3 ตัวเลือก

คอสดิน (Costin. 1970 : 353-358) ได้ศึกษาจำนวนตัวเลือกที่เหมาะสมของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนจิตวิทยาเบื้องต้น ซึ่งแบ่งเนื้อหา วิชาออกเป็น 4 เรื่อง คือ การรับรู้ (Perception) การเรียนรู้ (Learning) การจูงใจ (Motivation) และเชาวน์ปัญญา (Intelligence) จำนวนข้อในแบบทดสอบในแต่ละเรื่องมีไม่เท่ากัน คือ การรับรู้ และเชาวน์ปัญญามีอย่างละ 50 ข้อ ส่วนการเรียนรู้ และการจูงใจ มีอย่างละ 60 ข้อ เนื่องจากจำนวนชั่วโมงเรียนไม่เท่ากัน ในการศึกษาผู้วิจัยได้สุ่มแบบทดสอบที่แยกตามเรื่องออกมาทีละชุดและในแต่ละชุด แบ่งครึ่งจำนวนข้อสอบของแบบทดสอบออกเป็น 2 ส่วน ส่วนแรกมีจำนวนตัวเลือก 4 ตัวเลือก ส่วนหลังให้มีตัวเลือกเพียง 3 ตัวเลือก โดยใช้วิธีการสุ่มตัดตัวลงทั้งหนึ่งตัวแล้วนำแบบทดสอบทั้ง 4 ชุด ไปทดสอบกับนักศึกษาของมหาวิทยาลัยอิลลินอยส์ที่เรียนวิชานี้ จำนวน 200 คน โดยทดสอบในเทอมแรกกับนักศึกษา 2 ห้อง และเทอมหลัง 2 ห้อง ปรากฏว่าค่าเฉลี่ยของค่าอำนาจจำแนก และ

ค่าความยากของแบบทดสอบที่มี 2 ตัวเลือก มีค่าสูงกว่าแบบทดสอบที่มี 4 ตัวเลือกและค่าความเชื่อมั่นของแบบทดสอบที่มี 3 ตัวเลือกสูงกว่าแบบทดสอบที่มี 4 ตัวเลือก

บูเดสคู และเนโว (Budescu and Nevo. 1985 : 183-196) ได้ศึกษาจำนวนตัวเลือกที่เหมาะสมเพื่อหาข้อสรุปเกี่ยวกับสัดส่วนของเวลาในการทำข้อสอบกับตัวข้อสอบโดยใช้กลุ่มตัวอย่างจำนวน 1,018 คน ที่สอบเข้ามหาวิทยาลัยไฮฟา (The University of Haifa) โดยใช้แบบทดสอบ 3 ฉบับ คือ แบบทดสอบคำศัพท์ 60 ข้อ แบบทดสอบเหตุผลทางคณิตศาสตร์ 25 ข้อ และแบบทดสอบความเข้าใจในการอ่าน 24 ข้อ โดยครั้งแรกใช้แบบทดสอบ 5 ตัวเลือกแล้วนำมาแบ่งออกเป็นแบบทดสอบในสถานการณ์ต่าง ๆ 6 แบบ คือ แบบทดสอบที่มี 4 ตัวเลือก 3 ตัวเลือก และ 2 ตัวเลือก โดยการสุ่มตัวดวงตัวสุดท้ายออกจากแบบทดสอบ และวิธีการตัดตัวดวงที่มีค่าอำนาจจำแนกต่ำที่สุดออก ผลการศึกษาพบว่า แบบทดสอบทั้ง 3 ฉบับ ที่มีจำนวนตัวเลือกมากจะให้ค่าความแปรปรวน ค่าความเชื่อมั่น และค่าอำนาจจำแนกสูงกว่าแบบทดสอบที่มีจำนวนตัวเลือกน้อย เมื่อหาเวลาเฉลี่ยที่ใช้ตอบต่อตัวเลือก ปรากฏว่าส่วนใหญ่เป็นไปตามข้อสันนิษฐานของลอว์รี (Budescu and Nevo. 1985 : 183 ; Citing Lord. 1977, 1980) ดังสมการ

$$T = nat$$

เมื่อ T แทน เวลาทั้งหมดในการทำแบบทดสอบ

n แทน จำนวนข้อสอบ

a แทน จำนวนตัวเลือกในแต่ละข้อ

t แทน เวลาที่ใช้ตอบต่อตัวเลือก

ดังนั้น ถ้าเวลาทั้งหมดในการทำแบบทดสอบคงที่ แบบทดสอบที่มีจำนวนตัวเลือกมากกว่าจะใช้เวลาในการตอบต่อตัวเลือกน้อยกว่า ยกเว้นแบบทดสอบความเข้าใจในการอ่านที่ตัดตัวสุดท้ายทิ้ง พบว่าแบบทดสอบที่มี 4 ตัวเลือก ใช้เวลาเฉลี่ยในการตอบต่อตัวเลือกมากกว่าแบบทดสอบที่มี 3 ตัวเลือก และ 2 ตัวเลือก เมื่อกำหนดเวลาในการตอบไว้ในช่วง 0-10 นาที

จากการศึกษาวิจัยที่กล่าวมาข้างต้น จะเห็นได้ว่ามีความขัดแย้งกันในเรื่องค่าความเชื่อมั่น ค่าอำนาจจำแนก และค่าความยากของแบบทดสอบแบบเลือกตอบที่มีจำนวนตัวเลือกไม่เท่ากัน และแบบทดสอบที่มีความยาวไม่เท่ากัน ซึ่งเป็นคุณลักษณะสำคัญในการตรวจสอบคุณภาพของแบบทดสอบ โดยคุณลักษณะที่ได้นี้ส่วนมากเป็นการศึกษาตามแนวทฤษฎีคลาสสิกอล ซึ่งไม่สามารถสรุปได้อย่างชัดเจนว่าแบบทดสอบแบบเลือกตอบฉบับใดที่มีคุณภาพสูงกว่า เมื่อนำมา

เปรียบเทียบกับเพราะเป็นการ เปรียบเทียบคุณลักษณะที่สำคัญของการตรวจสอบคุณภาพของแบบทดสอบทีละประเด็นซึ่งในการศึกษาประสิทธิภาพสัมพัทธ์และคุณภาพของแบบทดสอบที่มีจำนวนข้อและจำนวนตัวเลือกไม่เท่ากัน ในทฤษฎีการตอบข้อคำถามยังไม่มีผู้ทำวิจัยมาก่อน ดังนั้นผู้วิจัยจึงมีความสนใจที่จะศึกษาเกี่ยวกับเรื่องนี้

สมมติฐานในการศึกษาค้นคว้า

1. ประสิทธิภาพสัมพัทธ์ของแบบทดสอบฉบับที่ 1 มี 60 ข้อ 3 ตัวเลือก กับฉบับที่ 2 มี 45 ข้อ 4 ตัวเลือก มีค่าสูงกว่า 1 ทุกระดับความสามารถ
2. ประสิทธิภาพสัมพัทธ์ของแบบทดสอบฉบับที่ 1 มี 60 ข้อ 3 ตัวเลือก กับฉบับที่ 3 มี 36 ข้อ 5 ตัวเลือก มีค่ามากกว่า 1 ทุกระดับความสามารถ
3. ประสิทธิภาพสัมพัทธ์ของแบบทดสอบฉบับที่ 2 มี 45 ข้อ 4 ตัวเลือก กับฉบับที่ 3 มี 36 ข้อ 5 ตัวเลือก มีค่ามากกว่า 1 ทุกระดับความสามารถ
4. แบบทดสอบฉบับที่ 1 มี 60 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 45 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 36 ข้อ 5 ตัวเลือก มีค่าความยากของข้อสอบแตกต่างกัน
5. แบบทดสอบ ฉบับที่ 1 มี 60 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 36 ข้อ 5 ตัวเลือก มีค่าอำนาจจำแนกของข้อสอบแตกต่างกัน
6. แบบทดสอบ ฉบับที่ 1 มี 60 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 36 ข้อ 5 ตัวเลือก มีค่าสัมประสิทธิ์การเดาของข้อสอบแตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาค เรียนที่ 1 ปีการศึกษา 2537 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ ใน จังหวัดสมุทรสาคร จำนวน 112 โรงเรียน จำนวนนักเรียน 5,246 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 1 ปีการศึกษา 2537 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ ใน จังหวัดสมุทรสาครจำนวนประมาณ 3,016คน โดยการสุ่มตัวอย่างแบบหลายขั้นตอน (Multistage Random Sampling) ซึ่งมีลำดับขั้นดังต่อไปนี้

ขั้นที่ 1 แบ่งโรงเรียนออกเป็น 3 ขนาด คือ โรงเรียนขนาดใหญ่ โรงเรียนขนาดกลาง และโรงเรียนขนาดเล็ก ตามเกณฑ์ของสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ ได้ดังนี้

1. โรงเรียนขนาดใหญ่มี 6 โรงเรียน มีจำนวนนักเรียน 1,438 คน
2. โรงเรียนขนาดกลางมี 35 โรงเรียน มีจำนวนนักเรียน 2,131 คน
3. โรงเรียนขนาดเล็กมี 71 โรงเรียน มีจำนวนนักเรียน 1,677 คน

ขั้นที่ 2 สุ่มโรงเรียนขนาดใหญ่มา 5 โรงเรียน โรงเรียนขนาดกลางมา 21 โรงเรียน และโรงเรียนขนาดเล็กมา 29 โรงเรียน ด้วยวิธีการสุ่มอย่างง่ายได้ จำนวนโรงเรียน 55 โรงเรียน จำนวนนักเรียน 3,016 คน ดังแสดงในตารางที่ 1

ขั้นที่ 3 สุ่มนักเรียนออกเป็น 3 กลุ่ม ด้วยวิธีการสุ่มอย่างง่ายซึ่งจะได้นักเรียนกลุ่มละประมาณ 1,005 คน

ตารางที่ 1 โรงเรียน อำเภอ และจำนวนนักเรียนที่เป็นกลุ่มตัวอย่างของการวิจัย

โรงเรียน	อำเภอ	จำนวนนักเรียน
ขนาดใหญ่		
วัดบางหญ้าแพรก	เมืองสมุทรสาคร	120
เอกชัย	เมืองสมุทรสาคร	244
อนุบาลสมุทรสาคร	เมืองสมุทรสาคร	110
วัดศรีสำราญราษฎร์บำรุง	กระทุ่มแบน	243
วัดอ้อมน้อย	กระทุ่มแบน	212
ขนาดกลาง		
วัดใหม่จอมปราสาท	เมืองสมุทรสาคร	46
วัดบางปิ้ง	เมืองสมุทรสาคร	68
สหกรณ์นิคมเกลือ	เมืองสมุทรสาคร	56
วัดโสภณาราม	เมืองสมุทรสาคร	87
วัดคลองครุ	เมืองสมุทรสาคร	60
วัดราษฎร์รังสรรค์	เมืองสมุทรสาคร	82
วัดโพธิ์แจ้ง	เมืองสมุทรสาคร	90
วัดท่าเสา	กระทุ่มแบน	50
วัดราษฎร์บำรุง	กระทุ่มแบน	42
บ้านคลองแค	กระทุ่มแบน	51
วัดนางสาว	กระทุ่มแบน	109
บ้านคอนไผ่	บ้านแพ้ว	48
วัดใหม่หนองพระอง	บ้านแพ้ว	47
วัดราษฎร์ศรัทธากระขาราม	บ้านแพ้ว	56
วัดกกกระบัตร์	บ้านแพ้ว	31
วัดบ้านแคลาข	บ้านแพ้ว	22
วัดหลักสองราษฎร์บำรุง	บ้านแพ้ว	57

ตาราง 1 (ต่อ)

โรงเรียน	อำเภอ	จำนวนนักเรียน
วัดเจ็ดริ้ว	บ้านแพ้ว	60
วัดหลักสี่ราษฎร์สโมสร	บ้านแพ้ว	120
วันครู 2500	บ้านแพ้ว	112
ทำนบแพ้ว	บ้านแพ้ว	43
<u>ขนาดเล็ก</u>		
วัดชีผ้าขาว	เมืองสมุทรสาคร	17
วัดเจริญสุขาราม	เมืองสมุทรสาคร	15
เทพนรรัตน์	เมืองสมุทรสาคร	23
หลวงแพทย์โกศลอุปถัมภ์	เมืองสมุทรสาคร	14
วัดบางปลา	เมืองสมุทรสาคร	52
บ้านอ้อมโรงหีบ	เมืองสมุทรสาคร	23
วัดชัยมงคล	เมืองสมุทรสาคร	26
บ้านยกบตร	เมืองสมุทรสาคร	14
วัดสามัคคีศรีธาราม	เมืองสมุทรสาคร	41
วัดศรีสุนธาราม	เมืองสมุทรสาคร	29
บ้านบางน้ำจืด	เมืองสมุทรสาคร	33
เมืองสมุทรสาคร	เมืองสมุทรสาคร	35
วัดใหญ่บ้านป้อ	เมืองสมุทรสาคร	26
วัดบ้านไร่	เมืองสมุทรสาคร	11
วัดบางพลี	เมืองสมุทรสาคร	12
วัดเกตุมดีศรีวราราม	เมืองสมุทรสาคร	26
วัดนาขวาง	เมืองสมุทรสาคร	6
วัดศรีเมือง	เมืองสมุทรสาคร	43
วัดกาหลง	เมืองสมุทรสาคร	40

ตาราง 1 (ต่อ)

โรงเรียน	อำเภอ	จำนวนนักเรียน
วัดอ่างทอง	กระทุ่มแบน	35
วัดสุวรรณรัตนาราม	กระทุ่มแบน	27
บ้านคลองกระทุ่มแบน	กระทุ่มแบน	43
วัดคอนโฆสิตาราม	บ้านแพ้ว	16
บ้านวังจรเข้	บ้านแพ้ว	23
วัดหนองบัว	บ้านแพ้ว	44
วัดรางต้นนิคประดิษฐ์	บ้านแพ้ว	9
วัดกระโจมทอง	บ้านแพ้ว	21
วัดธรรมเจดีย์ศรีพิพัฒน์	บ้านแพ้ว	21
ไทยรัฐวิทยา ๔	บ้านแพ้ว	25
รวม 55		3,016

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

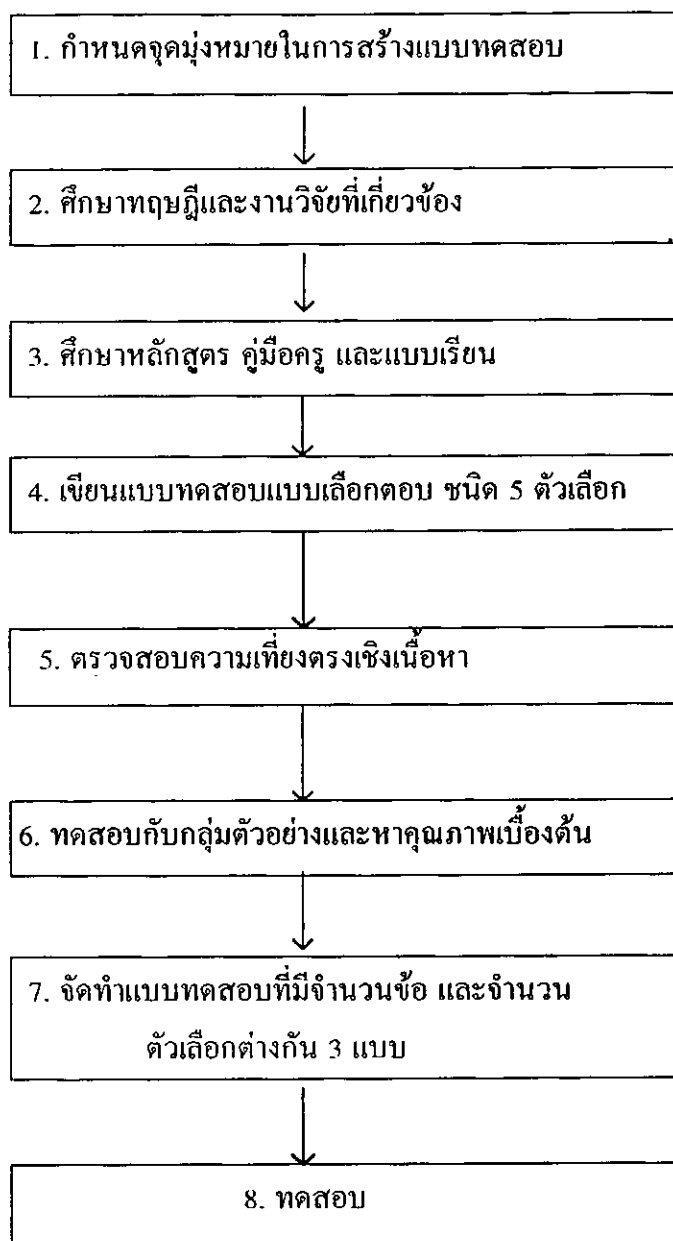
เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่อง ตัวประกอบของ จำนวนนับ และเศษส่วนที่ผู้วิจัยสร้างขึ้นเองตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ.2533) จำนวน 3 ฉบับ ดังนี้

1. ฉบับที่ 1 มี 60 ข้อ 3 ตัวเลือก
2. ฉบับที่ 2 มี 45 ข้อ 4 ตัวเลือก
3. ฉบับที่ 3 มี 36 ข้อ 5 ตัวเลือก

วิธีดำเนินการสร้างแบบทดสอบ

การสร้างแบบทดสอบสำหรับการศึกษาค้นคว้าครั้งนี้ ได้ดำเนินการสร้างตามขั้นตอน ดัง

ภาพประกอบ 5



ภาพประกอบ 5 ลำดับขั้นตอนการสร้างเครื่องมือที่ใช้ในการวิจัย

รายละเอียดขั้นตอนการสร้างแบบทดสอบ

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ เพื่อสร้างแบบทดสอบแบบเลือกตอบ

วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เพื่อเปรียบเทียบประสิทธิภาพสัมพันธ์ และคุณภาพของแบบทดสอบแบบเลือกตอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกัน

2. ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการสร้างแบบทดสอบแบบเลือกตอบทั้งของไทยและต่างประเทศ เพื่อเป็นแนวทางที่จะสร้างแบบทดสอบแบบเลือกตอบ ที่มีจำนวนข้อและจำนวนตัวเลือกต่างกัน

3. ศึกษาหลักสูตร คู่มือครู และแบบเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ.2533) เพื่อกำหนดจุดประสงค์ตามขอบเขตเนื้อหาวิชาในการสร้างแบบทดสอบครั้งนี้

4. เขียนแบบทดสอบแบบเลือกตอบชนิด 5 ตัวเลือก จำนวน 100 ข้อ ตามจุดประสงค์เชิงพฤติกรรมและเนื้อหาในขอบเขต

5. นำแบบทดสอบที่สร้างขึ้นไปให้ผู้เชี่ยวชาญทางเนื้อหา ตรวจสอบความเที่ยงตรงเชิงเนื้อหา และแก้ไขข้อบกพร่องของแบบทดสอบที่สร้างขึ้น

6. นำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่างที่สุ่มมาสำหรับทดลองเครื่องมือซึ่งเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 1 ปีการศึกษา 2537 ของโรงเรียน เอกชัย อำเภอมือง จังหวัดสมุทรสาคร จำนวน 219 คน แล้วนำผลไปวิเคราะห์หาค่าความยากและค่าอำนาจจำแนกของข้อสอบ โดยใช้เทคนิค 27 เปอร์เซนต์ และตารางสำเร็จรูปของ จุง เต ฟาน (Chung-teh-Fan) เพื่อคัดเลือกข้อสอบไว้ 60 ข้อ ซึ่งคำตอบถูกมีค่าความยากระหว่าง 0.20 - 0.83 และค่าอำนาจจำแนกระหว่าง 0.31 - 0.75

7. นำข้อสอบที่ไปทดลองมาแล้ว เลือกข้อสอบที่มีคุณภาพ 60 ข้อ ซึ่งแต่ละข้อมี 5 ตัวเลือก มาจัดทำเป็นแบบทดสอบ 3 ฉบับ ดังนี้

ฉบับที่ 1 แบบทดสอบ 60 ข้อ 3 ตัวเลือก ซึ่งได้จากการนำแบบทดสอบ 60 ข้อ 5 ตัวเลือก ที่คัดไว้แล้วมาตัดตัวลงของแต่ละข้อออก 2 ตัว โดยตัดตัวลงที่มีค่าอำนาจจำแนกน้อยที่สุดและรองลงมาออก

ฉบับที่ 2 แบบทดสอบ 45 ข้อ 4 ตัวเลือก ซึ่งได้จากการนำแบบทดสอบ 60 ข้อ 5 ตัวเลือก ที่คัดไว้แล้วมาตัดข้อสอบที่มีค่าอำนาจจำแนกน้อยที่สุดตามลำดับ ออกจำนวน 15 ข้อ ข้อสอบที่เหลือแต่ละข้อตัดตัวลงที่มีค่าอำนาจจำแนกน้อยที่สุดออก 1 ตัว

ฉบับที่ 3 แบบทดสอบ 36 ข้อ 5 ตัวเลือก ซึ่งได้จากการนำแบบทดสอบ 60 ข้อ 5 ตัวเลือก ที่คัดไว้แล้วมาตัดข้อสอบที่มีค่าอำนาจจำแนกน้อยที่สุดตามลำดับออก จำนวน 24 ข้อ

8. นำแบบทดสอบทั้ง 3 ฉบับ ไปทดสอบกับกลุ่มตัวอย่างที่สุ่มมาเพื่อการวิจัยตามที่แบ่งไว้ 3 กลุ่ม ให้แต่ละกลุ่มทำแบบทดสอบกลุ่มละ 1 ฉบับ นักเรียนแต่ละคนทำแบบทดสอบคนละ 1 ฉบับ

ตัวอย่างข้อคำถาม

แบบทดสอบทั้ง 3 ฉบับ มีจำนวนข้อและจำนวนตัวเลือกต่างกันดังตัวอย่างต่อไปนี้

ฉบับที่ 1 แบบทดสอบ จำนวน 60 ข้อ 3 ตัวเลือก

0) ห.ร.ม. ของเลขชุดใดมีค่ามากที่สุด

ก. 14, 21, 35

ข. 18, 27, 36

ค. 27, 57, 63 เฉลย ข.

ฉบับที่ 2 แบบทดสอบ จำนวน 45 ข้อ 4 ตัวเลือก

0) ห.ร.ม. ของเลขชุดใดมีค่ามากที่สุด

ก. 12, 18, 24

ข. 14, 21, 35

ค. 18, 27, 36

ง. 27, 57, 63 เฉลย ก.

ฉบับที่ 3 แบบทดสอบ จำนวน 36 ข้อ 5 ตัวเลือก

0) ห.ร.ม. ของเลขชุดใดมีค่ามากที่สุด

ก. 12, 18, 24

ข. 14, 21, 35

ค. 18, 27, 36

ง. 24, 40, 56

จ. 27, 57, 63 เฉลย ก.

วิธีดำเนินการเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลเพื่อการศึกษาครั้งนี้ ผู้วิจัยดำเนินการดังนี้

1. วางแผนในการดำเนินการสอบ โดยติดต่อขอความร่วมมือในการเก็บข้อมูลจาก ผู้อำนวยการการประถมศึกษาจังหวัด และหัวหน้าการประถมศึกษาอำเภอในจังหวัดสมุทรสาคร และติดต่อโรงเรียนที่เป็นกลุ่มตัวอย่าง เพื่อกำหนดวัน เวลาในการดำเนินการสอบ

2. จัดเตรียมเครื่องมือที่ใช้ในการศึกษาค้นคว้า ในที่นี้คือแบบทดสอบแบบเลือกตอบวิชาคณิตศาสตร์ จำนวน 3 ฉบับ คือ

2.1 ฉบับที่ 1 มี 60 ข้อ 3 ตัวเลือก

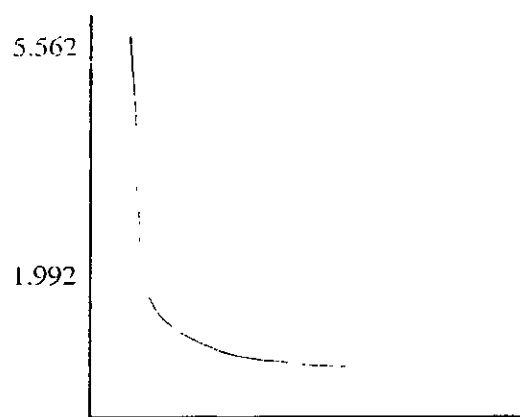
2.2 ฉบับที่ 2 มี 45 ข้อ 4 ตัวเลือก

2.3 ฉบับที่ 3 มี 36 ข้อ 5 ตัวเลือก

3. ผู้วิจัยนำแบบทดสอบทั้ง 3 ฉบับ ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าตามที่แบ่งไว้ 3 กลุ่ม ๆ ละเท่า ๆ กัน โดยให้นักเรียนแต่ละคนในแต่ละกลุ่ม ทำแบบทดสอบคนละ 1 ฉบับ

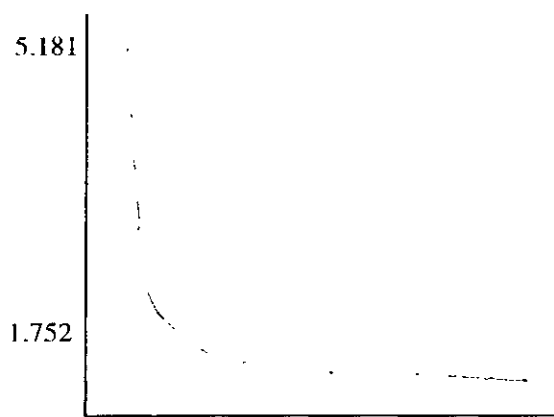
4. อธิบายให้นักเรียนในกลุ่มตัวอย่าง เข้าใจในวัตถุประสงค์ในการสอบและขอความร่วมมือในการสอบ เพื่อให้ได้ผลตรงตามความเป็นจริง

5. นำผลการสอบมาตรวจให้คะแนน แล้ววิเคราะห์แบบทดสอบด้วยการวิเคราะห์องค์ประกอบ เนื่องจากตามข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม ได้กำหนดไว้ว่า แบบทดสอบที่จะใช้วิเคราะห์ด้วยโมเดลต่าง ๆ ในกรอบทฤษฎีนี้ จะต้องมีความเป็นมิติตีเดียว (Hambleton and Swaminathan.1985 : 16 - 17) และสามารถทำการตรวจสอบความเป็นมิติตีเดียวได้ โดยการวิเคราะห์องค์ประกอบ เพื่อสังเกตค่าไอเกน ขององค์ประกอบที่มีค่าสูงสุดว่า แตกต่างจากค่าอื่น ๆ อย่างชัดเจนหรือไม่ ในขณะที่ค่าไอเกนขององค์ประกอบที่เหลือทั้งหมดมีค่าใกล้เคียงกัน และถ้าแบบทดสอบมีความเป็นมิติตีเดียวแล้ว แบบทดสอบนั้นจะมีความเป็นอิสระในการตอบข้อสอบ (Warm. 1978 : 104 ; Lord. 1980 : 21) จากผลการวิเคราะห์พบว่า แบบทดสอบทั้ง 3 ฉบับ มีความเป็นมิติตีเดียว และมีความเป็นอิสระในการตอบข้อสอบได้เสนอในรูปกราฟดังภาพประกอบ 6 , 7 มีรายละเอียดในตาราง 10 ภาคผนวก



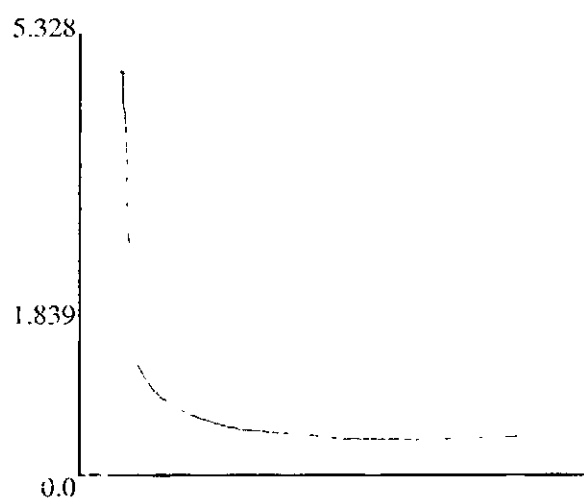
0.0

ฉบับที่ 1 มี 60 ข้อ 3 ตัวเลือก



0.0

ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก



ฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก

ภาพประกอบ 6 ค่าไอเกนขององค์ประกอบของแบบทดสอบทั้ง 3 ฉบับ

จากภาพประกอบ 6 แสดงให้เห็นว่าค่าไอเกนขององค์ประกอบที่ 1 ของแบบทดสอบทั้ง 3 ฉบับ สูงกว่าองค์ประกอบที่ 2 , 3 , 4 และองค์ประกอบที่ 2 , 3 , 4 มีค่าไอเกนใกล้เคียงกัน แสดงว่าแบบทดสอบทั้ง 3 ฉบับ มีองค์ประกอบเด่นเพียง 1 องค์ประกอบ จึงสรุปได้ว่าแบบทดสอบมีความเป็นมิติเดียวและมีความเป็นอิสระในการตอบข้อสอบ

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. สถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย (Mean) ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)
2. วิเคราะห์คุณภาพของเครื่องมือโดยใช้เทคนิค 27 เปอร์เซนต์ และตารางสำเร็จรูปของ จุง เต ฟาน (Chung-Teh-Fan) หาค่าความยาก (P) และค่าอำนาจจำแนก (r)
3. การหาค่าพารามิเตอร์ของข้อสอบ ตามทฤษฎีการตอบข้อคำถามคือ ค่าความยาก ค่าอำนาจจำแนก และค่าสัมประสิทธิ์การเดา หาด้วยโปรแกรมสำเร็จรูป ไบล็อก (Bilog)
4. ฟังก์ชันอินฟอร์เมชันของข้อสอบและแบบทดสอบ โดยใช้สูตรของเบิร์นบอม (ผจจ จิต อินทสุวรรณ. 2528 : 27-29 ; อ้างอิงมาจาก Birnbom. 1968)

สูตร ฟังก์ชันอินฟอร์เมชันของข้อสอบ

$$I(\theta, u_i) = \frac{(P'_i)^2}{P_i Q_i} \quad ; i = 1, 2, 3, \dots, n$$

เมื่อ $I(\theta, u_i)$ คือ ฟังก์ชันอินฟอร์เมชันของข้อสอบข้อที่ i

P_i คือ ความน่าจะเป็นที่ผู้ที่มีความสามารถ θ จะตอบข้อสอบข้อที่ i ถูก

Q_i คือ $1 - P_i$

P'_i คือ ความชันของโค้งลักษณะข้อสอบที่ระดับความสามารถ θ

สูตร ฟังก์ชันอินฟอร์เมชันของแบบทดสอบ

$$I(\theta) = \sum_{i=1}^n I(\theta, u_i) \quad ; i = 1, 2, 3, \dots, n$$

เมื่อ $I(\theta)$ คือ ฟังก์ชันอินฟอร์เมชันของแบบทดสอบ

5. การคำนวณค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่า ความสามารถระหว่างแบบทดสอบที่แตกต่างกัน 3 รูปแบบ โดยนำค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ ณ ระดับความสามารถต่าง ๆ 81 ระดับ มาเปรียบเทียบกับทีละคู่ โดยใช้สูตรในการคำนวณ ดังนี้ (Lord and Novick. 1968 : 472)

$$I(\theta, x_1, x_2) = \frac{I(\theta, x_1)}{I(\theta, x_2)}$$

เมื่อ $RE(\theta, x_1, x_2)$ แทน ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถ
ระหว่างแบบทดสอบรูปแบบที่ 1 (x_1) เทียบกับแบบทดสอบ
รูปแบบที่ 2 (x_2)

$I(\theta, x_1)$ แทน ค่าอินฟอร์เมชันของแบบทดสอบรูปแบบที่ 1 (x_1) ณ ระดับ
ความสามารถ θ

$I(\theta, x_2)$ แทน ค่าอินฟอร์เมชันของแบบทดสอบรูปแบบที่ 2 (x_2) ณ ระดับ
ความสามารถ θ

6. การทดสอบความแตกต่างค่าพารามิเตอร์ของข้อสอบโดยการวิเคราะห์ความแปรปรวนแบบทิศทางเดียว (One - Way Anova) และถ้าพบความแตกต่างอย่างมีนัยสำคัญทางสถิติ จะตรวจสอบความแตกต่างรายคู่ ด้วยการเปรียบเทียบพหุคูณ โดยวิธี เชฟเฟ (Scheff's Test) (ล้วน
สาขศ และอังคณา สาขาศ. 2528 : 95-119)

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

$$S = \sqrt{(K-1) F(\alpha, df_1, df_2)} \cdot \sqrt{MS_E \sum_{i=1}^k \frac{c_i^2}{n_i}}$$

เมื่อ MS_b คือ ความแปรปรวนระหว่างกลุ่ม

MS_w คือ ความแปรปรวนภายในกลุ่ม

F คือ ค่าการแจกแจงของ F ที่ α

S คือ ค่าวิกฤตของ Scheffe's

K คือ จำนวนกลุ่มตัวอย่างที่นำมาเปรียบเทียบกัน

$F(\alpha, df_1, df_2)$ คือ ค่าจากตาราง F

MS_E คือ ค่าความคลาดเคลื่อนของความแปรปรวนในตารางวิเคราะห์
ความแปรปรวน

C_i คือ สัมประสิทธิ์ Contrast ซึ่งมีค่า 1, -1, 1, -1

n_i คือ จำนวนคนในกลุ่มตัวอย่างแต่ละกลุ่ม

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการศึกษาครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามความมุ่งหมายและสมมติฐานของการศึกษากันคว่ำตามข้อต่อไปนี้

1. ค่าพารามิเตอร์ของข้อสอบ
2. สถิติพื้นฐานของค่าพารามิเตอร์ของข้อสอบ
3. ผลการศึกษาฟังก์ชันอินฟอร์เมชันของแบบทดสอบ
4. ผลการหาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ
5. ผลการทดสอบความแตกต่างของค่าความยากของข้อสอบ
6. ผลการทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบ
7. ผลการทดสอบความแตกต่างของค่าสัมประสิทธิ์การเดาของข้อสอบ

ผลการวิเคราะห์ข้อมูล มีรายละเอียดดังนี้

1. ค่าพารามิเตอร์ของข้อสอบ

จากการนำแบบทดสอบทั้ง 3 ฉบับ คือ ฉบับที่ 1 มี 60 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 45 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 36 ข้อ 5 ตัวเลือก ไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างจำนวน 3,016 คน โดยทดสอบกับนักเรียน จำนวน 1,006 คน, จำนวน 1,006 คน และ จำนวน 1,004 คน ตามลำดับ แล้วนำผลการสอบไปวิเคราะห์องค์ประกอบของการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว และความเป็นอิสระในการตอบข้อสอบ ซึ่งปรากฏผลดังตาราง 10 ในภาคผนวก ก. ในการหาค่าพารามิเตอร์ของข้อสอบได้นำผลการสอบไปวิเคราะห์ตามทฤษฎีการตอบข้อคำถามโดยใช้โปรแกรมสำเร็จรูป ไบล็อก (Bilog) Version 3.04 ซึ่งเมื่อทดสอบค่า Correlation biserial ของแบบทดสอบทั้ง 3 ฉบับ พบว่า ฉบับที่ 1 ข้อ 25 มีค่าติดลบ ฉบับที่ 2 ข้อ 11 มีค่าติดลบ และฉบับที่ 3 ข้อ 10 มีค่าติดลบ ซึ่งทำให้ไม่สามารถจะวิเคราะห์หาค่าพารามิเตอร์ได้ ดังนั้นจึงต้องตัดข้อเหล่านี้ออก จึงจะสามารถหาค่าพารามิเตอร์ a, b, c ของแบบทดสอบแต่ละแบบได้ ดังนั้น ฉบับที่ 1, ฉบับที่ 2, ฉบับที่ 3 เหลือข้อสอบ 59 ข้อ, 44 ข้อ, 35 ข้อ ตามลำดับ ดังแสดงในตาราง 11, 12, 13 ภาคผนวก ข.

2. สถิติพื้นฐานของค่าพารามิเตอร์ของข้อสอบแต่ละแบบ

ผลการวิเคราะห์ด้วยโปรแกรมสำเร็จรูป ใบล็อก ได้ค่าสถิติพื้นฐานของค่าพารามิเตอร์ ดังแสดงในตาราง 2 มีรายละเอียด ในตาราง 11,12 และ 13 ในภาคผนวก ข

ตาราง 2 ค่าสถิติพื้นฐานของความยาก อำนาจจำแนก และสัมประสิทธิ์การเดาของ แบบทดสอบ ทั้ง 3 ฉบับ

ค่าพารามิเตอร์	59 ข้อ 3 ตัวเลือก		44 ข้อ 4 ตัวเลือก		35 ข้อ 5 ตัวเลือก	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
ค่าความยาก	1.042	1.111	.951	1.224	.932	1.047
ค่าอำนาจจำแนก	1.077	.423	.982	.330	.997	.264
ค่าสัมประสิทธิ์การเดา	.318	.063	.239	.065	.177	.056

จากตาราง 2 แสดงให้เห็นว่า ค่าความยากเฉลี่ยของแบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีค่าเท่ากับ 1.042 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 1.111 ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีค่าเท่ากับ .951 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 1.224 และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าเท่ากับ .932 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.047

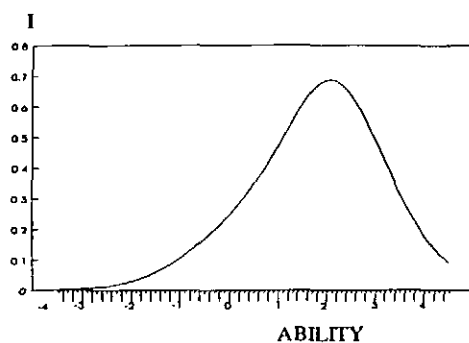
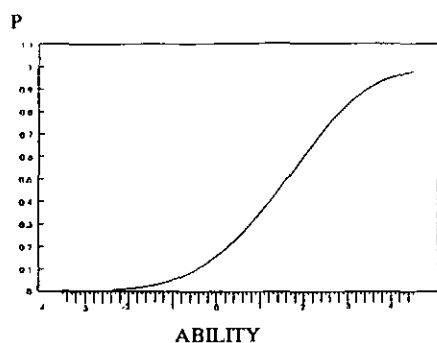
ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีค่าเท่ากับ 1.077 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ .423 ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีค่าเท่ากับ .982 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ .330 และ ฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าเท่ากับ .997 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ .264

ค่าสัมประสิทธิ์การเดาเฉลี่ยของแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีค่าเท่ากับ .318 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ .063 ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีค่าเท่ากับ .239 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ .065 และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าเท่ากับ .177 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ .056

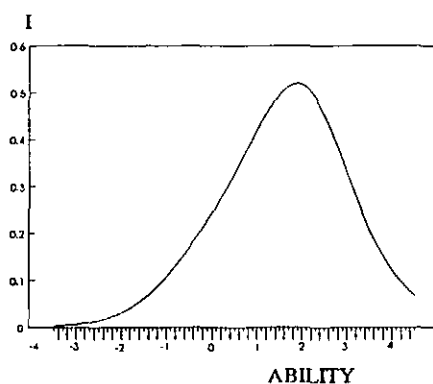
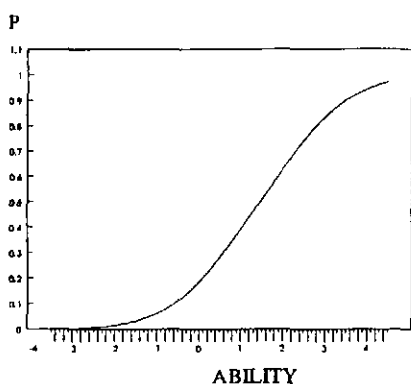
3. ผลการศึกษาค่าอินฟอร์เมชันของแบบทดสอบทั้ง 3 ฉบับ

ค่าอินฟอร์เมชันของแบบทดสอบตามทฤษฎีการตอบข้อคำถามถือว่ามีความเชื่อมั่นของแบบทดสอบตามทฤษฎีการทดสอบแบบคลาสสิกอล (Classical test theory) สำหรับค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบในการศึกษารั้งนี้ ผู้วิจัยคิดจากค่าเฉลี่ยของแบบทดสอบแต่ละระดับความสามารถ มีรายละเอียดในตาราง 12 ภาคผนวก

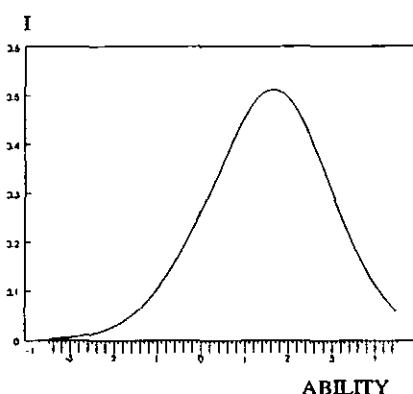
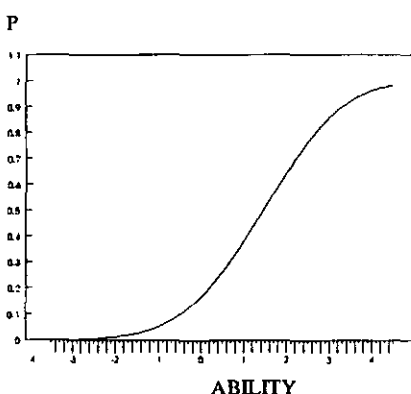
และได้นำผลการวิเคราะห์ความน่าจะเป็นในการตอบแบบทดสอบถูก (P) ณ ระดับความสามารถต่างๆของแบบทดสอบทั้ง 3 ฉบับมาเขียนโค้งลักษณะของแบบทดสอบ โดยค่า P หมายถึงสัดส่วนของ $\frac{\Sigma P(\theta)}{I}$ โดยที่ $\Sigma P(\theta)$ เป็นผลรวมของความน่าจะเป็นในการตอบข้อสอบแต่ละข้อได้ถูก ณ θ ใดๆ I หมายถึงจำนวนข้อในแบบทดสอบ มีรายละเอียดในตาราง 14 ภาคผนวก และได้นำความน่าจะเป็นในการตอบแบบทดสอบทั้ง 3 ฉบับมาเปรียบเทียบกัน เสนอในรูปกราฟดังภาพประกอบ 7 โดยเฉพาะอินฟอร์เมชันของแบบทดสอบทั้ง 3 ฉบับ เสนอในรูปกราฟที่นำมาซ้อนกันดังภาพประกอบ 8



ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก



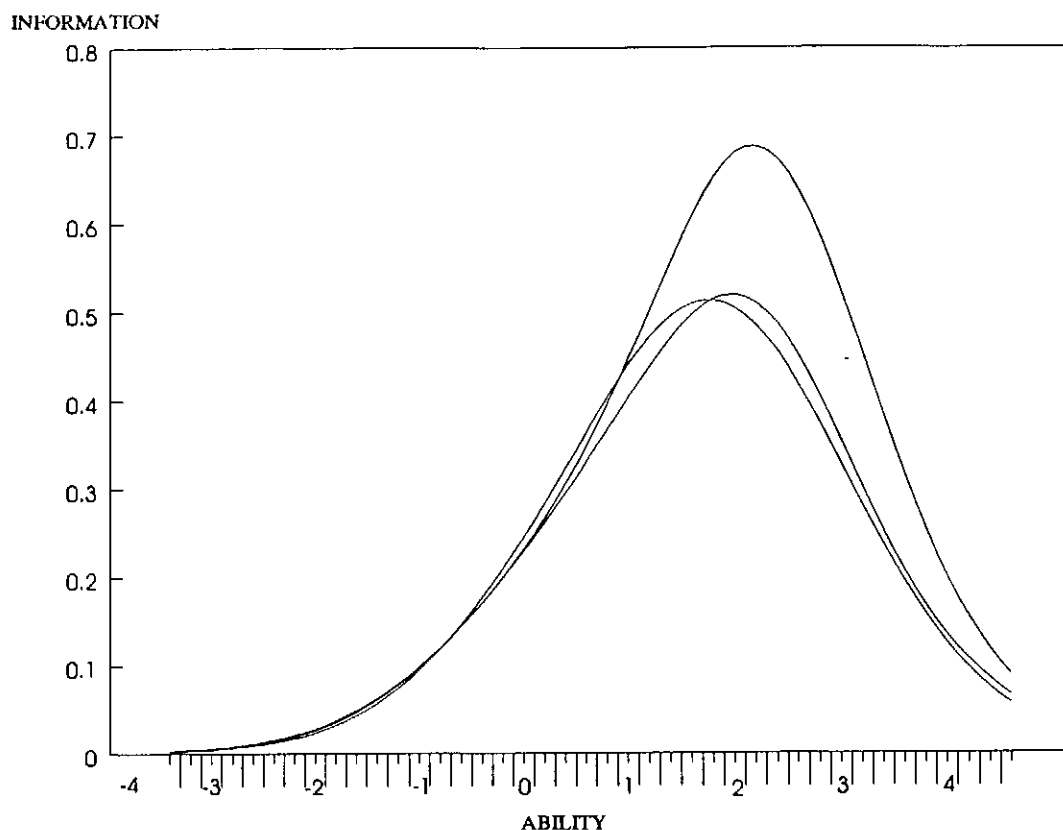
ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก



ฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก

ภาพประกอบ 7 โค้งลักษณะแบบทดสอบและโค้งอินฟอร์เมชันของแบบทดสอบทั้ง 3 ฉบับ

จากภาพประกอบ 7 แสดงถึงการเปรียบเทียบของโค้งลักษณะของแบบทดสอบ และโค้งอินฟอร์เมชันของแบบทดสอบทั้ง 3 ฉบับ จากโค้งลักษณะของแบบทดสอบทั้ง 3 ฉบับ แสดงว่า คนที่มีความสามารถต่ำ มีโอกาสที่จะตอบแบบทดสอบได้ถูก มีค่าน้อย คนที่มีความสามารถสูงมีโอกาสที่จะตอบแบบทดสอบได้ถูก มีค่ามาก คนที่มีความสามารถปานกลาง มีโอกาสที่จะตอบแบบทดสอบได้ถูกประมาณ 0.5 และจากโค้งอินฟอร์เมชัน แสดงให้เห็นว่าแบบทดสอบทั้ง 3 ฉบับ เหมาะสำหรับใช้ทดสอบกับนักเรียนที่มีระดับความสามารถสูง ๆ



ภาพประกอบ 8 การเปรียบเทียบโค้งอินฟอร์เมชันของแบบทดสอบทั้ง 3 ฉบับ

จากภาพประกอบ 8 แสดงให้เห็นว่า แบบทดสอบแบบเลือกตอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกันมีค่าฟังก์ชันอินฟอร์เมชันต่างกัน โดยแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีฟังก์ชันอินฟอร์เมชันของแบบทดสอบสูงสุด รองลงมาเป็นฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือกต่ำสุดซึ่งหมายถึงแบบทดสอบที่มีจำนวนข้อมากจำนวนตัวเลือกน้อยกว่า จะมีค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบสูงกว่าแบบทดสอบที่มีจำนวนข้อน้อยจำนวนตัวเลือกมาก หรือ แบบทดสอบที่มีจำนวนข้อมากตัวเลือกน้อยกว่า สามารถวัดนักเรียนได้คะแนนถูกต้องกว่าแบบทดสอบที่มีจำนวนข้อน้อยจำนวนตัวเลือกมากกว่า

4. ผลการหาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ

ผู้วิจัยได้คำนวณหาประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบระหว่างแบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก และแบบทดสอบฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก เทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก โดยการนำค่าฟังก์ชันอินฟอร์เมชัน

ของแบบทดสอบแต่ละฉบับ มาเปรียบเทียบกัน ปรากฏผลดังตาราง 3 มีรายละเอียดดังตาราง 14 ในภาคผนวก ก. และตาราง 15 ในภาคผนวก ง.

ตาราง 3 ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบระหว่าง แบบทดสอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกันของแบบทดสอบ 3 ฉบับ ที่ระดับความสามารถตั้งแต่ -4.00 ถึง 4.00 จำนวน 81 ระดับความสามารถ

ระดับ ความสามารถ	ค่าประสิทธิภาพสัมพัทธ์		
	ฉบับที่ 1 / ฉบับที่ 2	ฉบับที่ 1 / ฉบับที่ 3	ฉบับที่ 2 / ฉบับที่ 3
$\theta < -3.0$	0.9288 - 0.9356	1.1637 - 1.1450	1.2528 - 1.2237
$-3.00 \leq \theta < -1.00$	0.9373 - 1.0003	1.1409 - 0.9744	1.2173 - .9741
$-1.00 \leq \theta \leq 1.00$	1.0019 - 1.2272	0.9662 - 1.1985	.9644 - .9767
$1.00 < \theta \leq 3.00$	1.2483 - 1.5277	1.2364 - 1.6177	.9905 - 1.0589
$\theta > 3.00$	1.5236 - 1.3611	1.6127 - 1.5837	1.0585 - 1.1635

จากตาราง 3 แสดงให้เห็นช่วงการกระจายของค่าประสิทธิภาพสัมพัทธ์ ของการประมาณค่าความสามารถ ระหว่างแบบทดสอบที่แตกต่างกัน 3 ฉบับ ซึ่งนำมาเปรียบเทียบกับเกณฑ์ที่ 1 (Warm. 1978 : 76) ได้ผลดังนี้

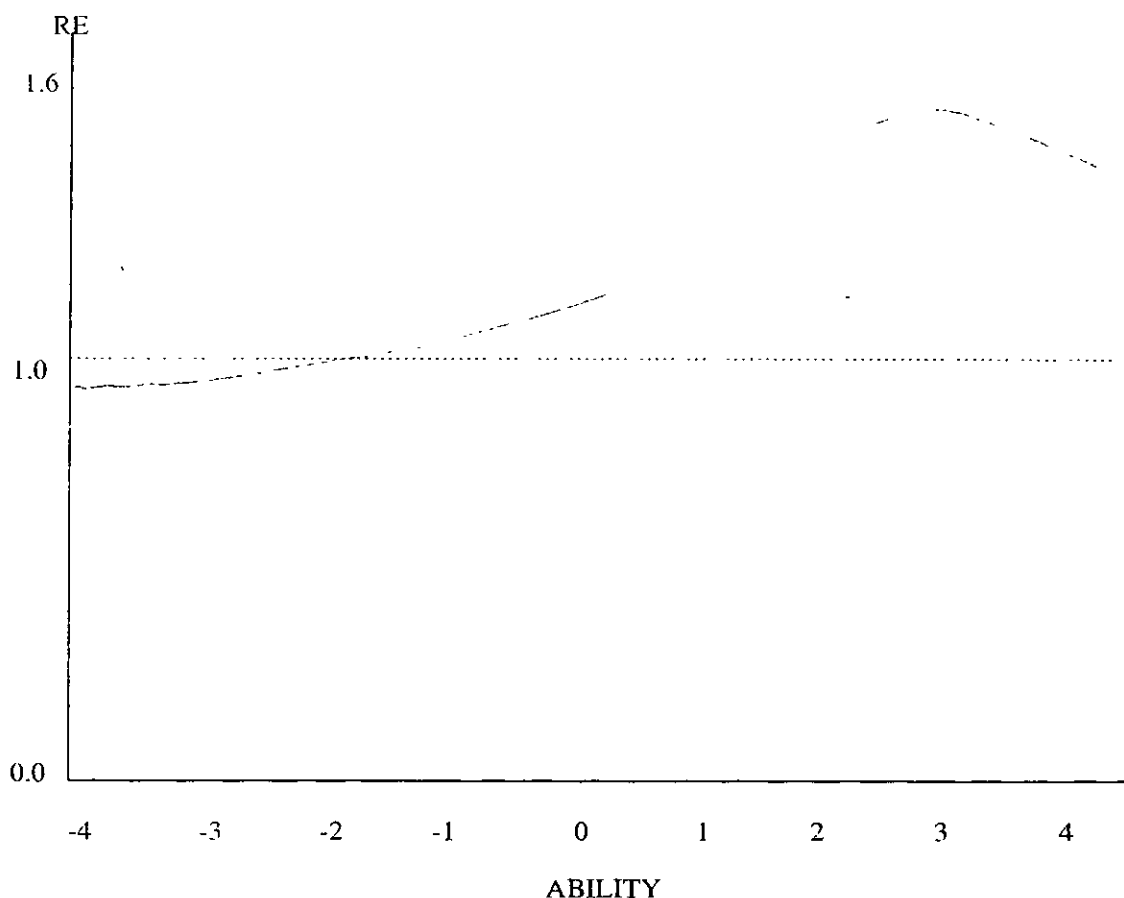
1. ค่าประสิทธิภาพสัมพัทธ์ ของการประมาณค่าความสามารถ ในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 2 ที่ 44 ข้อ 4 ตัวเลือก มีค่าต่ำกว่า 1 ที่ระดับความสามารถต่ำมาก ($\theta < -3.00$) แสดงว่าฉบับที่ 2 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่า ฉบับที่ 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก ส่วนที่ความสามารถระดับต่ำ มีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 และสูงกว่า 1 โดยมีค่าสูงกว่า 1 ที่ระดับความสามารถเท่ากับ -1.1 เพียงระดับเดียวเท่านั้น นอกนั้นที่ความสามารถตั้งแต่ -3.0 ถึง -1.2 ($-3.0 \leq \theta \leq -1.2$) จะมีค่าน้อยกว่า 1 แสดงว่าที่ความสามารถระดับต่ำนี้ส่วนมากมีค่าน้อยกว่า 1 แสดงว่า ฉบับที่ 2 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 1 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำ สำหรับที่ความสามารถระดับปานกลาง ($-1 \leq \theta \leq 1$) ระดับสูง ($1 < \theta \leq 3$) ระดับสูงมาก ($\theta > 3$) มีค่าประสิทธิภาพสัมพัทธ์สูงกว่า 1 ทุกระดับความสามารถ แสดงว่าฉบับที่ 1 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 2 ในกลุ่มผู้สอบที่มีความสามารถตั้งแต่ระดับปานกลางถึงระดับสูงมาก ดังนั้น จึงสรุปได้ว่าแบบทดสอบฉบับที่ 1 มี

ประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 2 ในกลุ่มผู้สอบที่มีความสามารถตั้งแต่ระดับปานกลางถึงระดับสูงมาก และฉบับที่ 2 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่า ฉบับที่ 1 ในกลุ่มผู้สอบที่มีความสามารถตั้งแต่ระดับต่ำลงมา

2. ค่าประสิทธิภาพสัมพัทธ์ ของการประมาณค่าความสามารถ ในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 2 มี 35 ข้อ 5 ตัวเลือก มีค่าสูงกว่า 1 ที่ระดับความสามารถต่ำมาก ($\theta < -3$) ระดับสูง ($1 < \theta \leq 3$) และระดับสูงมาก ($\theta > 3$) แสดงว่า ฉบับที่ 1 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่า ฉบับที่ 3 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมาก และสูงขึ้นไป ส่วนที่ระดับความสามารถต่ำ ($-3 \leq \theta < -1$) มีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 และสูงกว่า 1 โดยมีค่าต่ำกว่า 1 ที่ระดับความสามารถเท่ากับ -1.3, -1.2, -1.1 เพียง 3 ระดับเท่านั้น นอกนั้นที่ความสามารถตั้งแต่ -3 ถึง -1.4 ($-3 \leq \theta \leq -1.4$) จะมีค่าสูงกว่า 1 แสดงว่าที่ความสามารถระดับนี้ส่วนมากมีค่าสูงกว่า 1 แสดงว่าฉบับที่ 1 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่า ฉบับที่ 3 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำ สำหรับที่ระดับความสามารถปานกลางมีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 และสูงกว่า 1 โดยมีค่าสูงกว่า 1 ที่ระดับความสามารถเท่ากับ 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0 เพียง 7 ระดับเท่านั้น นอกนั้นที่ความสามารถตั้งแต่ -1 ถึง .3 ($-1 \leq \theta \leq .3$) จะมีค่าต่ำกว่า 1 แสดงว่าที่ความสามารถระดับปานกลางนี้ส่วนมากมีค่าต่ำกว่า 1 แสดงว่า ฉบับที่ 3 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่า ฉบับที่ 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ดังนั้น จึงสรุปได้ว่า แบบทดสอบฉบับที่ 1 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่า ฉบับที่ 3 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป และฉบับที่ 3 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง

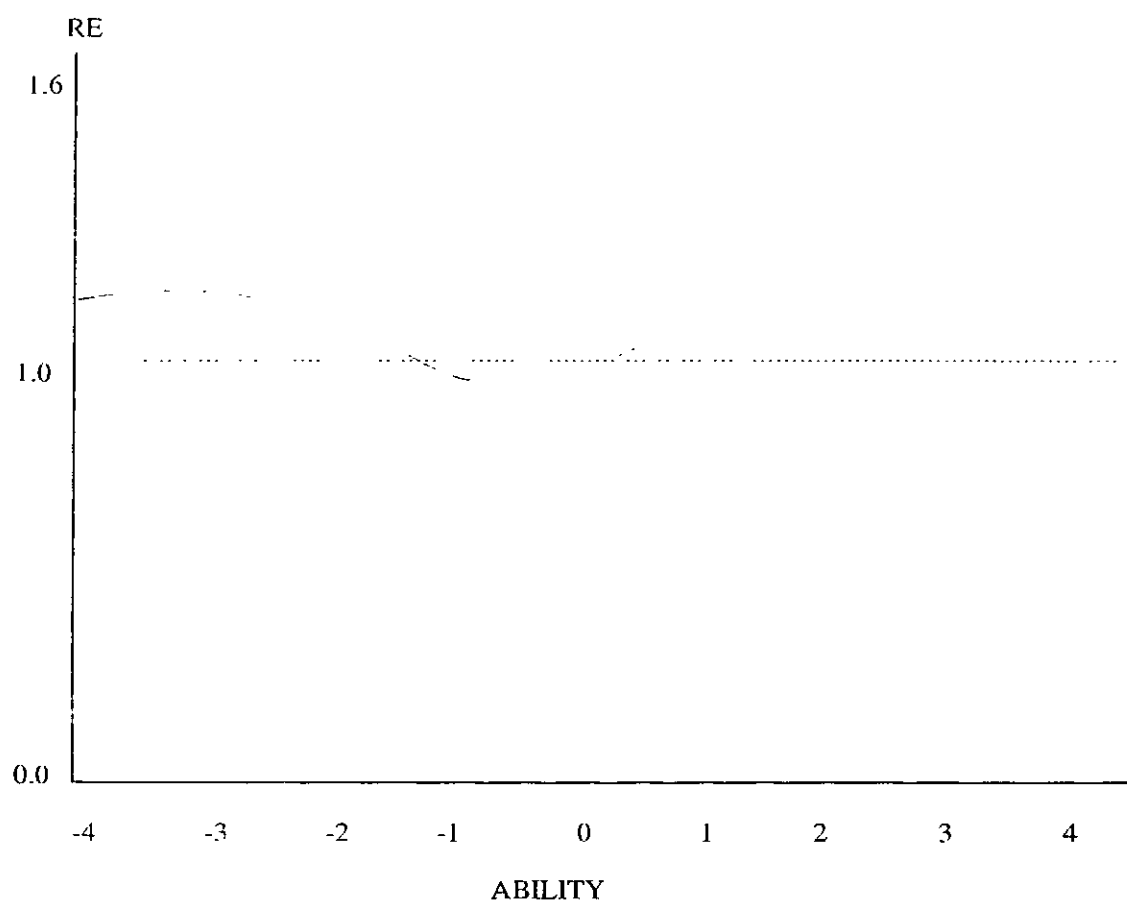
3. ค่าประสิทธิภาพสัมพัทธ์ ของการประมาณค่าความสามารถ ในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบ ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก เทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าต่ำกว่า 1 ที่ระดับความสามารถปานกลาง ($-1 \leq \theta \leq 1$) แสดงว่าฉบับที่ 3 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 2 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง และมีค่าสูงกว่า 1 ที่ระดับความสามารถต่ำมาก ($\theta < -3$) และระดับสูงมาก ($\theta > 3$) แสดงว่าฉบับที่ 2 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 3 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำมากและสูงมาก ส่วนที่ระดับความสามารถต่ำ ($-3 \leq \theta < -1$) มีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 และสูงกว่า 1 โดยมีค่าต่ำกว่า 1 มีระดับความสามารถเท่ากับ -1.3, -1.2, -1.1 เท่ากับ 0.9964, 0.9848, 0.9741 เพียง 3 ระดับเท่านั้น นอกนั้นที่ความสามารถตั้งแต่ -3 ถึง -1.4 ($-3 \leq \theta \leq -1.4$) จะมีค่าสูงกว่า 1 แสดงว่าที่ความสามารถระดับนี้ส่วนมากมีค่าสูงกว่า 1 แสดงว่า ฉบับที่ 2 มีประ

สัทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 3 ในกลุ่มผู้สอบที่มีความสามารถระดับต่ำ ส่วนที่ระดับความสามารถสูง ($1 < \theta \leq 3$) มีค่าประสิทธิภาพสัมพัทธ์ต่ำกว่า 1 และสูงกว่า 1 โดยมีค่าต่ำกว่า 1 ที่ระดับความสามารถเท่ากับ 1.1 เพียง 1 ระดับเท่านั้น นอกนั้นที่ความสามารถตั้งแต่ 1.2 ถึง 3 ($1.2 \leq \theta \leq 3$) จะมีค่าสูงกว่า 1 แสดงว่าที่ความสามารถระดับสูงนี้ ส่วนมากมีค่าสูงกว่า 1 แสดงว่า ฉบับที่ 2 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่า ฉบับที่ 3 ในกลุ่มผู้สอบที่มีความสามารถระดับสูง ดังนั้น จึงสรุปได้ว่า แบบทดสอบฉบับที่ 2 มีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 3 ในกลุ่มผู้สอบที่มีความสามารถตั้งแต่ระดับต่ำลงมา และระดับสูงขึ้นไป และฉบับที่ 3 มีประสิทธิภาพสัมพันธ์ในการประมาณค่าความสามารถสูงกว่า ฉบับที่ 2 ในกลุ่มผู้สอบ ที่มีความสามารถระดับปานกลาง และได้นำค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบที่แตกต่างกัน 3 แบบ มาเสนอในรูปแบบกราฟ ดังภาพประกอบ 9, 10, 11



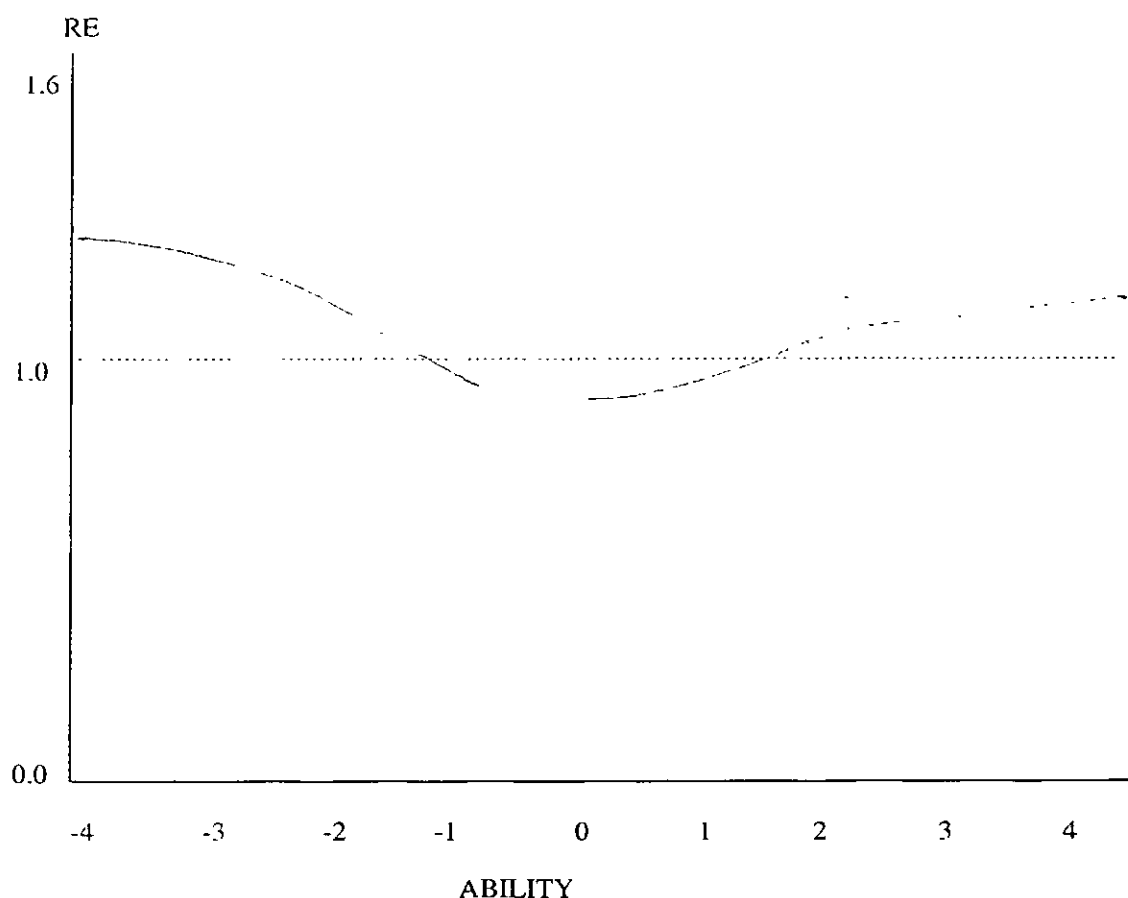
ภาพประกอบ 9 โค้งประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบระหว่าง
แบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก

จากภาพประกอบ 9 แสดงให้เห็นว่า ค่าประสิทธิภาพสัมพัทธ์ ของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีค่าสูงกว่า 1 ในช่วงระดับความสามารถตั้งแต่ -1.1 ถึง 4.00 แต่มีค่าต่ำกว่า 1 ในช่วงระดับความสามารถตั้งแต่ -4.00 ถึง -1.2 นั่นคือแบบทดสอบฉบับที่ 1 จะมีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 2 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลางถึงระดับสูงมาก ($\theta > -1.2$) แต่จะมีประสิทธิภาพในการประมาณค่าความสามารถต่ำกว่าฉบับที่ 2 ในกลุ่มผู้สอบที่มีความสามารถตั้งแต่ระดับต่ำลงมา ($\theta < -1.1$)



ภาพประกอบ 10 โค้งประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบระหว่าง
แบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก

จากภาพประกอบ 10 แสดงให้เห็นว่าค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือกเทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าสูงกว่า 1 ในช่วงระดับความสามารถตั้งแต่ -4.00 ถึง -1.4 และ 0.4 ถึง 4.0 แต่มีค่าต่ำกว่า 1 ในช่วง ระดับความสามารถตั้งแต่ -1.3 ถึง 0.3 นั่นคือ แบบทดสอบฉบับที่ 1 จะมีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 3 ในกลุ่มผู้สอบที่มีความสามารถตั้งแต่ระดับต่ำลงมา ($\theta < -1.3$) และระดับสูงขึ้นไป ($\theta > 0.3$) แต่จะมีประสิทธิภาพในการประมาณค่าความสามารถต่ำกว่าฉบับที่ 3 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ($-1.3 \leq \theta \leq 0.3$)



ภาพประกอบ 11 โค้งประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบระหว่าง
แบบทดสอบฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก เทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก

จากภาพประกอบ 11 แสดงให้เห็นว่าค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ ระหว่างแบบทดสอบฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือกเทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าสูงกว่า 1 ในช่วงความสามารถตั้งแต่ -4.0 ถึง -1.4 และ 1.2 ถึง 4.0 แต่มีค่าต่ำกว่า 1 ในช่วงระดับความสามารถตั้งแต่ -1.3 ถึง 1.1 นั่นคือแบบทดสอบฉบับที่ 2 จะมีประสิทธิภาพในการประมาณค่าความสามารถสูงกว่าฉบับที่ 3 ในกลุ่มผู้สอบที่มีความสามารถตั้งแต่ระดับต่ำลงมา ($\theta < -1.3$) และระดับสูงขึ้นไป ($\theta > 1.1$) แต่จะมีประสิทธิภาพในการประมาณค่าความสามารถต่ำกว่าแบบทดสอบจำนวน 35 ข้อ 5 ตัวเลือก ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ($-1.2 \leq \theta \leq 1.1$)

4. ผลการทดสอบความแตกต่างของค่าความยากของแบบทดสอบทั้ง 3 ฉบับ จากค่าพารามิเตอร์ความยากของข้อสอบของแบบทดสอบทั้ง 3 ฉบับ นำมาทดสอบความแตกต่างโดยการวิเคราะห์ความแปรปรวนแบบทิศทางเดียว (ONE WAY ANOVA) ได้ผลดังตาราง 4

ตาราง 4 การทดสอบความแตกต่างของค่าความยากของแบบทดสอบทั้ง 3 ฉบับ

แหล่งตัวแปร	SS	df	MS	F
ระหว่างกลุ่ม	0.3352	2	0.1676	0.1306
ภายในกลุ่ม	173.2931	135	1.2837	
รวม	173.6284	137		

จากตาราง 4 แสดงผลการทดสอบค่าความยากของแบบทดสอบทั้ง 3 ฉบับ ได้ค่า F เท่ากับ 0.1306 ซึ่งมีความแตกต่างอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 นั่นคือ แบบทดสอบทั้ง 3 ฉบับมีค่าความยากพอ ๆ กัน

5. ผลการทดสอบความแตกต่างของค่าอำนาจจำแนกของแบบทดสอบทั้ง 3 ฉบับ

จากค่าพารามิเตอร์อำนาจจำแนกของข้อสอบทั้ง 3 ฉบับ นำมาทดสอบความแตกต่างโดยการวิเคราะห์ความแปรปรวนแบบทิศทางเดียว (ONE WAY ANOVA) ได้ดังตาราง 5

ตาราง 5 การทดสอบความแตกต่างของค่าอำนาจจำแนกของข้อสอบทั้ง 3 ฉบับ

แหล่งตัวแปร	SS	df	MS	F
ระหว่างกลุ่ม	0.2713	2	0.1356	1.0500
ภายในกลุ่ม	17.4405	135	0.1292	
รวม	17.7118	137		

จากตาราง 5 ผลการทดสอบค่าอำนาจจำแนกของแบบทดสอบทั้ง 3 ฉบับได้ค่า F เท่ากับ 1.0500 ซึ่งมีความแตกต่างอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 นั่นคือแบบทดสอบทั้ง 3 ฉบับ มี

ค่าอำนาจจำแนกพอ ๆ กัน

6. ผลการทดสอบความแตกต่างค่าสัมประสิทธิ์การเคาของแบบทดสอบทั้ง 3 ฉบับ

จากค่าพารามิเตอร์สัมประสิทธิ์การเคาของแบบทดสอบทั้ง 3 ฉบับ นำมาทดสอบความแตกต่าง โดยการวิเคราะห์ความแปรปรวนแบบทิศทางเดียว (ONE WAY ANOVA) ได้ผลดังตาราง 6

ตาราง 6 การทดสอบความแตกต่างของค่าสัมประสิทธิ์การเคาของแบบทดสอบทั้ง 3 ฉบับ

แหล่งตัวแปร	SS	df	MS	F
ระหว่างกลุ่ม	0.4598	2	0.2299	60.2453*
ภายในกลุ่ม	0.5151	135	0.0038	
รวม	0.9749	137		

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 6 ผลการทดสอบค่าสัมประสิทธิ์การเคาของแบบทดสอบทั้ง 3 ฉบับ ได้ค่า F เท่ากับ 60.2453 ซึ่งมีความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

นั่นคือ แบบทดสอบ ทั้ง 3 ฉบับ จะต้องมีความแตกต่างอย่างน้อยหนึ่งคู่ ดังนั้นผู้วิจัยจึงทดสอบความแตกต่างรายคู่ด้วยการเปรียบเทียบพหุคูณ โดยวิธีเชฟเฟ ได้ผลดังตาราง 7

ตาราง 7 การทดสอบความแตกต่างรายคู่ด้วยการเปรียบเทียบพหุคูณ

แบบทดสอบ	$\bar{X}$	59ข้อ	44ข้อ	35ข้อ
		0.3185	0.2397	0.1770
59ข้อ	0.3185		0.0788*	0.1415*
44ข้อ	0.2397			0.1627*
35ข้อ	0.1770			

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 7 แสดงให้เห็นว่าค่าสัมประสิทธิ์การเดาของแบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก สูงกว่าฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือกสูงกว่าฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก เท่ากับ 0.0788, 0.1415 ตามลำดับ และฉบับที่ 2 มีค่าสัมประสิทธิ์การเดาสูงกว่าฉบับที่ 3 เท่ากับ 0.1627

ผู้วิจัยได้นำข้อสอบที่ซ้ำกันทุกข้อทั้ง 3 ฉบับ มาวิเคราะห์เพื่อหาค่าสถิติพื้นฐานและทดสอบความแตกต่างของค่าพารามิเตอร์พบว่าค่าความยาก ค่าอำนาจจำแนกและค่าสัมประสิทธิ์การเดาของแบบทดสอบทั้ง 3 ฉบับ ไม่ต่างจากผลของการวิเคราะห์ข้างต้น เพียงแต่ ค่าความยากมีแนวโน้มสูงขึ้นเมื่อแบบทดสอบมีจำนวนตัวเลือกเพิ่มขึ้น สำหรับค่าอำนาจจำแนกทุกฉบับมีค่าใกล้เคียง 1 และค่าสัมประสิทธิ์การเดาทุกฉบับมีค่าเฉลี่ยอยู่ที่ $(1.0 / \text{จำนวนตัวเลือก}) - 0.0.5$

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อศึกษาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบแบบเลือกตอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกัน
2. เพื่อศึกษาค่าความยากของข้อสอบแบบเลือกตอบที่มีจำนวนข้อ และจำนวนตัวเลือกต่างกัน
3. เพื่อศึกษาค่าอำนาจจำแนกของข้อสอบแบบเลือกตอบที่มีจำนวนข้อ และจำนวนตัวเลือกต่างกัน
4. เพื่อศึกษาค่าสัมประสิทธิ์การเดาของข้อสอบแบบเลือกตอบที่มีจำนวนข้อ และจำนวนตัวเลือกต่างกัน
5. เพื่อศึกษาฟังก์ชันอินฟอร์เมชันของแบบทดสอบแบบเลือกตอบที่มีจำนวนข้อ และจำนวนตัวเลือกต่างกัน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 1 ปีการศึกษา 2537 ของโรงเรียนสังกัดสำนักงานคณะกรรมการ การประถมศึกษาแห่งชาติ จังหวัดสมุทรสาคร จำนวน 3,016 คน ซึ่งเลือกมาโดยการสุ่มตัวอย่างแบบหลายขั้นตอน (Multistage Random Sampling)

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบแบบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่อง ตัวประกอบของจำนวนนับ และเศษส่วน ที่ผู้วิจัยสร้างขึ้นเองตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ.2533) จำนวน 3 ฉบับ ที่มีจำนวนข้อและจำนวนตัวเลือกต่างกัน ดังนี้

1. ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก
2. ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก
3. ฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก

วิธีดำเนินการเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลเพื่อการวิจัยครั้งนี้ ผู้วิจัยดำเนินการดังนี้

1. วางแผนในการดำเนินการสอบ โดยติดต่อขอความร่วมมือ ในการเก็บข้อมูลจากผู้ อำนวยการการประถมศึกษาจังหวัด และหัวหน้าการประถมศึกษาอำเภอ ในจังหวัดสมุทรสาคร และติดต่อโรงเรียนที่เป็นกลุ่มตัวอย่าง เพื่อกำหนดวัน เวลาในการดำเนินการสอบ
2. จัดเตรียมเครื่องมือที่ใช้ในการศึกษาค้นคว้า ในที่นี้คือแบบทดสอบแบบเลือกตอบวิชา คณิตศาสตร์ จำนวน 3 ฉบับ คือ
 - 2.1 ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก
 - 2.2 ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก
 - 2.3 ฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก
3. ผู้วิจัยนำแบบทดสอบทั้ง 3 ฉบับ ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า ตามที่แบ่งไว้ 3 กลุ่ม ๆ ละเท่า ๆ กัน โดยให้นักเรียนแต่ละคนในแต่ละกลุ่ม ทำแบบทดสอบ คนละ 1 ฉบับ
4. อธิบายให้นักเรียนในกลุ่มตัวอย่าง เข้าใจวัตถุประสงค์ในการสอบและขอความร่วมมือ ในการสอบ เพื่อให้ได้ผลตรงตามความเป็นจริง
5. นำผลการสอบมาตรวจให้คะแนน แล้ววิเคราะห์แบบทดสอบเพื่อตรวจสอบคุณ ลักษณะการวัดเพียงมิติเดียว และความเป็นอิสระในการตอบข้อสอบ โดยการวิเคราะห์องค์ ประกอบ (Factor Analysis) ซึ่งใช้โปรแกรม SPSS^x จากนั้นจึงวิเคราะห์แบบทดสอบด้วย โปรแกรมไบล็อก (Bilog) เพื่อหาค่าพารามิเตอร์ของข้อสอบ ค่าฟังก์ชันอินฟอเมชันของแบบ ทดสอบ ค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ และทดสอบสมมติฐานการวิจัยตามลำดับ

สรุปผลการวิเคราะห์ข้อมูล

1. ค่าพารามิเตอร์ของข้อสอบ

จากการวิเคราะห์ค่าพารามิเตอร์ของแบบทดสอบทั้ง 3 ฉบับ ด้วยโปรแกรมสำเร็จรูป

ไบล็อก (Bilog) เมื่อทดสอบค่า correlation biserial ของแบบทดสอบทั้ง 3 ฉบับ พบว่าแต่ละฉบับข้อสอบมีค่าคิดลบอย่างละ 1 ข้อ ซึ่งทำให้ไม่สามารถวิเคราะห์หาค่าพารามิเตอร์ได้ ดังนั้นจึงต้องตัดข้อเหล่านี้ออก จึงหาค่าพารามิเตอร์ได้ ทำให้แบบทดสอบ ฉบับที่ 1 เหลือข้อสอบ 59 ข้อ ฉบับที่ 2 เหลือข้อสอบ 44 ข้อ และฉบับที่ 3 เหลือข้อสอบ 35 ข้อ ได้ผลดังนี้

1.1 ค่าความยากของข้อสอบ

ค่าความยากของข้อสอบ แบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีค่าตั้งแต่ -.005 ถึง 3.488 มีค่าความยากเฉลี่ยของข้อสอบเท่ากับ 1.041 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.111 ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีค่าตั้งแต่ -.013 ถึง 3.567 มีความยากเฉลี่ยของข้อสอบเท่ากับ .951 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.224 และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าตั้งแต่ -.044 ถึง 3.043 มีค่าความยากเฉลี่ยของข้อสอบเท่ากับ .932 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.047 แสดงว่าแบบทดสอบฉบับที่ 1 มีความยากในการตอบถูกมากกว่า ฉบับที่ 2 และฉบับที่ 3 และฉบับที่ 2 มีความยากในการตอบถูกมากกว่า ฉบับที่ 3

1.2 ค่าอำนาจจำแนกของข้อสอบ

ค่าอำนาจจำแนกของข้อสอบ แบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีค่าตั้งแต่ .444 ถึง 2.528 มีค่าอำนาจจำแนกเฉลี่ยของข้อสอบเท่ากับ 1.077 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ .423 ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีค่าตั้งแต่ .369 ถึง 1.856 มีค่าอำนาจจำแนกเฉลี่ยของข้อสอบเท่ากับ .982 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ .330 และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าตั้งแต่ .470 ถึง 1.659 มีค่าอำนาจจำแนกเฉลี่ยของข้อสอบเท่ากับ .997 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ .264 แสดงว่าแบบทดสอบฉบับที่ 1 มีค่าอำนาจจำแนกของข้อสอบสูงกว่า ฉบับที่ 2 และฉบับที่ 3 และฉบับที่ 3 มีค่าอำนาจจำแนกของข้อสอบสูงกว่าฉบับที่ 2

1.3 ค่าสัมประสิทธิ์การเดาของข้อสอบ

ค่าสัมประสิทธิ์การเดาของข้อสอบ แบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีค่าตั้งแต่ .155 ถึง .484 มีค่าสัมประสิทธิ์การเดาเฉลี่ยของข้อสอบเท่ากับ .318 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ .063 ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีค่าตั้งแต่ .114 ถึง .342 มีค่าสัมประสิทธิ์การเดาเฉลี่ยของข้อสอบเท่ากับ .239 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ .065 และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าตั้งแต่ .072 ถึง .334 มีค่าสัมประสิทธิ์การเดาเฉลี่ยของข้อสอบเท่ากับ .177 ส่วนเบี่ยงเบน

มาตรฐานเท่ากับ .056 แสดงว่า ฉบับที่ 3 มีโอกาสในการเดาคำตอบได้ถูกต้องน้อยกว่าฉบับที่ 2 และฉบับที่ 1 และฉบับที่ 2 มีโอกาสในการเดาคำตอบได้ถูกต้องน้อยกว่าฉบับที่ 1

2. การศึกษาฟังก์ชันอินฟอร์เมชันของแบบทดสอบ ที่มีจำนวนข้อ และจำนวนตัวเลือก ต่างกัน

จากการศึกษาฟังก์ชันอินฟอร์เมชันของแบบทดสอบที่มีจำนวนข้อและจำนวนตัวเลือก ต่างกัน พบว่าแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีฟังก์ชันอินฟอร์เมชันของแบบทดสอบ สูงสุด เท่ากับ 0.6871 ที่ระดับความสามารถ (θ) 1.60 ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มี ฟังก์ชันอินฟอร์เมชันของแบบทดสอบรองลงมาเท่ากับ 0.5190 ที่ระดับความสามารถ (θ) 1.40 และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีฟังก์ชันอินฟอร์เมชันของแบบทดสอบต่ำสุด เท่ากับ 0.5120 ที่มีความสามารถ (θ) 1.20 แสดงว่าแบบทดสอบที่มีจำนวนข้อมากตัวเลือกน้อยกว่า จะมีฟังก์ชันอินฟอร์เมชันของแบบทดสอบสูงกว่า แบบทดสอบที่มีจำนวนข้อน้อยตัวเลือกมากกว่า ซึ่งฟังก์ชันอินฟอร์เมชันของแบบทดสอบ เป็นค่าความเชื่อมั่นในการวัดของแบบทดสอบ ณ ระดับความสามารถที่เหมาะสมที่สุดของแบบทดสอบฉบับนั้น โดยจุดที่โค้งที่สุดของฟังก์ชันอินฟอร์เมชันของแบบทดสอบอยู่ตรงกับระดับความสามารถใด แบบทดสอบฉบับนั้นก็เหมาะสำหรับใช้ทดสอบกับนักเรียนที่มีความสามารถในระดับนั้น

3.ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบ ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลางถึงระดับสูงมาก ($\theta > -1.2$) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถตั้งแต่ระดับต่ำลงมา ($\theta < -1.1$)

4.ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก เทียบกับ ฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าสูงกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถตั้งแต่ระดับต่ำลงมา ($\theta < 1$) และระดับสูงขึ้นไป ($\theta > 1$) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ($-1 \leq \theta \leq 1$)

5.ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก เทียบกับฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าสูงกว่า 1

ในกลุ่มผู้สอบที่มีความสามารถตั้งแต่ระดับต่ำลงมา ($\theta < -1$) และระดับสูงขึ้นไป ($\theta > 1$) แต่มีค่าต่ำกว่า 1 ในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง ($-1 \leq \theta \leq 1$)

6. การเปรียบเทียบค่าพารามิเตอร์ของข้อสอบทั้ง 3 ฉบับ

จากการศึกษาเปรียบเทียบค่าพารามิเตอร์ของข้อสอบที่คำนวณด้วยโปรแกรมไบล็อก (Bilog) ปรากฏว่าค่าความยากของข้อสอบ และค่าอำนาจจำแนกของข้อสอบทั้ง 3 แบบ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าสัมประสิทธิ์การเดาของข้อสอบ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยพบว่าแบบทดสอบฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีค่าสัมประสิทธิ์การเดาต่ำที่สุด และค่าสัมประสิทธิ์การเดาจะเพิ่มขึ้น เมื่อจำนวนข้อเพิ่มขึ้น และจำนวนตัวเลือกลดลงเป็น 44 ข้อ 4 ตัวเลือก และ 59 ข้อ 3 ตัวเลือกตามลำดับ นั่นคือแบบทดสอบที่มีจำนวนข้อน้อย ตัวเลือกมาก นักเรียนจะมีโอกาสในการเดาคำตอบได้ถูกต้องน้อยลง

อภิปรายผล

1. ค่าพารามิเตอร์ของข้อสอบ

1.1 ค่าความยากของข้อสอบ

ค่าความยากของข้อสอบแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือกฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก พบว่า ฉบับที่ 1 มีความยากในการตอบถูกมากกว่าฉบับที่ 2 และฉบับที่ 3 ฉบับที่ 2 มีความยากในการตอบถูกมากกว่า ฉบับที่ 3 ซึ่งแบบทดสอบที่มีจำนวนข้อมาก และจำนวนตัวเลือกน้อย จะต้องใช้ความคิด ในการค้นหาคำตอบของข้อสอบจำนวนมากข้อมากกว่าทำให้แบบทดสอบที่มีจำนวนข้อมาก จำนวนตัวเลือกน้อย มีความยากในการตอบถูกมากกว่าแบบทดสอบที่มีจำนวนข้อน้อย จำนวนตัวเลือกมาก

1.2 ค่าอำนาจจำแนกของข้อสอบ

ค่าอำนาจจำแนกของข้อสอบแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก พบว่าค่าอำนาจจำแนกของฉบับที่ 1 มีค่าอำนาจจำแนกของข้อสอบมากกว่าฉบับที่ 2 และฉบับที่ 3 และแบบทดสอบฉบับที่ 3 มีค่าอำนาจจำแนกมากกว่าฉบับที่ 2 ทั้งนี้เนื่องจากการศึกษาครั้งนี้ ใช้วิธีการตัดข้อและตัวลองที่มีค่าอำนาจจำแนกต่ำสุดออก ทำให้แบบทดสอบที่มีจำนวนข้อมาก ตัวเลือกน้อยกว่า มีค่าอำนาจจำแนกเฉลี่ยของข้อสอบรายข้อสูงกว่าแบบทดสอบที่มีจำนวนข้อน้อย จำนวนตัวเลือกมากกว่า และฉบับที่ 2

และฉบับที่ 3 ซึ่งมีจำนวนข้อใกล้เคียงกัน ก็มีโอกาสดำเนินการจำแนกเฉลี่ยของแบบทดสอบที่มีจำนวนข้อน้อย จำนวนตัวเลือกมากมีค่าสูงกว่าแบบทดสอบที่มีจำนวนข้อมากกว่า และตัวเลือกน้อยกว่าได้

1.3 ค่าสัมประสิทธิ์การเดาของข้อสอบ

ค่าสัมประสิทธิ์การเดาของข้อสอบแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก พบว่า ฉบับที่ 3 มีโอกาสในการเดาคำตอบได้ถูกต้อง น้อยกว่าฉบับที่ 2 และฉบับที่ 1 แบบทดสอบฉบับที่ 2 มีโอกาสในการเดาคำตอบได้ถูกต้องน้อยกว่าฉบับที่ 1 ทั้งนี้เนื่องจากจำนวนข้อ และจำนวนตัวเลือกต่างกัน ทำให้โอกาสในการเดาต่างกันถ้าแบบทดสอบมีจำนวนข้อน้อย แต่จำนวนตัวเลือกมาก โอกาสในการเดาถูกในแต่ละข้อก็จะลดลง

2.ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถในวิชาคณิตศาสตร์ ระหว่างแบบทดสอบทั้ง 3 ฉบับ พบว่าผลการวิจัยนี้ทั้งขัดแย้งและ สอดคล้องกับสมมติฐานที่ตั้งไว้ ซึ่งพิจารณาได้ดังนี้

ผลการวิจัยที่สอดคล้องกับสมมติฐาน มีดังนี้

2.1 แบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือกในกลุ่มผู้สอบที่มีความสามารถระดับปานกลางถึงระดับสูงมาก

2.2 แบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือกในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป

2.3 แบบทดสอบฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือกในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป

จากการวิจัยทั้ง 3 ข้อนี้พบว่า แบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือกมีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก และแบบทดสอบฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก เฉพาะในกลุ่มผู้สอบระดับความสามารถต่าง ๆ นั้นน่าจะเป็นเพราะว่า การหาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ

แสดงถึงการเปรียบเทียบคุณภาพของแบบทดสอบ ได้คำนวณมาจากค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ

โดยค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบนี้ มาจากผลรวมของค่าฟังก์ชันอินฟอร์เมชันของข้อสอบ ซึ่งในการวิจัยครั้งนี้ ค่าฟังก์ชันอินฟอร์เมชันของข้อสอบแต่ละระดับความสามารถคิดจากค่าเฉลี่ย ค่าฟังก์ชันอินฟอร์เมชัน เป็นฟังก์ชันของค่าพารามิเตอร์ของข้อสอบที่เป็นดัชนีกำหนดคุณภาพของข้อสอบแต่ละข้อในแบบทดสอบ ก็คือค่าอำนาจจำแนก ค่าความยากของข้อสอบ และค่าสัมประสิทธิ์การเดาของข้อสอบ ค่าพารามิเตอร์เหล่านี้มีผลโดยตรงต่อค่าฟังก์ชันอินฟอร์เมชันของข้อสอบเพิ่มขึ้น เมื่อนำมาหาค่าประสิทธิภาพสัมพัทธ์ของแบบทดสอบ ทำให้แบบทดสอบฉบับที่ 1 มีคุณภาพสูงกว่าฉบับที่ 2 และแบบทดสอบฉบับที่ 3 เนื่องจากแบบทดสอบฉบับที่ 1 มีค่าอำนาจจำแนกสูงกว่าฉบับที่ 2 และฉบับที่ 3 เฉพาะในกลุ่มผู้สอบระดับความสามารถปานกลางถึงสูงมาก ทั้งนี้อาจเนื่องจากแบบทดสอบที่มีจำนวนข้อมาก ตัวเลือกน้อยนั้น ผู้สอบที่มีความสามารถระดับปานกลางถึงระดับสูงมาก มีโอกาสที่จะตอบข้อสอบได้ถูก มีน้อยกว่าแบบทดสอบที่มีจำนวนข้อน้อย ตัวเลือกมาก (ตาราง 14 ในภาคผนวก จ.) และเนื่องจากแบบทดสอบที่มีจำนวนตัวเลือกน้อยนั้น ตัวเลือกแต่ละตัวจะมีค่าอำนาจจำแนกมาก เนื่องจากผู้วิจัยได้ตัดค่าตัวเลือกที่มีค่าอำนาจจำแนกต่ำสุดออกไปทีละตัว จึงทำให้แบบทดสอบ ที่มีจำนวนข้อมาก ตัวเลือกน้อย ตัวเลือกแต่ละตัวมีค่า อำนาจจำแนกมาก ส่วนแบบทดสอบที่มีจำนวนข้อน้อย มีตัวเลือกมากนั้น ค่าอำนาจจำแนกมีตั้งแต่ต่ำสุดถึงสูงสุด ด้วยเหตุผลดังกล่าว จึงทำให้แบบทดสอบที่มีจำนวนข้อมาก ตัวเลือกน้อยมีค่าฟังก์ชันอินฟอร์เมชันมากกว่าแบบทดสอบที่มีจำนวนข้อน้อย ตัวเลือกมาก

ผลการวิจัยที่ขัดแย้งกับสมมติฐาน มีดังนี้

2.1 แบบทดสอบฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือกในกลุ่มผู้สอบที่มีความสามารถระดับต่ำลงมา

2.2 แบบทดสอบฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงกว่าฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก และฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือกในกลุ่มผู้สอบที่มีความสามารถระดับปานกลาง

ผลจากการวิจัยที่ขัดแย้งกับสมมติฐานดังกล่าว เมื่อพิจารณาค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบทั้ง 3 แบบ พบว่า ค่าสูงสุดของค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบอยู่ที่ระดับความสามารถ 1.6, 1.4 และ 1.2 ตามลำดับ แสดงว่าแบบทดสอบทั้ง 3 ฉบับมีความเหมาะสมในการสอบวัดกับกลุ่มผู้สอบที่มีความสามารถระดับสูง ซึ่งแฮมเบิลตันและคูด (Hambleton and Cook 1977 : 86-87) ได้ กล่าวว่า ถ้าค่าฟังก์ชันอินฟอร์เมชันยิ่งมากค่าความคลาดเคลื่อนมาตรฐาน

ในการประมาณค่า ความสามารถก็จะยิ่งน้อยลง นั่นคือ ถ้าค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบมีค่าสูงสุด ณ ระดับความสามารถใด แสดงว่าแบบทดสอบนั้นมีประสิทธิภาพในการวัดสูงสุดที่ระดับความสามารถนั้น ก็จะทำให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า ณ ระดับความสามารถนั้นลดลง ทำให้มีความแม่นยำในการประมาณค่า ณ ระดับความสามารถนั้นสูง ดังนั้นเมื่อนำค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบทั้ง 3 ฉบับมาเปรียบเทียบทีละคู่ เพื่อหาคุณภาพของแบบทดสอบในกลุ่มผู้สอบความสามารถระดับต่ำลงมาและระดับสูงขึ้นไป ทำให้มีผลการวิจัยที่สอดคล้องกับสมมติฐานกล่าวคือแบบทดสอบฉบับที่ 1 และฉบับที่ 2 มีคุณภาพสูงกว่าฉบับที่ 3 แต่เมื่อนำมาเปรียบเทียบคุณภาพของแบบทดสอบในกลุ่มผู้สอบระดับความสามารถปานกลาง และกลุ่มผู้สอบระดับความสามารถต่ำนั้น อาจจะทำให้มีการแกว่ง (Fluctuation) ของคุณลักษณะภายในหรือความสามารถที่วัดนั้นไม่แจ่มแจ้งแบบปกติ จึงน่าจะให้มี ผลการวิจัยที่ขัดแย้งกับสมมติฐาน จึงไม่สามารถสรุปได้อย่างชัดเจน ในกลุ่มผู้สอบระดับความสามารถปานกลางและกลุ่มผู้สอบระดับความสามารถต่ำลงมา

3. การเปรียบเทียบค่าพารามิเตอร์ของข้อสอบ

3.1 ค่าความยากของข้อสอบ

ค่าความยากของข้อสอบของแบบทดสอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกันตามแนวทฤษฎีการตอบข้อคำถาม พบว่าค่าความยากของแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ผลการศึกษาค้นครั้งนี้ไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ แต่สอดคล้องกับการศึกษาของชาติชาย กู้กิตติไมตรี (2533 : 61) ที่พบว่าค่าความยากของแบบทดสอบแบบเลือกตอบที่มีจำนวนตัวเลือกต่างกัน โดยการใช้ทฤษฎีการตอบข้อคำถามนั้น แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.2 ค่าอำนาจจำแนกของข้อสอบ

ค่าอำนาจจำแนกของข้อสอบของแบบทดสอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกันตามแนวทฤษฎีการตอบข้อคำถาม พบว่า แบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ผลการศึกษาค้นครั้งนี้ ไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ แต่สอดคล้องกับการศึกษาของชาติชาย กู้กิตติไมตรี (2533 : 61) ที่พบว่าค่าอำนาจจำแนกของแบบทดสอบแบบเลือกตอบที่มี

จำนวนตัวเลือกต่างกัน โดยการใช้ทฤษฎีการตอบข้อคำถามนี้ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3.3 ค่าสัมประสิทธิ์การเดาของข้อสอบ

ค่าสัมประสิทธิ์การเดาของข้อสอบของแบบทดสอบที่มีจำนวนข้อและจำนวนตัวเลือกต่างกัน แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ผลการศึกษาค้นคว้าครั้งนี้สอดคล้องกับสมมติฐานที่ตั้งไว้ในข้อ 6 ซึ่งเป็นการสนับสนุนการศึกษานักเรียนชาติดิษาย กู้กิตติไมตรี (2533 : 61) ที่พบว่าค่าสัมประสิทธิ์การเดาของแบบทดสอบแบบเลือกตอบที่มีจำนวนตัวเลือกต่างกัน โดยการใช้ทฤษฎีการตอบข้อคำถามนั้น แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเป็นการสนับสนุนการศึกษามอนโร (Monroe, 1952 : 1507) ที่พบว่าแบบทดสอบที่มีจำนวนตัวเลือกมากขึ้นจะช่วยลดการเดาในการตอบแบบทดสอบได้ ทั้งนี้เนื่องจากข้อสอบที่มีจำนวนข้อมาก แต่จำนวนตัวเลือกน้อยกว่า ทำให้ผู้ตอบมีโอกาสในการเดาคำตอบได้ถูกต้องมากกว่าแบบทดสอบที่มีจำนวนข้อน้อย แต่จำนวนตัวเลือกมากกว่า ซึ่งยังสนับสนุนคำกล่าวของกรอนลันด์ (Gronlund, 1982 : 26) ที่ว่าแบบทดสอบ ยังมีตัวเลือกมากเท่าไร โอกาสในการเดาคำตอบที่ถูกต้อง ยังมีน้อยลง

ข้อเสนอแนะ

1. การสร้างแบบทดสอบตามกรอบของทฤษฎีการตอบข้อคำถาม เพื่อให้แบบทดสอบมีคุณภาพสูงสุดในกลุ่มผู้มีความสามารถระดับสูง ควรสร้างแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ส่วนผู้ที่มีความสามารถระดับปานกลาง และระดับต่ำ ควรสร้างแบบทดสอบ ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือกและฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก ตามลำดับ

2. ควรมีการศึกษาเรื่องจำนวนข้อและจำนวนตัวเลือกต่างกัน โดยใช้ทฤษฎีการตอบข้อคำถามอีก แต่ในการจัดรูปแบบทดสอบนั้น ให้พิจารณาตัดข้อและตัวเลือกที่มีค่าความยากสูงและต่ำออก เพื่อนำผลมาเปรียบเทียบกับการศึกษาครั้งนี้

3. ควรจะได้ทำการวิจัยทำนองเดียวกันนี้ โดยศึกษาในวิชาอื่น ๆ ได้แก่ สังคมศึกษา ภาษาไทย วิทยาศาสตร์ เป็นต้น

บรรณานุกรม

บรรณานุกรม

- จิราพร เนียมสุวรรณ. การเปรียบเทียบประสิทธิภาพสัมพัทธ์ และค่าความเที่ยงตรงของแบบเลือกตอบที่มีรูปแบบตัวเลือกแตกต่างกัน. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2532. อัดสำเนา.
- จักรกฤษณ์ สำราญใจ. ประสิทธิภาพเชิงสัมพัทธ์ของข้อสอบเลือกตอบชนิดตัดสินคำตอบทุกตัวเลือกเทียบกับข้อสอบเลือกตอบชนิดแบบฉบับในแบบสอบผลสัมฤทธิ์. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2530. อัดสำเนา.
- ชนะ ทาวงศ์. การศึกษาเปรียบเทียบคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ชนิดเลือกตอบที่มีจำนวนข้อและตัวเลือกแตกต่างกันวิชาวิทยาศาสตร์ชั้นประถมศึกษาปีที่ 4. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2521. อัดสำเนา.
- ชาติชาย กู้กิตติไมตรี. การเปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบที่มีจำนวนตัวเลือกแตกต่างกันโดยการใช้ทฤษฎีการตอบข้อคำถาม. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2533. อัดสำเนา.
- क्रम हाणुकरकुल. การเปรียบเทียบคุณภาพของแบบทดสอบชนิดเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 7 ที่มีจำนวนตัวเลือกไม่เท่ากัน. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัย ศรีนครินทรวิโรฒ ประสานมิตร, 2519. อัดสำเนา.
- ธีระศักดิ์ อูร์จนาพันธ์. การเปรียบเทียบคุณภาพของแบบทดสอบแบบเลือกตอบที่มี จำนวนตัวเลือกไม่เท่ากัน โดยการใช้ทฤษฎีการตอบสนองข้อสอบ. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2530. อัดสำเนา.
- บุญเรือน จรัสวิมล. การเพิ่มข้อสอบที่มีผลต่อค่าความเชื่อมั่นของแบบทดสอบ. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2521. อัดสำเนา.
- บุญเชิด ภิญโญนันตพงษ์. การสร้างแบบทดสอบ เอกสารคำสอนวิชาวัดผล 301. คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2525..
- ประคิษฐ์ เรืองตระกูล. การใช้โมเดลโลจิสติกเพื่อพัฒนาแบบสอบวินิจฉัยวิชาคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1. วิทยานิพนธ์ ปรียญมหาบัณฑิต กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2529. อัดสำเนา.

ผจงจิต อินทสุวรรณ. "Latent Trait Theory," วารสารวัดผลการศึกษา 3 (มกราคม - เมษายน 2525) : 51-67 ทฤษฎีการตอบข้อคำถาม. กรุงเทพมหานคร : สถาบันวิจัยพฤติกรรมศาสตร์ มหาวิทยาลัยศรีนครินทร วิโรฒ ประสานมิตร, 2528.

ล้วน สายยศ และ อังคณา สายยศ. หลักการวิจัยทางการศึกษา. ศึกษาพร, 2531.

วิภาดา วัฒนกุลกิตต์. การเปรียบเทียบค่าความยาก ค่าความเชื่อมั่นและสัดส่วน การเดาของแบบทดสอบเลือกตอบที่มีตัวเลือกต่างกัน. ปริญญาณีพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2529. อัดสำเนา.

วดาภรณ์ พูนผลอำนวย. ผลของรูปแบบตัวเลือกต่างกันที่มีต่อรูปแบบการตอบสนองข้อสอบ ไต่ถึงลักษณะของข้อสอบ ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ. วิทยานิพนธ์ ปริญญาดุชฎีบัณฑิต กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2533. อัดสำเนา.

สงบ ลักษณะ. "ความสามารถเกี่ยวกับการตอบข้อสอบ," วารสารวัดผลการศึกษา. 4(1) : 47-54, 88-93 : พฤษภาคม - สิงหาคม, 2525.

สำเริง บุญเรืองรัตน์. "การพัฒนาทฤษฎีเลเห็นเทรทเพื่อวิเคราะห์ข้อสอบ," วารสารวัดผลการศึกษา. 4(2) : 1-3 ; กันยายน - ธันวาคม, 2525.

อวยพร วิบูลย์กาญจน์. การเปรียบเทียบผลการวิเคราะห์แบบสอบอุปมาอุปไมยด้วยคลาสสิกอลโมเดลกับราชภัฏโมเดล. วิทยานิพนธ์ ก.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2526. อัดสำเนา.

อนันต์ ศรีโสภ. ทฤษฎีการวัดและการทดสอบ. กรุงเทพฯ : ไทยวัฒนาพานิช, 2525.

อุทุมพร จามรมาน. "การวัดความสามารถที่แท้จริง," บทคัดสรรทางวิชาการทดสอบ. โครงการพัฒนาแบบทดสอบ, กรุงเทพฯ : มหาวิทยาลัยสุโขทัยธรรมมาธิราช, 93-100 ; มิถุนายน 2529.

Birnbaum, A. "Some Latent Trait Model and Thier Use in Infering an Examine's ability," In , Statistical Theories of Mental Test Scores. pp. 395-479 ; by Lord, F.M., and Novick M.R. Readings, Massachusetts : Addison. Wesley, 1968.

Budescu, David. V and Barch. Nevo. "Optimal Number of Options : An Investigation of the Assumption of Proportionality," Journal of Educational Measurement. 22:183-186 ; Fall, 1985.

Costin, F. "The Optimal Number of Alternative In Multiple Choice Achievement Test : Some Empirical Evidence For A Mathematical Roff," Educational and Psychological Measurement. 30 : 353-358 ; 1970.

Chase, Clinton, I. Measurement for Educational Evaluation. Philippines, Addison - Wesley Publishing Co., Inc., 1974, 312 p.

Ebel, R.L. Measuring Educational Achievement. Englewood Cliffs ; Prentice-Hall, 1965, 481p.

Gronlund, Norman. E. Constructing Achievement Test. 3rd, ed., Engle - Wood Cliffs, Prentice - Hall, 1982.

Freurd, John E. Modern Elementary Statistics. 5th, ed., London Prentice - Hall, Inc., 1979.

Hambleton, R.K. and H.Swaminathan, Item Response Theory Principles and Applications. 2nd, ed., Kluwer - Nijhoff, Publishing. 1985.

Kendall, L.M. "The Effects Varying Time Limits On Test Validity," Educational and Psychological Measurement. 24 : 789-798 ; 1964.

Lord, F.M. Application of Item Response Theory To Practical Testing Problems. New Jersuy : Lawrence Erlbaum. Associates, Publishers, 1980.

Lord, F.M. and M.R. Novick. Statistical Theories of Mental Test Scores. Massachusetts : Addison - Wesley, Publishing Company Inc., 1968.

Lyman, Howard B. Test Scores and What They Mean. Prentice - Hall, Inc., Englewood Cliff, N.J. 1963.

Remmers, H.H. and R.H. Adkins. "Reliability of Multiple - Choice Measurement Instruments As A Function of The Spearman - Brown Formula IV," Journal of Educational Psychology. 33 : 385-390 ; 1952.

Tversky, A. "On The Optimal Number of Alternatives At a Choice Point," Journal of Mathematical Psychology. 1 : 386-391 ; 1964.

Williams, B.J. and R.L. Ebel. "The Effect of Varying The Number of Alternatives Per Item On The Multiple-Choice Vocabulary Test Items," The Fourteenth Year Book of The National Council On Measurements. 63-65 ; 1957.

Warm, Thomas A. A Primer of Item Response Theory. Oklahoma U.S. Coast Guard Institute, 1978.

Zimmerman, W.S. and L.G. Humphreys. "Item Reliability As A Function of The Omission of Mislead," American Psychologist. 8 : 460-461 ; 1953.

ภาคผนวก ก.

ผลการวิเคราะห์องค์ประกอบที่มีค่าไอเกนมากกว่า 1
เพื่อตรวจสอบความเป็นมิติเดียวของแบบทดสอบ จำนวน 3 ฉบับ

ตาราง 8 ค่าไอเกน (มากกว่า 1) เปอร์เซนต์ของความแปรปรวนและเปอร์เซนต์ของความแปรปรวนสะสม จากการนำแบบทดสอบ จำนวน 3 แบบ มาวิเคราะห์องค์ประกอบ

แบบทดสอบ	องค์ประกอบ ที่	ค่าไอเกน	เปอร์เซนต์ของ ความแปรปรวน	เปอร์เซนต์ของ ความแปรปรวน สะสม
59 ข้อ 3 ตัวเลือก	1	5.56197	9.3	9.3
	2	1.99220	3.3	12.6
	3	1.57029	2.6	15.2
	4	1.53587	2.6	17.8
	5	1.48315	2.5	20.2
	6	1.40692	2.3	22.6
	7	1.37507	2.3	24.9
	8	1.34094	2.2	27.1
	9	1.31680	2.2	29.3
	10	1.29957	2.2	31.5
	11	1.26840	2.1	33.6
	12	1.23877	2.1	35.6
	13	1.19685	2.0	37.6
	14	1.18517	2.0	39.6
	15	1.14877	1.9	41.5
	16	1.14130	1.9	43.4
	17	1.09093	1.8	45.3
	18	1.07361	1.8	47.0
	19	1.04479	1.7	48.8
	20	1.03741	1.7	50.5
	21	1.01729	1.7	52.2
	22	1.01289	1.7	53.9
	23	1.00204	1.7	55.6

ตาราง 8 (ต่อ)

แบบทดสอบ	องค์ประกอบ ที่	ค่าไอเกน	เปอร์เซ็นต์ของ ความแปรปรวน	เปอร์เซ็นต์ของความ แปรปรวนสะสม
44 ข้อ 4 ตัวเลือก	1	5.18127	11.5	11.5
	2	1.75166	3.9	15.4
	3	1.49250	3.3	18.7
	4	1.38471	3.1	21.8
	5	1.33265	3.0	24.8
	6	1.29601	2.9	27.6
	7	1.26796	2.8	30.5
	8	1.23830	2.8	33.2
	9	1.18124	2.6	35.8
	10	1.16606	2.6	38.4
	11	1.11198	2.5	40.9
	12	1.06462	2.4	43.3
	13	1.05911	2.4	45.6
	14	1.03644	2.3	47.9
	15	1.01858	2.3	50.2
35 ข้อ 5 ตัวเลือก	1	5.32792	14.8	14.8
	2	1.83862	5.1	19.9
	3	1.27440	3.5	23.4
	4	1.26402	3.5	27.0
	5	1.20976	3.4	30.3
	6	1.17764	3.3	33.6
	7	1.15162	3.2	36.8
	8	1.10210	3.1	39.9
	9	1.04135	2.9	42.7
	10	1.02037	2.8	45.6

ภาคผนวก ข.

ค่าพารามิเตอร์ของแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก
ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก

ตาราง 9 ค่าพารามิเตอร์ของแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก

ข้อ	พารามิเตอร์			ข้อ	พารามิเตอร์		
	a	b	c		a	b	c
1	.496	-.249	.368	26	.557	-.009	.310
2	.870	-.129	.221	27	.729	.008	.224
3	.658	.178	.238	28	.924	1.511	.398
4	1.020	-.970	.225	29	.636	-.008	.418
5	.918	-1.616	.279	30	1.282	1.127	.406
6	.983	-.005	.367	31	1.452	.732	.348
7	1.037	1.748	.328	32	1.185	1.026	.303
8	.947	1.598	.283	33	1.187	.630	.334
9	1.121	1.533	.422	34	1.316	.168	.307
10	.735	1.658	.484	35	1.032	1.631	.307
11	1.547	2.471	.310	36	1.277	1.307	.351
12	1.228	-.877	.200	37	1.131	1.655	.363
13	.568	-.828	.367	38	1.106	1.582	.280
14	.758	.465	.345	39	1.690	1.891	.257
15	.985	-.792	.286	40	2.000	1.493	.273
16	.675	.767	.311	41	.959	1.669	.395
17	1.098	-1.034	.279	42	1.803	1.084	.309
18	.481	-.152	.398	43	1.531	1.420	.294
19	1.128	1.533	.155	44	1.839	1.671	.317
20	1.741	1.547	.318	45	1.426	1.563	.266
21	.927	-.107	.276	46	.768	2.792	.336
22	.564	.795	.179	47	.660	1.512	.377
23	.936	.439	.294	48	1.483	2.402	.292
24	1.271	-.064	.359	49	1.378	2.158	.381
25	-	-	-	50	.929	1.121	.328

ตาราง 9 (ต่อ)

ข้อ	พารามิเตอร์			ข้อ	พารามิเตอร์		
	a	b	c		a	b	c
51	1.666	1.466	.304	56	1.290	2.599	.387
52	.595	.702	.279	57	.592	1.606	.345
53	.444	2.743	.427	58	.950	2.699	.322
54	2.528	2.290	.281	59	1.001	3.488	.315
55	.736	2.118	.340	60	.777	1.688	.323

ตาราง 10 ค่าพารามิเตอร์ของแบบทดสอบฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก

ข้อ	พารามิเตอร์			ข้อ	พารามิเตอร์		
	a	b	c		a	b	c
1	.666	-.145	.284	26	.636	.067	.267
2	.959	.103	.114	27	.639	.219	.207
3	.724	.110	.152	28	.426	3.066	.500
4	.947	-.726	.201	29	.623	-.234	.223
5	1.042	-1.588	.260	30	.889	1.016	.249
6	.725	-.342	.260	31	1.330	.798	.290
7	.954	1.703	.167	32	1.322	.959	.198
8	.736	1.863	.237	33	1.320	.801	.263
9	1.496	1.571	.316	34	1.161	1.147	.206
10	.950	3.416	.342	35	.833	2.521	.194
11	-	-	-	36	1.297	2.099	.303
12	1.088	-.818	.160	37	.938	2.081	.270
13	.494	-.403	.281	38	.955	2.237	.235
14	.937	.146	.220	39	1.630	2.167	.228
15	.783	-.892	.194	40	1.658	1.583	.187
16	1.019	.939	.273	41	.895	2.298	.234
17	.775	-1.033	.241	42	1.856	1.299	.292
18	.269	-.107	.276	43	.908	1.826	.261
19	1.162	1.483	.162	44	.904	1.878	.217
20	1.108	1.659	.267	45	1.399	1.514	.262
21	.705	.273	.211	46	-	-	-
22	.713	1.189	.119	47	-	-	-
23	1.013	.550	.186	48	-	-	-
24	1.389	-.013	.238	49	-	-	-
25	.817	3.567	.291	50	-	-	-
51	-	-	-	56	-	-	-

ตาราง 10 (ต่อ)

ข้อ	พารามิเตอร์			ข้อ	พารามิเตอร์		
	a	b	c		a	b	c
52	-	-	-	57	-	-	-
53	-	-	-	58	-	-	-
54	-	-	-	59	-	-	-
55	-	-	-	60	-	-	-

ตาราง 11 ค่าพารามิเตอร์ของแบบทดสอบฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก

ข้อ	พารามิเตอร์			ข้อ	พารามิเตอร์		
	a	b	c		a	b	c
1	.470	.679	.180	26	.833	.381	.243
2	1.003	.420	.072	27	-	-	-
3	.723	-.102	.116	28	-	-	-
4	.828	-.359	.164	29	-	-	-
5	1.023	-1.255	.156	30	-	-	-
6	.947	-.044	.334	31	1.659	.666	.287
7	.953	1.820	.166	32	1.394	1.009	.247
8	.791	1.537	.202	33	1.330	.659	.236
9	.779	1.220	.128	34	1.488	1.216	.176
10	-	-	-	35	.670	2.336	.136
11	1.304	2.740	.255	36	1.138	1.592	.166
12	1.027	-.649	.109	37	.882	2.001	.201
13	.643	-.212	.144	38	1.255	1.907	.136
14	1.038	.436	.172	39	1.026	1.715	.175
15	.947	-.647	.187	40	1.346	1.454	.171
16	.666	.916	.121	41	1.164	2.002	.227
17	.763	1.311	.165	42	1.151	1.185	.207
18	-	-	-	43	-	-	-
19	.910	2.031	.147	44	-	-	-
20	1.111	1.884	.176	45	-	-	-
21	-	-	-	46	-	-	-
22	.729	1.077	.074	47	-	-	-
23	.785	.577	.183	48	-	-	-
24	1.059	.226	.133	49	-	-	-
25	1.046	3.043	.204	50	-	-	-

ตาราง 11 (ต่อ)

ข้อ	พารามิเตอร์			ข้อ	พารามิเตอร์		
	a	b	c		a	b	c
51	-	-	-	56	-	-	-
52	-	-	-	57	-	-	-
53	-	-	-	58	-	-	-
54	-	-	-	59	-	-	-
55	-	-	-	60	-	-	-

ภาคผนวก ก.

ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ (TIF) และค่าความคลาดเคลื่อนมาตรฐานในการ
ประมาณค่าความสามารถของผู้สอบ (SEE) ระหว่างแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก
ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก ที่ระดับความสามารถตั้งแต่ -
4.00 ถึง 4.00 จำนวน 81 ระดับความสามารถ

ตาราง 12 ค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ (TIF) และค่าความคลาดเคลื่อนมาตรฐาน
ในการประมาณค่าความสามารถของผู้สอบ (SEE) ระหว่างแบบทดสอบฉบับที่ 1 มี 59 ข้อ
3 ตัวเลือก ฉบับที่ 1 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 1 มี 35 ข้อ 5 ตัวเลือก ที่ระดับความ
สามารถตั้งแต่ -4.00 ถึง 4.00 จำนวน 81 ระดับความสามารถ

ระดับความ สามารถ (θ)	59 ข้อ 3 ตัวเลือก		44 ข้อ 4 ตัวเลือก		35 ข้อ 5 ตัวเลือก	
	TIF	SEE	TIF	SEE	TIF	SEE
-4.00	0.0028	2.4742	0.0030	2.7605	0.0024	3.4656
-3.00	0.0033	2.2755	0.0035	2.5394	0.0028	3.1859
-3.80	0.0039	2.0933	0.0042	2.3366	0.0033	2.9293
-3.70	0.0046	1.9262	0.0049	2.1507	0.0039	2.6937
-3.60	0.0054	1.7730	0.0058	1.9804	0.0047	2.4777
-3.50	0.0064	1.6326	0.0068	1.8243	0.0055	2.2796
-3.40	0.0075	1.5040	0.0080	1.6814	0.0065	2.0979
-3.30	0.0088	1.3846	0.0095	1.5506	0.0077	1.9314
-3.20	0.0104	1.2784	0.0111	1.4309	0.0090	1.7788
-3.10	0.0122	1.1798	0.0130	1.3215	0.0106	1.6390
-3.00	0.0143	1.0897	0.0152	1.2216	0.0125	1.5111
-2.90	0.0167	1.0073	0.0178	1.1304	0.0147	1.3940
-2.80	0.0193	0.9321	0.0207	1.0472	0.0172	1.2870
-2.70	0.0227	0.8635	0.0241	0.9714	0.0202	1.1892
-2.60	0.0264	0.8010	0.0279	0.9024	0.0236	1.0999
-2.50	0.0306	0.7441	0.0322	0.8396	0.0275	1.0184
-2.40	0.0354	0.6923	0.0371	0.7826	0.0321	0.9442
-2.30	0.0407	0.6153	0.0425	0.7300	0.0372	0.8765
-2.20	0.0467	0.6025	0.0486	0.6839	0.0430	0.8151
-2.10	0.0533	0.5638	0.0553	0.6413	0.0496	0.7592
-2.00	0.0607	0.5286	0.0626	0.6027	0.0567	0.7085

ตาราง 12 (ต่อ)

ระดับความ สามารถ (๐)	59 ข้อ 3 ตัวเลือก		44 ข้อ 4 ตัวเลือก		35 ข้อ 5 ตัวเลือก	
	TIF	SEE	TIF	SEE	TIF	SEE
-1.90	0.0687	0.4967	0.0705	0.5677	0.0651	0.6625
-1.80	0.0774	0.4679	0.0791	0.5360	0.0741	0.6209
-1.70	0.0869	0.4417	0.0884	0.5071	0.0840	0.5832
-1.60	0.0970	0.4180	0.0983	0.4809	0.0948	0.5491
-1.50	0.1078	0.3966	0.1088	0.4571	0.1064	0.5182
-1.40	0.1192	0.3771	0.1199	0.4353	0.1189	0.4903
-1.30	0.1312	0.3594	0.1317	0.4155	0.1321	0.4650
-1.20	0.1438	0.3434	0.1440	0.3973	0.1462	0.4421
-1.10	0.1569	0.3287	0.1569	0.3806	0.1610	0.4212
-1.00	0.1706	0.3152	0.1703	0.3653	0.1766	0.4023
-0.90	0.1849	0.3028	0.1842	0.3512	0.1928	0.3850
-0.80	0.1998	0.2913	0.1987	0.3382	0.2097	0.3691
-0.70	0.2153	0.2806	0.2136	0.3262	0.2272	0.3546
-0.60	0.2316	0.2705	0.2291	0.3150	0.2453	0.3413
-0.50	0.2487	0.2610	0.2450	0.3046	0.2640	0.3290
-0.40	0.2667	0.2521	0.2613	0.2949	0.2832	0.3176
-0.30	0.2856	0.2436	0.2781	0.2859	0.3029	0.3071
-0.20	0.3055	0.2355	0.2952	0.2775	0.3230	0.2974
-0.10	0.3265	0.2278	0.3127	0.2696	.03433	0.2885
0.00	0.3485	0.2205	0.3305	0.2622	0.3637	0.2803
0.10	0.3717	0.2135	0.3485	0.2554	0.3839	0.2728
0.20	0.3960	0.2069	0.3667	0.2490	0.4037	0.2660
0.30	0.4214	0.2006	0.3850	0.2430	0.4227	0.2600
0.40	0.4477	0.1946	0.4032	0.2374	0.4406	0.2546
0.50	04749	0.1889	0.4211	0.2323	0.4572	0.2500

ตาราง 12 (ต่อ)

ระดับความ สามารถ (θ)	59 ข้อ 3 ตัวเลือก		44 ข้อ 4 ตัวเลือก		35 ข้อ 5 ตัวเลือก	
	TIF	SEE	TIF	SEE	TIF	SEE
0.60	0.5027	0.1836	0.4386	0.2276	0.4720	0.2460
0.70	0.5306	0.1787	0.4554	0.2234	0.4849	0.2427
0.80	0.5583	0.1742	0.4710	0.2197	0.4955	0.2401
0.90	0.5850	0.1702	0.4851	0.2165	0.5036	0.2382
1.00	0.6102	0.1667	0.4973	0.2138	0.5092	0.2369
1.10	0.6331	0.1636	0.5071	0.2117	0.5120	0.2362
1.20	0.6528	0.1611	0.5143	0.2102	0.5120	0.2362
1.30	0.6686	0.1592	0.5183	0.2094	0.5093	0.2369
1.40	0.6799	0.1579	0.5190	0.2093	0.5036	0.2382
1.50	0.6862	0.1572	0.5162	0.2098	0.4953	0.2402
1.60	0.6871	0.1571	0.5097	0.2112	0.4842	0.2429
1.70	0.6824	0.1576	0.4996	0.2133	0.4706	0.2464
1.80	0.6724	0.1588	0.4861	0.2162	0.4546	0.2507
1.90	0.6572	0.1606	0.4693	0.2201	0.4366	0.2558
2.00	0.6373	0.1630	0.4498	0.2248	0.4167	0.2618
2.10	0.6133	0.1662	0.4278	0.2305	0.3954	0.2688
2.20	0.5856	0.1701	0.4040	0.2372	0.3731	0.2767
2.30	0.5551	0.1747	0.3788	0.2449	0.3500	0.2857
2.40	0.5223	0.1801	0.3529	0.2538	0.3266	0.2958
2.50	0.4880	0.1864	0.3266	0.2638	0.3031	0.3070
2.60	0.4527	0.1935	0.3005	0.2750	0.2800	0.3195
2.70	0.4172	0.2016	0.2751	0.2874	0.2573	0.3332
2.80	0.3819	0.2107	0.2506	0.3011	0.2354	0.3484
2.90	0.3475	0.2209	0.2274	0.3060	0.2143	0.3651
3.00	0.3142	0.2323	0.2057	0.3324	0.1942	0.3835

ตาราง 12 (ต่อ)

ระดับความ สามารถ (θ)	59 ข้อ 3 ตัวเลือก		44 ข้อ 4 ตัวเลือก		35 ข้อ 5 ตัวเลือก	
	TIF	SEE	TIF	SEE	TIF	SEE
3.10	0.2825	0.2449	0.1854	0.3501	0.1752	0.4038
3.20	0.2527	0.2590	0.1668	0.3692	0.1573	0.4263
3.30	0.2249	0.2745	0.1496	0.3897	0.1405	0.4510
3.40	0.1993	0.2917	0.1340	0.4119	0.1249	0.4783
3.50	0.1758	0.3105	0.1197	0.4358	0.1105	0.5085
3.60	0.1544	0.3313	0.1066	0.4617	0.0973	0.5418
3.70	0.1352	0.3541	0.0948	0.4897	0.0853	0.5786
3.80	0.1180	0.3791	0.0840	0.5202	0.0745	0.6192
3.90	0.1026	0.4065	0.0742	0.5534	0.0648	0.6640
4.00	0.0889	0.4366	0.0653	0.5898	0.0562	0.7133

ภาคผนวก ง.

ค่าประสิทธิภาพสัมพัทธ์ ของการประมาณค่าความสามารถของผู้สอบระหว่างแบบทดสอบ
ฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก
ที่ระดับความสามารถตั้งแต่ -4.00 ถึง 4.00 จำนวน 81 ระดับความสามารถ

ตาราง 13 ค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าความสามารถของผู้สอบ ระหว่าง
แบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 1 มี
35 ข้อ 5 ตัวเลือก ที่ระดับความสามารถตั้งแต่ -4.00 ถึง 4.00 จำนวน 81 ระดับความสามารถ

ระดับความ สามารถ (θ)	ค่าประสิทธิภาพสัมพัทธ์		
	59 ข้อ 3 ตัวเลือก	44 ข้อ 4 ตัวเลือก	35 ข้อ 5 ตัวเลือก
-4.00	0.9288	1.1637	1.2528
-3.90	0.9285	1.1629	1.2525
-3.80	0.9289	1.1620	1.2509
-3.70	0.9296	1.1602	1.2480
-3.60	0.9303	1.1584	1.2452
-3.50	0.9312	1.1568	1.2424
-3.40	0.9321	1.1543	1.2384
-3.30	0.9332	1.1515	1.2340
-3.20	0.9342	1.1485	1.2295
-3.10	0.9356	1.1450	1.2237
-3.00	0.9373	1.1409	1.2173
-2.90	0.9391	1.1361	1.2007
-2.80	0.9412	1.1309	1.2015
-2.70	0.9436	1.1250	1.0922
-2.60	0.9464	1.1185	1.1817
-2.50	0.9495	1.1112	1.1703
-2.40	0.9529	1.1032	1.1577
-2.30	0.9567	1.0947	1.1442
-2.20	0.9607	1.0855	1.1299
-2.10	0.9651	1.0758	1.1148
-2.00	0.9695	1.0657	1.0992

ตาราง 13 (ต่อ)

ระดับความ สามารถ (θ)	ค่าประสิทธิภาพสัมพัทธ์		
	59 ข้อ 3 ตัวเลือก	44 ข้อ 4 ตัวเลือก	35 ข้อ 5 ตัวเลือก
-1.90	0.9741	1.0553	1.0834
-1.80	0.9787	1.0447	1.0676
-1.70	0.9830	1.0340	1.0520
-1.60	0.9870	1.0234	1.0369
-1.50	0.9907	1.0130	1.0226
-1.40	0.9938	1.0027	1.0090
-1.30	0.9964	0.9928	0.9964
-1.20	0.9985	0.9833	0.9848
-1.10	1.0003	0.9744	0.9741
-1.00	1.0019	0.9662	0.9644
-0.90	1.0036	0.9590	0.9556
-0.80	1.0055	0.9528	0.9476
-0.70	1.0080	0.9478	0.9404
-0.60	1.0112	0.9442	0.9338
-0.50	1.0154	0.9422	0.9279
-0.40	1.0207	0.9416	0.9226
-0.30	1.0272	0.9428	0.9179
-0.20	1.0350	0.9458	0.9136
-0.10	1.0441	0.9509	0.9107
0.00	1.0547	0.9583	0.9086
0.10	1.0666	0.9682	0.9077
0.20	1.0799	0.9809	0.9083
0.30	1.0946	0.9968	0.9107
0.40	1.1105	1.0161	0.9149
0.50	1.1277	1.0388	0.9211

ตาราง 13 (ต่อ)

ระดับความ สามารถ (θ)	ค่าประสิทธิภาพสัมพัทธ์		
	59 ข้อ 3 ตัวเลือก	44 ข้อ 4 ตัวเลือก	35 ข้อ 5 ตัวเลือก
0.60	1.1460	1.0649	1.9292
0.70	1.1653	1.0944	1.9391
0.80	1.1854	1.1268	1.9506
0.90	1.2060	1.1617	1.9632
1.00	1.2272	1.1985	1.9767
1.10	1.2483	1.2364	1.9905
1.20	1.2693	1.2748	1.0043
1.30	1.2899	1.3129	1.0179
1.40	1.3100	1.3500	1.0305
1.50	1.3294	1.3855	1.0422
1.60	1.3480	1.4190	1.0526
1.70	1.3660	1.4502	1.0616
1.80	1.3834	1.4790	1.0691
1.90	1.4004	1.5054	1.0750
2.00	1.4170	1.5293	1.0793
2.01	1.4334	1.5508	1.0819
2.20	1.4495	1.5697	1.0829
2.30	1.4652	1.5860	1.0825
2.40	1.4802	1.5994	1.0805
2.50	1.4941	1.6098	1.0774
2.60	1.5065	1.6171	1.0734
2.70	1.5167	1.6212	1.0690
2.80	1.5239	1.6225	1.0647
2.90	1.5278	1.6211	1.0611
3.00	1.5277	1.6177	1.0589

ตาราง 13 (ต่อ)

ระดับความ สามารถ (๒)	ค่าประสิทธิภาพสัมพัทธ์		
	59 ข้อ 3 ตัวเลือก	44 ข้อ 4 ตัวเลือก	35 ข้อ 5 ตัวเลือก
3.10	1.5236	1.6127	1.0585
3.20	1.5052	1.6070	1.0605
3.30	1.5030	1.6011	1.0652
3.40	1.4874	1.5955	1.0727
3.50	1.4689	1.5907	1.0829
3.60	1.4484	1.5869	1.0956
3.70	1.4267	1.5844	1.1105
3.80	1.4044	1.5830	1.1272
3.90	1.3824	1.5828	1.1450
4.00	1.3611	1.5837	1.1635

ภาคผนวก จ.

ความน่าจะเป็นในการตอบแบบทดสอบได้ถูก ของแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3 ตัวเลือก
ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก จำนวน 81 ระดับความสามารถ

ตาราง 14 ความน่าจะเป็นในการตอบแบบทดสอบได้ถูกของแบบทดสอบฉบับที่ 1 มี 59 ข้อ 3
ตัวเลือก ฉบับที่ 2 มี 44 ข้อ 4 ตัวเลือก และฉบับที่ 3 มี 35 ข้อ 5 ตัวเลือก จำนวน 81 ระดับ
ความสามารถ

ระดับความ สามารถ (θ)	แบบทดสอบ			ระดับความ สามารถ (θ)	แบบทดสอบ		
	59 ข้อ	44 ข้อ	35 ข้อ		59 ข้อ	44 ข้อ	35 ข้อ
-4.00	0.0010	0.0012	0.0009	-1.70	0.0401	0.0472	0.0385
-3.90	0.0012	0.0014	0.0011	-1.60	0.0459	0.0540	0.0444
-3.80	0.0015	0.0017	0.0013	-1.50	0.0523	0.0616	0.0510
-3.70	0.0017	0.0020	0.0015	-1.40	0.0595	0.0700	0.0585
-3.60	0.0020	0.0024	0.0018	-1.30	0.0673	0.0793	0.0668
-3.50	0.0024	0.0038	0.0021	-1.20	0.0760	0.0895	0.0759
-3.40	0.0028	0.0033	0.0025	-1.10	0.0854	0.1006	0.0860
-3.30	0.0034	0.0040	0.0030	-1.00	0.0957	0.1126	0.0971
-3.20	0.0040	0.0047	0.0035	-0.90	0.1068	0.1257	0.1092
-3.10	0.0047	0.0055	0.0042	-0.80	0.1187	0.1397	0.1222
-3.00	0.0055	0.0065	0.0049	-0.70	0.1316	0.1546	0.1363
-2.90	0.0065	0.0076	0.0058	-0.60	0.1453	0.1705	0.1514
-2.80	0.0076	0.0090	0.0069	-0.50	0.1599	0.1874	0.1676
-2.70	0.0090	0.0106	0.0081	-0.40	0.1754	0.2051	0.1847
-2.60	0.0105	0.0124	0.0095	-0.30	0.1917	0.2236	0.2029
-2.50	0.0123	0.0145	0.0112	-0.20	0.2089	0.2429	0.2221
-2.40	0.0144	0.0169	0.0132	-0.10	0.2270	0.2630	0.2422
-2.30	0.0168	0.0198	0.0155	0.00	0.2459	0.2836	0.2633
-2.20	0.0195	0.0230	0.0181	0.10	0.2655	0.3049	0.2852
-2.10	0.227	0.0267	0.211	0.20	0.2859	0.3266	0.3080
-2.00	0.0263	0.0309	0.0246	0.30	0.3071	0.3488	0.3316
-1.90	0.0304	0.0357	0.0287	0.40	0.3290	0.3713	0.3559

ตาราง 14 (ต่อ)

ระดับความ สามารถ (θ)	แบบทดสอบ			ระดับความ สามารถ (θ)	แบบทดสอบ		
	59 ข้อ	44 ข้อ	35 ข้อ		59 ข้อ	44 ข้อ	35 ข้อ
0.50	0.3516	0.3942	0.3809	2.30	0.7973	0.7985	0.8243
0.60	0.3748	0.4173	0.4064	2.40	0.8167	0.8158	0.8419
0.70	0.3987	0.4407	0.4324	2.50	0.8348	0.8320	0.8583
0.80	0.4232	0.4642	0.4588	2.60	0.8517	0.8473	0.8735
0.90	0.4483	0.4879	0.4854	2.70	0.8673	0.8615	0.8874
1.00	0.4739	0.5118	0.5122	2.80	0.8817	0.8748	0.9002
1.10	0.4999	0.5356	0.5391	2.90	0.8949	0.8871	0.9112
1.20	0.5263	0.5595	0.5659	3.00	0.9070	0.8984	0.9225
1.30	0.5528	0.5833	0.5926	3.10	0.9179	0.9089	0.9320
1.40	0.5795	0.6070	0.6189	3.20	0.9278	0.9185	0.9406
1.50	0.6061	0.6305	0.6448	3.30	0.9367	0.9274	0.9483
1.60	0.6324	0.6536	0.6703	3.40	0.9447	0.9355	0.9551
1.70	0.6584	0.6763	0.6950	3.50	0.9518	0.9428	0.9611
1.80	0.6838	0.6985	0.7190	3.60	0.9581	0.9496	0.9664
1.90	0.7084	0.7201	0.7421	3.70	0.9637	0.9556	0.9711
2.00	0.7322	0.7410	0.7643	3.80	0.9686	0.9611	0.9752
2.10	0.7550	0.7610	0.7854	3.90	0.9729	0.9660	0.9787
2.20	0.7768	0.7802	0.8054	4.00	0.9767	0.9704	0.9818

ภาคผนวก ง.

- ค่าความยาก ค่าอำนาจจำแนก และค่าความยากมาตรฐานของเครื่องมือ
- ตัวอย่างแบบทดสอบที่ใช้ในการวิจัย

ตาราง 15 ค่าความยาก ค่าอำนาจจำแนก และค่าความยากมาตรฐานของเครื่องมือ

ข้อ	ตัวเลือก	p	r	Δ	ข้อ	ตัวเลือก	p	r	Δ
1	1	.18	-.12	16.7	6	(1)	.55	.54	12.5
	(2)	.55	.54	12.5		2	.12	-.44	17.8
	3	.14	-.65	17.4		3	.11	.42	17.8
	4	.09	-.22	18.3		4	.10	-.26	18.1
	5	.05	-.39	19.7		5	.09	-.07	18.2
2	1	.06	-.10	19.1	7	1	.19	-.36	16.5
	2	.12	-.34	17.6		2	.21	-.30	16.3
	3	.21	-.53	16.2		(3)	.25	.68	15.6
	4	.11	-.30	17.9		4	.13	.09	17.4
	(5)	.40	.70	14.0		5	.14	-.25	17.2
3	1	.07	-.60	18.9	8	1	.13	-.36	17.4
	2	.10	-.19	18.2		2	.02	.00	-
	(3)	.56	.55	12.3		(3)	.73	.52	10.5
	4	.08	-.52	18.7		4	.08	-.52	18.7
	5	.17	-.40	16.8		5	.03	.00	-
4	1	.14	-.25	17.2	9	1	.17	-.05	16.7
	2	.17	-.44	16.9		2	.15	-.14	17.2
	(3)	.53	.51	12.7		3	.10	-.57	18.2
	4	.05	-.19	19.5		4	.17	-.21	16.8
	5	.08	-.17	18.7		(5)	.36	.51	14.4
5	1	.10	-.19	18.2	10	1	.15	-.67	17.1
	(2)	.58	.53	12.2		2	.03	.00	-
	3	.10	.38	18.2		(3)	.53	.60	12.7
	4	.11	.24	17.9		4	.18	-.12	16.7
	5	.11	.42	17.9		5	.07	-.26	19.1

ตาราง 15 (ต่อ)

ข้อ	ตัวเลือก	p	r	Δ	ข้อ	ตัวเลือก	p	r	Δ
11	1	.30	-.27	15.1	16	1	.17	-.40	16.8
	(2)	.50	.50	13.0		2	.03	.00	-
	3	.01	.00	-		(3)	.67	.61	11.3
	4	.07	-.50	18.9		4	.07	-.26	19.1
	5	.10	-.19	18.2		5	.07	-.50	18.9
12	1	.08	-.52	18.7	17	1	.07	-.50	18.9
	2	.04	.00	-		2	.08	-.13	18.5
	3	.07	-.50	18.9		(3)	.66	.55	11.4
	(4)	.78	.75	9.9		4	.11	-.24	17.9
	5	.05	-.41	19.6		5	.07	-.50	18.9
13	1	.04	.00	-	18	1	.20	-.19	16.3
	2	.07	-.50	18.9		2	.13	-.09	17.4
	3	.08	-.31	18.7		(3)	.59	.47	12.1
	4	.10	-.19	18.2		4	.08	-.31	18.7
	(5)	.72	.53	10.7		5	.05	-.19	19.5
14	1	.16	-.42	17.0	19	(1)	.35	.63	14.6
	2	.12	-.34	17.6		2	.09	-.07	18.2
	3	.06	-.46	19.2		3	.10	-.57	18.1
	4	.06	-.10	19.1		4	.12	-.27	17.7
	(5)	.57	.57	12.3		5	.27	-.21	15.4
15	1	.09	-.22	18.3	20	1	.23	-.12	15.9
	(2)	.52	.64	12.8		2	.14	-.34	17.3
	3	.13	-.21	17.5		(3)	.05	-.19	19.5
	4	.13	-.47	17.5		4	.38	.60	14.2
	5	.09	-.55	18.4		5	.13	-.48	17.4

ตาราง 15 (ต่อ)

ข้อ	ตัวเลือก	p	r	Δ	ข้อ	ตัวเลือก	p	r	Δ
21	1	.11	-.42	17.9	26	1	.05	.41	19.6
	(2)	.54	.45	12.6		(2)	.68	.52	11.2
	3	.08	-.34	18.5		3	.11	.42	17.9
	4	.17	-.21	16.9		4	.05	.19	19.5
	5	.08	.00	-		5	.11	.24	17.9
22	1	.13	-.21	17.5	27	1	.18	-.22	16.7
	2	.26	-.60	15.6		2	.15	-.35	17.2
	3	.11	-.13	17.9		3	.08	-.13	18.5
	4	.05	-.19	19.5		(4)	.50	.44	13.0
	(5)	.37	.72	14.3		5	.06	.00	19.2
23	1	.19	-.37	16.5	28	1	.14	-.25	17.2
	2	.07	-.26	19.1		2	-	-	-
	3	.14	-.11	17.3		3	.32	-.23	14.9
	(4)	.50	.50	13.0		(4)	.47	.44	13.3
	5	.08	-.31	18.7		5	.05	-.41	19.6
24	1	.11	-.24	17.9	29	1	.12	-.34	17.6
	2	.15	-.52	17.1		2	.09	-.55	18.4
	3	.10	-.38	18.2		(3)	.65	.49	11.5
	4	.05	-.41	19.6		4	.10	-.26	18.1
	(5)	.55	.69	12.5		5	.01	.00	-
25	1	.28	.23	15.3	30	1	.25	-.24	15.7
	2	.18	-.34	16.7		2	.15	-.28	17.1
	3	.17	-.56	16.8		3	.09	-.22	18.3
	(4)	.23	.44	16.0		(4)	.46	.48	13.5
	5	.08	.17	18.7		5	.02	.00	-

ตาราง 15 (ต่อ)

ข้อ	ตัวเลือก	p	r	Δ	ข้อ	ตัวเลือก	p	r	Δ
31	1	.05	-.41	19.6	36	1	.23	-.25	16.0
	(2)	.64	.64	11.5		2	.11	-.13	17.9
	3	.18	-.42	16.7		3	.16	.42	17.0
	4	.07	-.50	18.9		(4)	.31	.49	15.0
	5	.05	-.19	19.5		5	.16	.08	16.9
32	1	.09	-.07	18.2	37	1	.08	-.31	18.7
	(2)	.52	.67	12.8		(2)	.47	.50	13.3
	3	.11	-.42	17.9		3	.19	-.26	16.5
	4	.09	-.55	18.4		4	.09	-.55	18.4
	5	.12	-.45	17.7		5	.14	.00	17.3
33	1	.10	-.38	18.2	38	1	.05	-.41	19.6
	2	.12	-.44	17.8		2	.19	.07	16.5
	3	.05	-.19	19.5		3	.29	-.17	15.2
	(4)	.69	.58	11.0		(4)	.20	.60	16.4
	5	.04	.00	-		5	.21	-.31	16.2
34	1	.28	.23	15.3	39	1	.15	.14	17.2
	(2)	.43	.50	13.7		(2)	.30	.57	15.1
	3	.09	-.55	18.4		3	.26	-.41	15.6
	4	.11	.00	17.9		4	.05	-.41	19.6
	5	.07	-.26	19.1		5	.18	-.22	16.7
35	1	.32	.23	14.9	40	1	.20	-.28	16.4
	2	.21	-.40	16.3		2	.19	-.49	16.5
	3	.09	-.36	18.4		3	.08	.00	18.6
	4	.12	-.27	17.7		(4)	.44	.51	13.6
	(5)	.22	.50	16.2		5	.05	.00	19.6

ตาราง 15 (ต่อ)

ข้อ	ตัวเลือก	p	r	Δ	ข้อ	ตัวเลือก	p	r	Δ
41	1	.11	-.24	17.9	46	1	.06	-.10	19.1
	2	.17	-.32	16.8		2	.12	-.34	17.6
	(3)	.35	.48	14.6		3	.21	-.53	16.2
	4	.26	-.24	15.6		4	.11	-.30	17.9
	5	.05	.19	19.5		(5)	.40	.70	14.0
42	1	.21	-.40	16.3	47	1	.16	-.18	16.9
	(2)	.37	.71	14.3		2	.12	-.27	17.7
	3	.18	-.42	16.7		3	.53	.44	12.7
	4	.14	-.11	17.3		4	.11	-.42	17.9
	5	.04	.00	-		(5)	.06	.00	19.2
43	1	.13	-.09	17.4	48	1	.12	.44	17.8
	2	.11	-.13	17.9		2	.21	.15	16.2
	3	.18	-.47	16.7		(3)	.13	-.21	17.5
	(4)	.38	.48	14.2		4	.24	-.14	15.8
	5	.16	-.08	16.9		5	.26	.40	15.6
44	1	.18	-.12	16.7	49	1	.18	-.14	16.6
	(2)	.36	.38	14.5		2	.17	.00	16.8
	3	.14	-.25	17.2		(3)	.26	.31	15.6
	4	.22	-.22	16.1		4	.13	-.48	17.4
	5	.09	.07	18.2		5	.22	.15	16.0
45	(1)	.24	.47	.47	50	1	.10	-.19	18.2
	2	.17	.00	.00		(2)	.54	.43	12.6
	3	.27	-.13	-.13		3	.19	-.16	16.5
	4	.12	-.18	-.18		4	.12	-.45	17.7
	5	.12	-.18	-.18		5	.03	.00	-

ตาราง 15 (ต่อ)

ข้อ	ตัวเลือก	p	r	Δ	ข้อ	ตัวเลือก	p	r	Δ
51	1	.09	-.53	18.4	56	1	.12	-.44	17.8
	2	.15	-.04	17.1		2	.23	.11	16.0
	(3)	.47	.41	13.3		3	.15	.14	17.2
	4	.15	-.04	17.1		(4)	.29	.32	15.2
	5	.12	.27	17.7		5	.18	-.34	16.7
52	1	.32	.06	14.8	57	1	.04	.00	-
	(2)	.32	.39	14.8		2	.20	-.39	16.4
	3	.07	-.26	19.1		(3)	.46	.39	13.4
	4	.16	-.18	16.9		4	.12	-.16	17.8
	5	.09	-.55	18.4		5	.16	.08	16.9
53	1	.33	.03	14.7	58	1	.20	-.12	16.4
	2	.08	-.31	18.7		2	.24	.00	15.8
	3	.10	-.38	18.2		(3)	.34	.35	14.6
	(4)	.41	.41	13.9		4	.13	-.31	17.5
	5	.05	-.41	19.6		5	.07	-.26	19.1
54	(1)	.40	.34	14.0	59	1	.12	-.18	17.7
	2	.10	.19	18.2		(2)	.29	.39	15.2
	3	.13	-.21	17.5		3	.25	-.21	15.7
	4	.24	-.21	15.7		4	.16	-.08	16.9
	5	.12	-.27	17.7		5	.15	-.05	17.1
55	1	.16	-.18	16.9	60	(1)	.43	.39	13.8
	2	.16	-.18	16.9		2	.13	-.21	17.5
	3	.37	.10	14.3		3	.15	-.39	17.2
	4	.06	-.46	19.2		4	.16	-.08	16.9
	(5)	.21	.40	16.3		5	.11	.00	17.9

หมายเลขข้อสอบ.....

แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6

เรื่อง - ตัวประกอบของจำนวนนับ

- เศษส่วน

คำชี้แจง

1. แบบทดสอบฉบับนี้มีข้อคำถาม 60 ข้อ ให้เวลาทำ 1 ชั่วโมง 30 นาที
2. ข้อคำถามทั้งหมดเป็นแบบปรนัยชนิดเลือกตอบ 3 ตัวเลือกให้นักเรียน พิจารณาเลือกคำตอบที่ถูกต้องที่สุด หรือเหมาะสมที่สุดเพียงคำตอบเดียวแล้วนำไปตอบในกระดาษคำตอบที่จัดไว้ให้
3. โปรดอย่าขีดเขียนหรือทำเครื่องหมายใด ๆ ลงในแบบทดสอบฉบับนี้ ถ้านักเรียนต้องการจะทด ให้ทดลงในหลังกระดาษคำตอบ
4. ถ้าพบว่ามีข้อคำถามข้อใดยากอย่าเดาคำตอบ หรืออย่าท้อใจ ให้เว้นข้ามไปทำข้อถัดไปก่อน เพราะอาจจะมีข้อที่ง่ายหรือข้อที่นักเรียนทำได้อยู่ตอนท้าย ๆ ก็ได้ เมื่อมีเวลาเหลือจึงค่อยกลับมาทำข้อที่เว้นไว้
5. ก่อนที่จะทำแบบทดสอบฉบับนี้ ให้นักเรียนกรอกรายละเอียดในส่วนบนของกระดาษคำตอบให้เรียบร้อย
6. วิธีตอบ
 - 6.1 เมื่อนักเรียนได้คำตอบหรือตัวเลือกใดแล้ว ให้ทำเครื่องหมาย X ทับบนหมายเลขของตัวเลือกในกระดาษคำตอบ ดังตัวอย่าง

0. (ก) (ข) (ค) (ง) (จ)

00. (ก) (ข) (ค) (ง) (จ)
 - 6.2 ถ้าต้องการเปลี่ยนคำตอบ ให้ทำเครื่องหมาย = ทับคำตอบเดิมเสียก่อน แล้วทำเครื่องหมายกากบาท X ทับหมายเลขของตัวเลือกที่ต้องการ เช่น ข้อ 0. ถ้าต้องการเปลี่ยนคำตอบจาก (ข) เป็น (ง) ให้ทำดังนี้

0. (ก) (ข) (ค) (ง) (จ)

1. ถ้า $8 : 2 = 4$

คำกล่าวในข้อใดถูกต้องที่สุด

- ก. 2 เป็นตัวประกอบของ 4
- ข. 2 เป็นตัวประกอบของ 8
- ค. 4 เป็นตัวประกอบของ 2
- ง. 8 เป็นตัวประกอบของ 2
- จ. 8 เป็นตัวประกอบของ 4

2. จำนวนใดมี 13 เป็นตัวประกอบ

- ก. 225
- ข. 285
- ค. 345
- ง. 385
- จ. 455

3. จำนวนใดไม่เป็นตัวประกอบของ 102

- ก. 2
- ข. 3
- ค. 13
- ง. 17
- จ. 15

4. ข้อใดมี 4 เป็นตัวประกอบทุกจำนวน

- ก. 8, 12, 18
- ข. 16, 20, 26
- ค. 24, 28, 52
- ง. 40, 56, 62
- จ. 48, 72, 86

5. จำนวน 2, 3, 4 เป็นตัวประกอบ
ของจำนวนในข้อใด

- ก. 8
- ข. 12
- ค. 15
- ง. 18
- จ. 26

6. จำนวนในข้อใดเป็นจำนวนเฉพาะ

- ก. 17
- ข. 27
- ค. 57
- ง. 77
- จ. 87

7. จำนวนในข้อใดไม่เป็นจำนวนเฉพาะ

- ก. 97
- ข. 101
- ค. 201
- ง. 307
- จ. 401

8. ข้อใดเป็นจำนวนเฉพาะทั้งหมด

- ก. 31, 41, 51
- ข. 33, 47, 53
- ค. 37, 47, 59
- ง. 39, 49, 57
- จ. 49, 57, 69

9. คำกล่าวข้อใดไม่ถูกต้อง

- ก. 2, 19 เป็นตัวประกอบเฉพาะของ 38
- ข. 5, 13 เป็นตัวประกอบเฉพาะของ 65
- ค. 3, 31 เป็นตัวประกอบเฉพาะของ 93
- ง. 3, 37 เป็นตัวประกอบเฉพาะของ 111
- จ. 2, 33 เป็นตัวประกอบเฉพาะของ 198

10. จำนวนเฉพาะระหว่าง 29 ถึง 42 มีกี่จำนวน

- ก. 1
- ข. 2
- ค. 3
- ง. 4
- จ. 5

11. ตัวประกอบเฉพาะของ 20 เมื่อคูณกันแล้วจะได้ผลลัพธ์เป็นเท่าไร

- ก. 20
- ข. 10
- ค. 5
- ง. 4
- จ. 2

12. จำนวนใดที่มี 2, 5, 8 และ 10 เป็นตัวประกอบ

- ก. 20
- ข. 24
- ค. 30
- ง. 40
- จ. 48

13. จำนวน 552 แยกเป็นตัวประกอบได้ตามข้อใด

- ก. $2 \times 2 \times 3 \times 21$
- ข. $2 \times 2 \times 3 \times 23$
- ค. $2 \times 3 \times 3 \times 23$
- ง. $2 \times 3 \times 7 \times 21$
- จ. $2 \times 2 \times 2 \times 3 \times 23$

14. ข้อใดแยกตัวประกอบได้ถูกต้อง

- ก. $48 = 6 \times 8$
- ข. $42 = 2 \times 21$
- ค. $36 = 2 \times 2 \times 9$
- ง. $32 = 2 \times 2 \times 8$
- จ. $28 = 2 \times 2 \times 7$

15. ข้อใดเป็นตัวประกอบร่วมของ 39 กับ 78

- ก. 11
- ข. 13
- ค. 15
- ง. 17
- จ. 19

16. ตัวประกอบร่วมที่มากที่สุดของ 36, 68, 92 คือข้อใด
 ก. 2
 ข. 3
 ค. 4
 ง. 6
 จ. 7
17. 15 เป็น ห.ร.ม. ของจำนวนในข้อใด
 ก. 36, 48, 72
 ข. 36, 54, 72
 ค. 15, 75, 90
 ง. 54, 72, 126
 จ. 72, 120, 132
18. ห.ร.ม. ของ 24, 36, 48 และ 60 คือข้อใด
 ก. 4
 ข. 8
 ค. 12
 ง. 36
 จ. 72
19. ถ้า $K = 2^3 \times 3$, $P = 2^4 \times 3$, $Q = 2^2 \times 5$ ห.ร.ม. Q ของ K,P, คือข้อใด
 ก. 2^2
 ข. $2^2 \times 3^2$
 ค. $2^3 \times 3^2$
 ง. $2 \times 3^2 \times 5$
 จ. $2^2 \times 3^2 \times 5^2$
20. ห.ร.ม. ของ 57 และ 60 ต่างกับ ห.ร.ม. ของ 12 และ 60 อยู่เท่าไร
 ก. 3
 ข. 5
 ค. 7
 ง. 9
 จ. 12
21. จำนวนที่มากที่สุดที่หาร 26, 76 และ 131 แล้วเหลือเศษ 1 ทุกจำนวน คือ
 ก. 3
 ข. 5
 ค. 15
 ง. 25
 จ. 26
22. จำนวนในข้อใดเป็นตัวคูณร่วมของ 6, 7, 14
 ก. 1
 ข. 2
 ค. 14
 ง. 28
 จ. 42
23. ตัวคูณร่วมที่น้อยที่สุดของ 3,4,6 คือ
 ก. 3
 ข. 4
 ค. 6
 ง. 12
 จ. 24

24. จำนวนที่มี 60 เป็นตัวคูณร่วมน้อยที่สุดคือ
 ก. 24, 48, 72
 ข. 2, 11, 24
 ค. 3, 8, 24
 ง. 8, 12, 48
 จ. 4, 10, 12
25. จำนวนนับที่น้อยที่สุดที่หารด้วย $2^2 \times 3^4$, $2^3 \times 3^2$ และ $2^4 \times 3^4$ ได้ลงตัวคือ
 ก. $2^2 \times 3^2$
 ข. $2^2 \times 3^4$
 ค. $2^3 \times 3^4$
 ง. $2^4 \times 3^4$
 จ. $2^3 \times 3^4$
26. ค.ร.น. ของ 12, 16, 24, 36 คือจำนวนใด
 ก. 124
 ข. 144
 ค. 154
 ง. 164
 จ. 174
27. จงหาจำนวนที่น้อยที่สุดเมื่อนำ 4, 6, 8, 12 ไปหาร แล้วเหลือเศษ 3 ทุกตัว
 ก. 15 ค. 23
 ข. 19 ง. 27 จ. 31
28. ค.ร.น. ของ 30 และ 40 เป็นกี่เท่าของ ห.ร.ม. ของจำนวนทั้งสอง
 ก. 3 เท่า
 ข. 6 เท่า
 ค. 10 เท่า
 ง. 12 เท่า
 จ. 24 เท่า
29. ถ้า $9 = 3 \times 3$ และ $15 = 3 \times 5$ แล้ว ค.ร.น. และ ห.ร.ม. คู่กันจะได้จำนวนใด
 ก. 130
 ข. 132
 ค. 135
 ง. 140
 จ. 145
30. ผลคูณของ ค.ร.น. และ ห.ร.ม. ของเลข 2 จำนวน เป็น 216 ถ้าจำนวนหนึ่งเป็น 18 จงหาจำนวน อีกจำนวนหนึ่ง
 ก. 6
 ข. 8
 ค. 10
 ง. 12
 จ. 15
31. ถ้า $\frac{28}{26} = \frac{[]}{144}$ จำนวนใน [] คือข้อใด
 ก. 110 ง. 116
 ข. 112 จ. 118
 ค. 114

32. ถ้า $\frac{540}{804} = \frac{45}{[]}$ จำนวนใน []

คือข้อใด

ก. 65

ข. 67

ค. 69

ง. 71

จ. 73

33. ถ้า $35\frac{6}{11} = 35\frac{42}{[]}$ จำนวน [] คือข้อใด

ก. 71

ข. 73

ค. 75

ง. 77

จ. 79

34. เศษส่วนในข้อใดมีค่าเท่ากับ $\frac{5}{9}$

ก. $\frac{20}{45}$

ข. $\frac{30}{54}$

ค. $\frac{40}{55}$

ง. $\frac{25}{35}$

จ. $\frac{50}{65}$

35. เศษส่วนอย่างต่ำของ $\frac{144}{169}$ ตรงกับ

ข้อใด

ก. $\frac{12}{13}$

ข. $\frac{16}{17}$

ค. $\frac{72}{91}$

ง. $\frac{24}{29}$

จ. $\frac{144}{169}$

36. จะต้องใส่เครื่องหมาย < ลงใน []

ข้อใดจึงจะถูกต้อง

ก. $\frac{5}{8} [] \frac{1}{6}$

ข. $\frac{11}{12} [] \frac{7}{8}$

ค. $\frac{4}{7} [] \frac{3}{8}$

ง. $\frac{4}{15} [] \frac{2}{5}$

จ. $\frac{5}{9} [] \frac{7}{13}$

37. จะต้องใส่เครื่องหมาย > ลงใน []

ข้อใดจึงจะถูกต้อง

ก. $\frac{7}{24} [] \frac{4}{9}$

ข. $\frac{17}{21} [] \frac{6}{27}$

ค. $\frac{4}{15} [] \frac{7}{18}$

ง. $\frac{7}{28} [] \frac{5}{12}$

จ. $\frac{5}{11} [] \frac{3}{5}$

38. ข้อใดเปรียบเทียบได้ถูกต้อง

ก. $\frac{7}{8} > \frac{14}{15}$

ข. $\frac{1}{8} > \frac{5}{12}$

ค. $\frac{11}{15} < \frac{7}{10}$

ง. $\frac{9}{17} < \frac{6}{11}$

จ. $\frac{10}{23} > \frac{7}{9}$

39. จงหาค่าของ $\left(\frac{2}{5} - \frac{1}{4}\right) + \left(4\frac{1}{4} - 3\frac{1}{2}\right)$

ก. $\frac{3}{20}$

ข. $\frac{9}{10}$

ค. $\frac{12}{20}$

ง. $\frac{3}{4}$

จ. $\frac{6}{20}$

40. จงหาค่าของ $11 - \left(3\frac{1}{7} - 2\frac{3}{4}\right)$

ก. 22

ข. 24

ค. 26

ง. 28

จ. 30

41. ถ้า $P = 6$, $Q = 3$ จงหาค่าของ $\frac{\frac{P}{Q} + \frac{Q}{P}}{\frac{P}{Q} - \frac{Q}{P}}$

ก. $\frac{3}{5}$

ข. $\frac{4}{15}$

ค. $\frac{5}{3}$

ง. $\frac{11}{15}$

จ. $\frac{15}{4}$

42. จงหาค่าของ $\frac{\frac{4}{\frac{25}{3}}}{\frac{3}{10}}$

ก. $\frac{7}{10}$

ข. $\frac{8}{15}$

ค. $\frac{9}{10}$

ง. $\frac{3}{8}$

จ. $\frac{11}{15}$

43. จงหาค่าของ $\left(\frac{9}{14} \times 1\frac{2}{5}\right) + 5\frac{1}{7}$

ก. $\frac{3}{7}$

ข. $\frac{6}{7}$

ค. $\frac{19}{21}$

ง. $\frac{7}{40}$

จ. $\frac{9}{10}$

44. จงหาค่าของ $\frac{1\frac{1}{3} + 1\frac{1}{9}}{1\frac{2}{3} - \frac{4}{9}}$

ก. 1

ข. 2

ค. $\frac{1}{2}$ ง. $\frac{2}{9}$ จ. $\frac{9}{2}$

45. จงหาค่าของ $\frac{1\frac{1}{6} + \frac{7}{10}}{2\frac{1}{3} - 1\frac{2}{5}}$

ก. 1

ข. 2

ค. 3

ง. 4

จ. 5

46. จำนวนใดที่มีตัวประกอบเฉพาะ
ไม่ซ้ำกัน 4 ตัว

ก. 12

ข. 68

ค. 102

ง. 210

จ. 260

47. ตัวประกอบเฉพาะของ 135

ก. 5, 9

ข. 3, 9

ค. 3, 5

ง. 9, 27

จ. 5, 27

48. ข้อใดแยกตัวประกอบไม่ถูกต้อง

ก. $124 = 2^2 \times 31$

ข. $196 = 2^2 \times 7^2$

ค. $280 = 2^3 \times 5 \times 7$

ง. $360 = 2^3 \times 3^2 \times 5$

จ. $648 = 2^3 \times 3^2$

49. ข้อใดแยกตัวประกอบไม่ถูกต้อง

ก. $550 = 2 \times 5 \times 5 \times 11$

ข. $546 = 2 \times 3 \times 7 \times 13$

ค. $420 = 2 \times 5 \times 6 \times 7$

ง. $330 = 2 \times 3 \times 5 \times 11$

จ. $270 = 2 \times 3 \times 3 \times 3 \times 5$

50. ถ้า $a = 2$, $b = 3$, $c = 5$ แล้วข้อใด
เป็นการแยกตัวประกอบของ 450

- ก. $a \times b \times c^2$
- ข. $a \times b^2 \times c^2$
- ค. $a^2 \times b^2 \times c$
- ง. $a^2 \times b \times c$
- จ. $a^2 \times b \times c^2$

51. จำนวน 729 เขียนเป็นเลขยกกำลัง
คือข้อใด

- ก. 3^4
- ข. 3^5
- ค. 3^6
- ง. 3^7
- จ. 3^8

52. จำนวนที่มากที่สุดที่หาร 30 เหลือ เศษ 2
และหาร 127 เหลือเศษ 1 คือ

- ก. 7
- ข. 14
- ค. 16
- ง. 18
- จ. 31

53. ถ้า ห.ร.ม ของ x และ y คือ z ข้อใดกล่าว
ถูกต้อง

- ก. x และ y เป็นตัวประกอบของ z
- ข. x เป็นตัวประกอบของ z
- ค. y เป็นตัวประกอบของ z
- ง. z เป็นตัวประกอบของ x และ y
- จ. x เป็นตัวประกอบของ y

54. จำนวน 324 แยกเป็นตัวประกอบได้
ตามข้อใด

- ก. $2^2 \times 3^4$
- ข. $3^2 \times 6^2$
- ค. $2^2 \times 9^2$
- ง. $2^2 \times 3^2 \times 9$
- จ. $2^3 \times 3 \times 27$

55. จำนวนในข้อใดมีความแตกต่างระหว่าง
ค.ร.น กับ ห.ร.ม มากที่สุด

- ก. 2, 4, 6
- ข. 2, 3, 4
- ค. 1, 3, 5
- ง. 2, 3, 6
- จ. 3, 5, 6

56. ข้อใดเปรียบเทียบไม่ถูกต้อง

ก. $\frac{9}{4} = \frac{27}{12}$

ข. $\frac{13}{15} < \frac{9}{10}$

ค. $\frac{2}{3} > \frac{8}{21}$

ง. $\frac{1}{7} > \frac{3}{20}$

จ. $\frac{11}{12} > \frac{7}{10}$

57. เศษส่วนในข้อใดมีค่าเท่ากัน

ก. $\frac{2}{3} = \frac{6}{9} = \frac{8}{11}$

ข. $\frac{4}{7} = \frac{12}{21} = \frac{20}{30}$

ค. $\frac{5}{6} = \frac{10}{12} = \frac{25}{30}$

ง. $\frac{7}{8} = \frac{14}{15} = \frac{21}{24}$

จ. $\frac{1}{5} = \frac{8}{45} = \frac{11}{55}$

58. จงหาค่าของ $\left(\frac{3}{8} + 1\frac{1}{2}\right) \times 2\frac{1}{6}$

ก. $2\frac{1}{16}$

ข. $3\frac{5}{16}$

ค. $4\frac{1}{16}$

ง. $5\frac{7}{16}$

จ. $4\frac{2}{16}$

59. จงหาค่าของ $\left(\frac{3}{7} \times \frac{5}{12}\right) \div \left(2\frac{3}{5} - 2\frac{1}{7}\right)$

ก. $\frac{5}{9}$

ข. $\frac{25}{64}$

ค. $2\frac{7}{9}$

ง. $\frac{24}{35}$

จ. $\frac{25}{12}$

60. จงหาค่าของ $\frac{2}{3 + 1\frac{1}{5}}$

ก. $\frac{10}{21}$

ข. $\frac{11}{32}$

ค. $\frac{7}{8}$

ง. $\frac{9}{13}$

จ. $\frac{20}{21}$

ประวัติย่อของผู้วิจัย

ชื่อ - สกุล นางพรศรี จิมแก้ว

เกิดวันที่ 7 เดือนมิถุนายน พุทธศักราช 2498

สถานที่เกิด อำเภอเขียร์ใหญ่ จังหวัดนครศรีธรรมราช

สถานที่อยู่ปัจจุบัน บ้านเลขที่ 7/663 หมู่บ้านบัวขาว เขตมีนบุรี กรุงเทพฯ 10510

โทร 517-3502

ตำแหน่งหน้าที่การงาน หัวหน้าฝ่ายส่งเสริมและพัฒนาวิชาชีพครู

สถานที่ทำงานปัจจุบัน สำนักงานเลขาธิการคุรุสภา ถนนนครราชสีมา เขตดุสิต

กรุงเทพฯ โทร 282-6846 ต่อ 540

ประวัติการศึกษา

พ.ศ. 2516 มัธยมศึกษาตอนปลาย (วิทย์-คณิต)

โรงเรียนสตรีประเทืองวิทย์ กรุงเทพฯ

พ.ศ. 2518 ป.กศ.สูง (คณิตศาสตร์ - วิทยาศาสตร์)

วิทยาลัยครูสวนสุนันทา

พ.ศ. 2521 กศ.บ. (คณิตศาสตร์ - วัดผล)

มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

พ.ศ. 2538 กศ.ม. (การวัดผลการศึกษา)

มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร