

การเปรียบเทียบผลการเทียบคะแนนระหว่างวิธีการตอบข้อคำถามที่ใช้คะแนนจริง,
คะแนนสังเกต และวิธีอีควิเปอร์เซนไทล์

ปริญญานิพนธ์
ของ
ยุพิน หลายเจริญ

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
พฤษภาคม 2545
ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

371-2721

ข 421 ก

อ.3

การเปรียบเทียบผลการเทียบคะแนนระหว่างวิธีการตอบข้อคำถามที่ใช้คะแนนจริง,
คะแนนสังเกต และวิธีอิตีเปอร์เซนไทล์

บทคัดย่อ

ของ

ยุพิน หลายเจริญ

31 ก.ค. 2545

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

พฤษภาคม 2545

๒ 148967

ยุพิน หลายเจริญ. (2545). การเปรียบเทียบผลการเทียบคะแนนระหว่างวิธีการตอบข้อคำถามที่ใช้

คะแนนจริง, คะแนนสังเกต และวิธีอิกวิเปอร์เซนไทล์. ปรินซ์นิพนธ์ กศ.ม. (การวัดผลการ
ศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม :
รองศาสตราจารย์ อังคณา สายยศ , ดร.ชูศักดิ์ ชัมภลขิต.

การศึกษานี้มีจุดมุ่งหมายเพื่อเปรียบเทียบผลการเทียบคะแนนระหว่างวิธีการตอบข้อคำถามที่ใช้คะแนนจริง, คะแนนสังเกต, และวิธีอิกวิเปอร์เซนไทล์ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2542 ของสำนักงานการประถมศึกษาจังหวัดหนองคาย จำนวน 1,062 คน ได้มาโดยการสุ่มแบบสองขั้นตอน (Two – Stage Random Sampling) เครื่องมือที่ใช้ในการศึกษาเป็นแบบทดสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่อง สมการ ที่ผู้วิจัยสร้างขึ้นจากการประมาณค่าพารามิเตอร์ของข้อสอบ แล้วแบ่งออกเป็น 2 ฉบับ ๆ ละ 30 ข้อ

ผลการวิจัยสรุปได้ดังนี้

1. คะแนนที่ได้จากวิธีการเทียบคะแนนจริงกับวิธีเทียบคะแนนสังเกตมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนวิธีเทียบคะแนนสังเกตกับวิธีอิกวิเปอร์เซนไทล์ , วิธีเทียบคะแนนจริงกับวิธีอิกวิเปอร์เซนไทล์แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ความแตกต่างของคะแนนความคลาดเคลื่อนที่ได้จากวิธีเทียบคะแนนจริงกับวิธีอิกวิเปอร์เซนไทล์มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนวิธีเทียบคะแนนจริงกับวิธีเทียบคะแนนสังเกต, วิธีเทียบคะแนนสังเกตกับวิธีอิกวิเปอร์เซนไทล์มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. ความเพียงพอของวิธีการเทียบคะแนนจริงสูงกว่าวิธีเทียบอิกวิเปอร์เซนไทล์ และวิธีเทียบคะแนนสังเกต ตามลำดับ แต่เมื่อเปรียบเทียบกับเกณฑ์แล้ว ทุกวิธีให้ผลเป็นที่น่าพอใจอย่างยิ่ง

**A COMPARISON OF EFFECT OF EQUATING SCORES BETWEEN ITEM
RESPONSES THEORY USING TRUE-SCORE , OBSERVED-SCORE METHODS
AND EQUIPERCENTILE METHOD**

**AN ABSTRACT
BY
YUPIN LAICHAROEN**

**Present in partial fulfillment of the requirements
for the Master of Education degree in Educational Measurement
at Srinakharinwirot University
May 2002**

Yupin Laicharoen. (2545). *A comparison of effect of equating scores between Item Response Theory using True score ,Observed score methods and Equipercentile method* .Master thesis .M.Ed. (Educational Measurement).Bangkok : Graduate School ,Srinakharinwirot University.Advisor Committee : Assoc.Prof.Aungkana Saiyos , Dr.Choosak Khampalikit.

The purpose of this study were to compare the effect of equating score between Item Response Theory using True-score ,Observed-score method and Equipercentile method. The sample were 1,062 students of Pathomsuksa 6 in the academic year 1999 form primary school in Nongkhai Province and selected by Two-stage random sampling .The instrument were 2 subtests that composed Test X and Test Y of Mathematics achievement test that contains 30 multiple-choice items with 5 alternative per item .

The result of this study showed that ;

1. The equating score between Item Response Theory using True-score method and Observed score method were statistically significant different at .05 level , but the Item Response Theory using True-score method and Equipercentile method, Item Response Theory using Observed-score method and Equipercentile method not statistically significant different.
2. The standard error of equating Item Response Theory using True-score method and Equipercentile method were statistically significant different at .01 level. But the standard error of equating score of Item Response Theory using True-score method and Observed-score method, Item Response Theory using Observed-score method and Equipercentile method not statistically significant different.
3. The discrepancy of equating of Item Response Theory using True-score more higher than Equipercentile method , Item Response Theory using Observed-score method , and all procedures had satisfied of criterion discrepancy index in equating scores.

ปริญญานิพนธ์ฉบับนี้ได้รับทุนสนับสนุนการวิจัย
ทุน ดร.ชวาล แพรัตกุล
มหาวิทยาลัยศรีนครินทรวิโรฒ
ประจำปีการศึกษา 2543

ผู้วิจัยขอกราบขอบพระคุณไว้เป็นอย่างสูง ณ โอกาสนี้

ปริญญานิพนธ์

เรื่อง

การเปรียบเทียบผลการเทียบคะแนนระหว่างวิธีการตอบข้อคำถามที่ใช้คะแนนจริง,
คะแนนสังเกต และวิธีอิกวิเปอร์เซนไทล์

ของ

นางสาวยุพิน หลายเจริญ

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....โพนวิทย์ น.น.น......คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร.นภาพร หะวานนท์)
วันที่ 10 เดือน พฤษภาคม พ.ศ. 2545

คณะกรรมการสอบปริญญานิพนธ์

.....อ.อ.อ.อ.อ......ประธาน
(รองศาสตราจารย์อังกณา สายยศ)

.....อ.อ.อ.อ.อ......กรรมการ
(อาจารย์ ดร. ชูศักดิ์ ชัมภลขิต)

.....อ.อ.อ.อ.อ......กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ชูศรี วงศ์รัตน)

.....อ.อ.อ.อ.อ......กรรมการที่แต่งตั้งเพิ่มเติม
(ผู้ช่วยศาสตราจารย์ ดร.นิคม ดังคะพิภพ)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดีด้วยความช่วยเหลือและแก้ไขข้อบกพร่องต่างๆ อย่างดียิ่งจาก รองศาสตราจารย์อังคณา สายยศ และอาจารย์ ดร.ชูศักดิ์ ชัมภลลิขิต ผู้วิจัยขอกราบขอบพระคุณไว้เป็นอย่างสูง ณ โอกาสนี้

ขอกราบขอบพระคุณ รองศาสตราจารย์ชูศรี วงศ์รัตนะ ผู้ช่วยศาสตราจารย์ ดร. นิคม ตั้งคะพิภพ รองศาสตราจารย์ นิภา ศรีไพโรจน์ ผู้ช่วยศาสตราจารย์ระวีวรรณ พันธุ์พานิช อาจารย์ชวลิต รวยอาจิน อาจารย์ทรงสิทธิ์ เจริญดี และผู้อำนวยการประยูร นามวงศ์ ที่ให้ความช่วยเหลือในแนะนำและแก้ไขข้อบกพร่องต่างๆ ของแบบทดสอบ และปริญญานิพนธ์

ขอขอบคุณผู้บริหารโรงเรียนและคณาจารย์ ขอบใจนักเรียนทุกคนที่ให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี

ขอขอบคุณ คุณอนุพันธ์ หลายเจริญ ที่ให้การสนับสนุนด้านการเงินในการเรียนและการทำปริญญานิพนธ์ และขอบคุณครอบครัว "หลายเจริญ" และครอบครัว "ศรีสุนทร" ที่เป็นกำลังใจแก่ผู้วิจัยเสมอมา

ขอขอบคุณสมาชิกครอบครัว ร้อยเอกหญิงวรรณรัตน์ ใจเชื้อกุล ที่เอื้อเฟื้อสถานที่และเครื่องคอมพิวเตอร์ ในการวิเคราะห์ข้อมูลและการจัดพิมพ์ปริญญานิพนธ์

ปริญญานิพนธ์เล่มนี้ ได้รับทุนสนับสนุนการวิจัยจาก ทุน ดร.ชวาล แพรัตกุล ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

คุณประโยชน์ของปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่ บิดา มารดา และผู้มีพระคุณทุกท่าน ที่ได้ให้คำแนะนำ จนทำให้ปริญญานิพนธ์ฉบับนี้สมบูรณ์ทุกประการ

ยุพิน หลายเจริญ

สารบัญ

บทที่		หน้า
1	บทนำ.....	1
	ภูมิหลัง.....	1
	ความมุ่งหมายของการศึกษาค้นคว้า.....	3
	ความสำคัญของการศึกษาค้นคว้า.....	3
	ขอบเขตของการศึกษาค้นคว้า.....	4
	เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	4
	ถำนิยาศัพท์เฉพาะ.....	5
2	เอกสารและงานวิจัยที่เกี่ยวข้อง.....	8
	แนวคิดเกี่ยวกับการเทียบคะแนน.....	8
	ทฤษฎีที่เกี่ยวข้องกับการเทียบคะแนน.....	12
	งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนน.....	24
	สมมติฐานของการศึกษาค้นคว้า.....	31
3	วิธีดำเนินการศึกษาค้นคว้า.....	33
	ประชากร.....	33
	กลุ่มตัวอย่าง.....	33
	เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	35
	วิธีดำเนินการสร้างแบบทดสอบ.....	36
	วิธีดำเนินการเก็บรวบรวมข้อมูล.....	40
	การวิเคราะห์ข้อมูลและสถิติที่ใช้.....	40
4	ผลการวิเคราะห์ข้อมูล.....	47
	สัญลักษณ์และอักษรที่ใช้ในการวิเคราะห์ข้อมูล.....	47
	การเสนอผลการวิเคราะห์ข้อมูล.....	47
	ผลการวิจัย.....	47
5	สรุปผล อภิปรายผล และข้อเสนอแนะ.....	62
	ความมุ่งหมายของการศึกษาค้นคว้า.....	62
	กลุ่มตัวอย่าง.....	62
	เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	62

สารบัญ (ต่อ)

บทที่	หน้า
5 (ต่อ) วิธีดำเนินการเก็บรวบรวมข้อมูล.....	62
สรุปผลการวิจัย.....	63
อภิปรายผล.....	64
ข้อเสนอแนะ.....	65
 บรรณานุกรม.....	 66
 ภาคผนวก.....	 70
 ประวัติย่อผู้วิจัย.....	 113

บัญชีตาราง

ตาราง	หน้า
1 รายชื่อโรงเรียนและจำนวนนักเรียนในกลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า.....	33
2 คำนวณสถิติพื้นฐานของแบบทดสอบวิชาคณิตศาสตร์ ชั้น ป.6 เรื่องสมการ.....	48
3 คะแนนสมมูลของการเทียบคะแนนจากแบบทดสอบ ฉบับ X ไปสู่ฉบับ Y ที่ระดับคะแนน และความสามารถเดียวกันตามวิธีเทียบคะแนนจริง.....	51
4 คะแนนสมมูลของการเทียบคะแนนจากแบบทดสอบฉบับ X ไปสู่ฉบับ Y ที่ระดับคะแนน และความสามารถเดียวกันตามวิธีเทียบคะแนนสังเกต.....	53
5 คะแนนสมมูลของการเทียบคะแนนแบบทดสอบ ฉบับ X และแบบทดสอบฉบับ Y ที่ระดับ เปอร์เซ็นต์ไทล์เดียวกัน.....	54
6 คะแนนสมมูลของการเทียบคะแนนทั้งสามวิธี.....	55
7 การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากวิธีเทียบทั้งสามวิธี.....	56
8 ความคลาดเคลื่อนที่ได้จากวิธีเทียบทั้งสามวิธี.....	57
9 การเปรียบเทียบความคลาดเคลื่อนที่ได้จากวิธีเทียบคะแนนทั้งสามวิธี.....	59
10 ความเพียงพอของการเทียบคะแนน.....	60
11 ตารางวิเคราะห์รายละเอียด เนื้อหาเรื่อง สมการและการแก้สมการ.....	72
12 ค่าไอเกนขององค์ประกอบที่มีค่า เกิน 1.0 ของแบบทดสอบวิชาคณิตศาสตร์ ชั้นป.6.....	84
13 ผลการวิเคราะห์ค่าพารามิเตอร์จากการทดลองเครื่องมือ.....	86
14 ค่าพารามิเตอร์ของแบบทดสอบทั้งสองฉบับ.....	90
15 การคำนวณคะแนนที่ได้จากการเทียบคะแนนจริง.....	95
16 การคำนวณคะแนนที่ได้จากการเทียบคะแนนสังเกต.....	99
17 การคำนวณคะแนนที่ได้จากการเทียบคะแนนโดยใช้ตำแหน่งเปอร์เซ็นต์ไทล์ที่เท่ากัน.....	101
18 ตารางการเทียบคะแนนจากฉบับ X และฉบับ Y ตามตำแหน่งเปอร์เซ็นต์ไทล์.....	102
18 – 1 ตารางเทียบคะแนนจากฉบับ X และ ฉบับ Y	103
19 การเทียบคะแนนตามตำแหน่งเปอร์เซ็นต์ไทล์.....	105
20 การคำนวณคะแนนที่ได้จากการเทียบคะแนนทั้งสามวิธี.....	107
21 ค่าดัชนีความสอดคล้อง (IOC).....	109

บัญชีภาพประกอบ

ภาพประกอบ

หน้า

1	แบบแผนการเก็บรวบรวมข้อมูลเพื่อเทียบคะแนนแบบทดสอบ.....	11
2	การเทียบคะแนนตามทฤษฎีการตอบข้อสอบ ระหว่างแบบทดสอบ ฉบับ X และฉบับ Y	18
3	แผนผังการสร้างแบบทดสอบวิชาคณิตศาสตร์ ชั้น ป. 6	36
4	คะแนนจริงของผู้สอบแบบทดสอบทั้งสองฉบับที่ระดับคะแนนและความสามารถต่างๆ.....	50
5	คะแนนที่คาดหวังของผู้สอบแบบทดสอบทั้งสองฉบับที่คะแนนโอกาสในการตอบถูก.....	52
6	ความคลาดเคลื่อนของการเทียบคะแนนทั้งสองวิธีที่ระดับคะแนนและความสามารถเดียวกัน.....	58
7	โค้งลักษณะเฉพาะของแบบทดสอบทั้งฉบับ.....	92
8	โค้งอินฟอร์เมชันของแบบทดสอบที่ระดับความสามารถต่างๆ.....	93
9	โค้งความถี่สะสม.....	104

บทที่ 1

บทนำ

ภูมิหลัง

การวัดผลและการประเมินผลการศึกษานับเป็นเครื่องมือสำคัญที่ช่วยปรับปรุงคุณภาพของการศึกษา เพราะผลจากการวัดผลและประเมินผลการศึกษาจะเป็นพื้นฐานสำหรับการตัดสินใจของครู และ นักการศึกษา เพื่อใช้ในการปรับปรุงวิธีสอน การใช้สื่อการสอน แบบเรียน หลักสูตร การแนะแนว การจัดระบบการบริหารโรงเรียน ตลอดจนการปรับปรุงวิธีเรียนของผู้เรียนให้มีประสิทธิภาพยิ่งขึ้น ซึ่งจะส่งผลในการดำเนินการจัดการศึกษาให้เป็นไปตามแนวทางและบรรลุผลตามจุดมุ่งหมายที่หลักสูตรต้องการ (สำราญ มีแจ้ง .2534: 1) การประเมินผลการจัดการศึกษาที่ผ่านมาวิธีการหนึ่งก็คือ การตรวจสอบคุณภาพการศึกษา โดยสอบวัดผลสัมฤทธิ์ทางการเรียนและข้อมูลที่ได้จากการสอบจะช่วยให้ ครูผู้สอนทราบถึงผลการสอนในช่วงระยะเวลาหนึ่งมาแล้ว ซึ่งจะช่วยให้สามารถวินิจฉัยความสามารถทางการเรียนของผู้เรียนและสามารถจัดการเรียนการสอนให้กับผู้เรียนแตกต่างกันไปตามปัญหาที่พบได้อย่างเหมาะสม หากการ ติดตามผลการจัดการศึกษาดังกล่าวได้กระทำอย่างต่อเนื่องโดยทำการสอบวัดกับนักเรียนในกลุ่มอายุ และระดับชั้นต่างกัน ตลอดจนการวัดความเปลี่ยนแปลงในทางอ้อมหรือเสื่อมโทรมของผลการศึกษาในช่วงเวลาต่างๆ กันแล้ว จะทำให้ได้ข้อมูลที่มีประโยชน์ต่อการกำหนดนโยบายและแนวปฏิบัติในการจัดการศึกษาให้มีความเหมาะสมยิ่งขึ้น (สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ. 2532 : 133 - 134)

การทดสอบก็เป็นสิ่งจำเป็นในวงการศึกษามากกว่าจะเป็นด้านการสอบเพื่อเข้าศึกษาต่อในระดับมหาวิทยาลัย การสอบคัดเลือกเข้าทำงาน การสอบเลื่อนตำแหน่ง การสอบเลื่อนระดับของหน่วยงานต่างๆ ซึ่งมีการสอบทุกปี อาจมีการใช้แบบทดสอบที่มีลักษณะแตกต่างกันไป เพื่อป้องกันการรั่วไหลของข้อสอบ หรือการขวนขวายหาคำตอบจากแบบทดสอบจากปีก่อนๆ ของผู้สอบ ทำให้เกิดการได้เปรียบเสียเปรียบกัน เพราะฉะนั้นในแต่ละหน่วยงานจึงต้องมีการสร้างแบบทดสอบที่มีลักษณะของโครงสร้างและเนื้อหาตลอดจนค่าสถิติต่างๆ ให้อยู่ในระดับเดียวกันกับชุดก่อนๆ เพื่อให้แน่ใจว่า วัดในสิ่งที่ต้องการวัดและมีความเชื่อถือได้ ตลอดจนสามารถเปรียบเทียบผลสอบกับการสอบครั้งก่อนๆ ได้ จึงนำแนวคิดของการเทียบคะแนนเข้ามาช่วยในการเทียบผลการสอบในระดับต่างๆ

แม้กระทั่งการประเมินคุณภาพการเรียนของนักเรียนนั้นก็มีหน่วยงานที่เกี่ยวข้องดำเนินการหลายหน่วยงาน ได้แก่ กรมวิชาการ หน่วยศึกษานิเทศก์ กรมสามัญศึกษา กองวิชาการ สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ นอกจากนี้ยังมีการประเมินผลสัมฤทธิ์ทางการเรียนของนักเรียนในรูปแบบงานวิจัยของอาจารย์ นิสิต นักศึกษาในสถาบันการศึกษาต่างๆ ซึ่งส่วนใหญ่จะนำผลสัมฤทธิ์ทางการเรียนของปีการศึกษาที่ต่างกัน กลุ่มตัวอย่างที่ต่างกัน และใช้เครื่องมือที่ต่างกันมาเปรียบเทียบกันโดยตรง โดยทั่วไปคะแนนที่ได้จากเครื่องมือวัดที่ต่างกันมิได้แทนความสามารถเดียวกัน ดังนั้นการนำคะแนนที่ได้ไปเปรียบเทียบกันโดยตรง จึงทำให้ผลการเปรียบเทียบขาดความแม่นยำและเกิดความคลาดเคลื่อนสูง เช่นเดียวกับการประเมินผลในระดับโรงเรียน ระดับกลุ่มโรงเรียน ระดับอำเภอ และระดับจังหวัด เครื่องมือที่ใช้วัดส่วนใหญ่เป็นแบบทดสอบที่ครูสร้างขึ้นเอง ถึงแม้จะมีคุณภาพในระดับหนึ่ง แต่ก็ไม่สามารถจะนำมาเปรียบเทียบกันโดยตรงได้ (ล้วน สายยศ และอังคณา สายยศ.2527:1) การแก้ปัญหาดังกล่าว คือ มีการรวมกลุ่มกันสร้างแบบทดสอบ โดยทำเป็นกลุ่มโรงเรียน หรือขยายกลุ่มให้ใหญ่ขึ้นและเพื่อไม่ให้เกิดปัญหาข้อสอบรั่ว หรือเกิด การได้เปรียบเสียเปรียบกันในระหว่างผู้เข้าสอบ ในกรณีที่มีการสอบไม่พร้อมกัน

ดังนั้น ความต้องการประการหนึ่งของผู้ที่เกี่ยวข้องกับวงการทดสอบคือ การที่จะมีแบบทดสอบที่สามารถวัดคุณลักษณะเดียวกันหลายๆฉบับ โดยที่สามารถนำผลสอบแต่ละฉบับนั้นมาเปรียบเทียบและแปลความหมายได้ ทั้งนี้เพื่อประโยชน์ในการดำเนินการสอบให้ยุติธรรมเกิดความเหมาะสมกับสภาพการสอบในแต่ละสถานการณ์ ความต้องการนี้นำไปสู่แนวความคิดในการสร้างแบบทดสอบที่เรียกว่า "แบบทดสอบคู่ขนาน" อย่างไรก็ตามความพยายามในการสร้างแบบทดสอบคู่ขนานมักจะประสบปัญหาไม่สามารถสร้างแบบทดสอบให้คู่ขนานอย่างแท้จริง คือ ไม่สามารถนำเอาคะแนนดิบที่ได้จากแบบคู่ขนานนั้นมาเปรียบเทียบกันได้โดยตรง เพราะในทางปฏิบัตินั้นสามารถทำได้เพียงแต่ความใกล้เคียงด้านเนื้อหา ค่าสถิติของข้อสอบ และแบบทดสอบทั้งฉบับเท่านั้น (Gulliksen.1950: 174) จึงมีผู้คิดค้นวิธีการที่จะทำให้คะแนนจากแบบทดสอบที่วัดคุณลักษณะเดียวกัน แต่ต่างฉบับกัน สามารถนำผลมาเปรียบเทียบกันได้โดยมีความหมายขึ้น วิธีการดังกล่าว คือ การเทียบมาตรา (Equating) ซึ่ง ชูศักดิ์ ชัมภลขิต (2527 : 144) ได้สรุปว่า การเทียบมาตราเป็นกระบวนการที่เกี่ยวข้องกับกิจกรรม 2 ประการ คือ

- 1) กระบวนการที่ทำให้แบบทดสอบ 2 ฉบับใดๆ มีความทัดเทียมกันในเชิงของโครงสร้าง และ
- 2) การใช้วิธีการทางสถิติเพื่อปรับ (Adjust) คะแนนที่ได้จากแบบทดสอบแต่ละฉบับให้อยู่ในมาตราเดียวกันและเทียบกันได้ และข้อตกลงเบื้องต้นของการเทียบมาตราซึ่งเสนอโดย เลวิน (Holland and Rubin. 1982 : 10; citing Levin.n.d) ก็คือ แบบทดสอบ 2 ฉบับใดๆ จะเทียบกันได้ ก็ต่อเมื่อแบบทดสอบ 2 ฉบับนั้นคู่ขนานกันในด้านต่อไปนี้

1. โครงสร้าง (Structure)
2. เวลา (Timing)
3. ชนิดของข้อสอบ (Item Type)
4. รูปแบบ (Format)
5. เนื้อหาวิชา (Subject Matter)

วิธีการเทียบคะแนนที่นิยมใช้โดยทั่วไปมี 3 วิธี คือ

1. การเทียบคะแนนเชิงเส้นตรง (Linear Equating) วิธีนี้อาศัยการแปลงคะแนนผ่านคะแนนมาตรฐาน (Z-score) ซึ่งมีข้อตกลงเบื้องต้นว่า การกระจายของคะแนนทั้งสองกลุ่มต้องเหมือนกัน
2. การเทียบคะแนนโดยใช้เปอร์เซ็นต์ไทล์ที่เท่ากัน (Equipercentile Equating) เป็นการแปลงคะแนนเชิงเส้นโค้ง (Curvilinear Transformation) ซึ่งแองกอฟฟ์ (Anggoff.1971:563) ได้ให้นิยามเกี่ยวกับการเทียบคะแนนโดยใช้เปอร์เซ็นต์ไทล์ว่า "คะแนนสองตัวใดๆจะสามารถเทียบเท่ากันได้ก็ต่อเมื่อ คะแนนสองตัวนั้นต่างก็อยู่ในตำแหน่งเปอร์เซ็นต์ไทล์ที่เท่ากันของกลุ่มที่กำหนดให้"
3. การเทียบคะแนนโดยใช้โค้งลักษณะข้อสอบ (Item Characteristic Curve Equating) วิธีนี้เป็นการเทียบมาตราโดยใช้คะแนนจริง (True-score) ที่ระดับค่าประมาณความสามารถแท้ (Latent Trait) เดียวกันระหว่างแบบทดสอบต่างชุดนั้น (Lord. 1980:199-201)

สำหรับการศึกษาในเชิงประเมินว่าวิธีการเทียบคะแนนวิธีใดจะให้ผลการเทียบคะแนนที่ถูกต้องหรือมีประสิทธิภาพมากกว่ากันนั้น อาจกระทำได้ 2 แบบ คือ ประเมินก่อนการเทียบคะแนน โดยพิจารณาจากการแจกแจงคะแนนที่ได้จากการสอบด้วยแบบทดสอบต่างชุดว่าเป็นไปตามนิยามของการกำหนดคะแนนสมมูลของวิธีใดยอมแสดงว่าวิธีการนั้นเป็นวิธีการที่เหมาะสมมากกว่าอีกแบบหนึ่งก็คือ การประเมินหลังการเทียบคะแนน โดยเปรียบเทียบคะแนนที่ได้กับคะแนนเกณฑ์(คะแนนเกณฑ์ที่ใช้อาจได้มาจากกลุ่มตัวอย่าง

อื่นที่เท่าเทียมกัน หรือได้มาจากวิธีการอื่นที่เชื่อถือได้) หากพบว่าวิธีการเทียบคะแนนวิธีใดให้ผลการเทียบคะแนนที่ใกล้เคียงคะแนนเกณฑ์มากกว่า แสดงว่าวิธีนั้นมีประสิทธิภาพ ซึ่งแนวคิดเชิงทฤษฎีและนิยามของการเทียบคะแนนพอสรุปได้ว่า การเทียบคะแนนเชิงเส้นตรงมีความเหมาะสมในการแปลงคะแนนอย่างมาก เมื่อการแจกแจงคะแนนของแบบทดสอบต่างชุดที่นำมาเทียบคะแนนนั้นมีรูปร่างเหมือนกัน (Angoff, 1971; Braun and Holland, 1982:23) แต่ถ้าวรูปร่างของการแจกแจงคะแนนจากแบบทดสอบต่างชุดกันนั้นแตกต่างกันแล้วการเทียบคะแนนแบบอิกวิเปอร์เซนไทล์จะให้ผลดีกว่าการเทียบคะแนนเชิงเส้นตรง (Angoff, 1971:564) ส่วนการเทียบคะแนนโดยอาศัยทฤษฎีการตอบข้อคำถามนั้น ลอร์ด (Lord, 1980:203) ได้ศึกษาทฤษฎีการเทียบคะแนนดิบของแบบทดสอบสองชุด ด้วยวิธีการเทียบคะแนนจริงและคะแนนสังเกต พบว่ามีข้อได้เปรียบและเสียเปรียบ วิธีการเทียบคะแนนจริงให้ความหมายของคะแนนสมมูลที่อยู่เหนือค่าเฉลี่ยการเดา ส่วนวิธีการประมาณค่าคะแนนสังเกต ก็มีจุดอ่อนที่เป็นเพียงการเทียบมาตรงอย่างประมาณเท่านั้น ต่อข้อคำถามที่ว่า วิธีใดจะให้ผลการเทียบคะแนนที่ดีกว่านั้น ลอร์ด กล่าวว่า ทั้งสองวิธีมีความสอดคล้องกันมาก แต่การสรุปอ้างอิงต้องทำอย่างพิถีพิถัน ส่วน ฮาน, โคเลนและโพลแมนน์ (Tianqi Han, Michael Kolen and John Pohlmann, 1997 : 105 - 121) ได้ศึกษาผลการเทียบคะแนนระหว่างวิธีการเทียบคะแนนตามทฤษฎีการตอบข้อคำถามที่ใช้คะแนนจริง, คะแนนสังเกต และวิธีอิกวิเปอร์เซนไทล์ พบว่า โดยเฉลี่ยแล้ว การเทียบคะแนนจริงให้ผลที่คงเส้นคงวามากกว่าอีกสองวิธี แต่เสนอแนะให้มีการตรวจสอบแนวโน้มเหล่านี้

ด้วยเหตุนี้ผู้วิจัยจึงมีความสนใจที่จะศึกษาเปรียบเทียบผลของการเทียบคะแนนระหว่างวิธีที่ใช้คะแนนจริง คะแนนสังเกตตามทฤษฎีการตอบข้อคำถามและวิธีอิกวิเปอร์เซนไทล์ว่าวิธีใดจะให้การเทียบที่มีความเหมาะสมมากกว่ากัน

ความมุ่งหมายของการศึกษาค้นคว้า

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยมีความมุ่งหมายเพื่อศึกษาเปรียบเทียบผลการเทียบคะแนนแบบทดสอบวิชาคณิตศาสตร์ จำนวน 2 ฉบับด้วยวิธีการเทียบคะแนนจริง และการเทียบคะแนนสังเกตตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์ โดยมีจุดมุ่งหมายเฉพาะดังนี้

1. เพื่อเปรียบเทียบคะแนนที่เทียบได้จากแบบทดสอบวิชาคณิตศาสตร์ จากวิธีการเทียบคะแนนจริง คะแนนสังเกต ตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์
2. เพื่อเปรียบเทียบความคลาดเคลื่อนของการเทียบคะแนนจริง คะแนนสังเกตตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์
3. เพื่อเปรียบเทียบความเพี้ยนของการเทียบคะแนนจริง คะแนนสังเกตตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์

ความสำคัญของการศึกษาค้นคว้า

ผลจากการศึกษาค้นคว้าครั้งนี้ ทำให้ทราบว่า วิธีการเทียบคะแนนโดยใช้คะแนนจริง คะแนนสังเกตตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์ วิธีการใดทั้ง 3 วิธี จะเป็นวิธีที่ให้ผลในการเทียบคะแนนเหมาะสมที่สุด อันจะเป็นประโยชน์ในการตัดสินใจที่จะเลือกวิธีการเทียบคะแนนไปใช้ในการเทียบคะแนนการสอบได้อย่างมีความหมายและมีประสิทธิภาพในการสอบแต่ละครั้งต่อไป

ขอบเขตของการศึกษาค้นคว้า

1. ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2542 สังกัดสำนักงานการประถมศึกษาจังหวัดหนองคาย จำนวน 485 โรงเรียน จำนวน 12,707 คน

2. กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 254 สังกัด สำนักงานการประถมศึกษาจังหวัดหนองคาย โดยวิธีการสุ่มสองขั้นตอน (Two – stage Random Sampling) คือสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีขนาดโรงเรียนเป็นชั้น (Strata) มีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) ได้กลุ่มตัวอย่างจำนวน 1,062 คน จาก 47 โรงเรียน

3. ตัวแปรที่ใช้ในการศึกษา

3.1 ตัวแปรอิสระ ได้แก่ วิธีการเทียบคะแนน 3 วิธี

- 3.1.1 การเทียบคะแนนตามทฤษฎีการตอบข้อคำถาม แบบเทียบจากคะแนนจริง
- 3.1.2 การเทียบคะแนนตามทฤษฎีการตอบข้อคำถาม แบบเทียบจากคะแนนการสังเกต
- 3.1.3 การเทียบคะแนนตามวิธีอควิเปอร์เซนไทล์

3.2 ตัวแปรตาม คือ ผลการเทียบคะแนนซึ่งพิจารณาจาก

- 3.2.1 คะแนนที่เทียบได้ในแต่ละวิธี
- 3.2.2 ความคลาดเคลื่อนของการเทียบคะแนน
- 3.2.3 ความเพียงพอของการเทียบคะแนน

4. เนื้อหา ที่ใช้ในการออกข้อสอบ คือ เนื้อหาวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่อง สมการ ตามหลักสูตรประถมศึกษา 2521 (ฉบับปรับปรุง พ.ศ. 2533) เรื่องเศษส่วน ซึ่งมีเนื้อหาย่อยดังนี้

- 4.1 สมการ
- 4.2 สมการที่เป็น
- 4.3 สมการที่มีตัวไม่ทราบค่า
- 4.4 คำตอบของสมการ
- 4.5 สมบัติการเท่ากันเกี่ยวกับการบวกและลบ
- 4.6 การแก้สมการโดยใช้การเท่ากันเกี่ยวกับการบวกและลบ
- 4.7 สมบัติการเท่ากันเกี่ยวกับการคูณและหาร
- 4.8 การแก้สมการโดยใช้สมบัติการเท่ากันเกี่ยวกับการคูณและหาร
- 4.9 การแก้โจทย์ปัญหาด้วยสมการ
- 4.10 การเขียนสมการจากข้อความที่กำหนดให้

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นแบบทดสอบแบบเลือกตอบ ชนิด 5 ตัวเลือก วิชา คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่อง สมการ ตามหลักสูตรประถมศึกษา 2521 (ฉบับปรับปรุง พ.ศ. 2533) จำนวน 2 ฉบับๆละ 30 ข้อ

คำนิยามศัพท์เฉพาะ

1. การเทียบคะแนน หมายถึง กระบวนการที่ใช้วิธีการทางสถิติเพื่อแปลงคะแนนที่ได้จากแบบทดสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 จำนวน 2 ฉบับ (ฉบับ X และฉบับ Y) ให้มีมาตราเดียวกัน และนำมาเปรียบเทียบกันได้

2. การเทียบคะแนนโดยใช้ทฤษฎีการตอบข้อคำถาม หมายถึง การเทียบมาตราในเชิงโครงสร้างที่ใช้คะแนนความสามารถของผู้สอบตามแบบจำลองโลจิสติกสามพารามิเตอร์ และคำนวณความสามารถที่แท้จริงจากโด่งลักษณะเฉพาะของข้อคำถามซึ่งมีวิธีการเทียบคะแนน 2 วิธี คือ

2.1 การเทียบจากคะแนนจริง หมายถึง การเทียบคะแนนโดยใช้คะแนนของแบบทดสอบ 2 ฉบับ (ฉบับ X และฉบับ Y) ที่ใช้ความสัมพันธ์เชิงคณิตศาสตร์ระหว่างความสามารถ θ และคำนวณคะแนนตามข้อที่ตอบถูกจากสูตร

$$\xi_x = \sum_{i=1}^n P_i(\theta_x)$$

$$\xi_y = \sum_{j=1}^m P_j(\theta_y)$$

2.2 การเทียบจากคะแนนการสังเกต หมายถึง การเทียบคะแนนโดยใช้คะแนนที่สอบได้มาคำนวณหาค่าคะแนนสังเกตจากแบบทดสอบสองฉบับ (ฉบับ X และฉบับ Y) ในแต่ละความสามารถ θ แล้วนำมาเทียบคะแนนโดยใช้สูตร

$$\sum_{x=0}^n f(x/\theta) t^x = \prod_{i=1}^n [Q_i(\theta) + tP_i(\theta)]$$

$$\sum_{y=0}^n f(y/\theta) t^y = \prod_{i=1}^n [Q_i(\theta) + tP_i(\theta)]$$

3. การเทียบคะแนนตามวิธีอีควิเปอร์เซนไทล์ หมายถึง วิธีการทางสถิติที่ใช้ปรับคะแนนของแบบทดสอบฟอร์มหนึ่งให้อยู่ในมาตราเดียวกับคะแนนของแบบทดสอบอีกฟอร์มหนึ่งโดยเทียบที่ตำแหน่งเปอร์เซนไทล์เดียวกันของทั้งสองกลุ่ม มีลักษณะเป็นการแปลงคะแนนเชิงเส้นโค้ง ซึ่งถือว่าคะแนนที่ได้จากแบบทดสอบต่างฉบับกันจะเทียบเท่ากันได้ถ้าตรงกับคะแนน ณ ตำแหน่งเดียวกัน

4. ผลการเทียบคะแนน หมายถึง ค่าที่พิจารณาได้จากวิธีเทียบคะแนนดังนี้

4.1 คะแนนที่เทียบได้ในแต่ละวิธี หมายถึง คะแนนที่เทียบได้จากการเทียบมาตราในวิธีการเทียบคะแนนวิธี ที่ 2.1, 2.2 และ 3

4.2 ความคลาดเคลื่อนของการเทียบคะแนน หมายถึง ความคลาดเคลื่อนที่เกิดจากการนำคะแนนสอบจากแบบทดสอบฉบับ X ไปแปลงคะแนนเป็นคะแนนที่เทียบได้แล้วมีค่าเบี่ยงเบนออกจากคะแนนเดิมซึ่งหาได้จากสูตรความคลาดเคลื่อนที่คิดเป็นรากที่สองของค่าเฉลี่ยของความคลาดเคลื่อน ซึ่งตรวจสอบได้จากสูตรดังต่อไปนี้ (Macro. 1977 : 157)

$$SEE = \left[\sum_{i=1}^n (Y_{ij} - Y_{io})^2 / N \right]^{1/2}$$

เมื่อ SEE คือ Standard Error of Equating

Y_{ij} คือ คะแนน y (คะแนนจากการสอบ) ของผู้สอบคนที่ i จากวิธีการเทียบคะแนนที่ j ของแบบทดสอบ X

Y_{io} คือ คะแนน y (คะแนนที่เทียบได้) ของผู้สอบคนที่ i จากวิธีการเทียบคะแนนที่ j ที่ได้จากการแปลงคะแนนที่เทียบเท่ากัน

N คือจำนวนผู้สอบที่เป็นกลุ่มเทียบคะแนน

4.3 ความเพียงพอของการเทียบคะแนน หมายถึง ค่าที่แสดงถึงคุณภาพของการเทียบคะแนน โดยยึดหลักการประเมินของปีเตอร์เซนและคนอื่นๆ (Petersen and others. 1982 : 93-94) ซึ่งมีสูตร ดังนี้ คือ

$$c = \sum (x_i - x_i^*)^2 / nS_x^2$$

เมื่อ c คือ ค่าความเพียงพอของการเทียบคะแนน

x_i คือ คะแนน x ของผู้สอบคนที่ i จากแบบทดสอบชุด X

x_i^* คือ คะแนน x ของผู้สอบคนที่ i จากการแปลงคะแนน y

n คือ จำนวนคนในกลุ่มตัวอย่างสอบทานผล

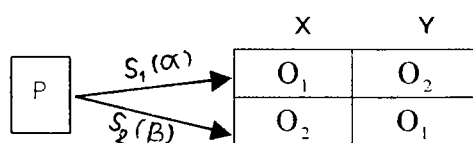
S_x^2 คือค่าความแปรปรวนของคะแนนจากแบบทดสอบ X

และมีเกณฑ์ในการพิจารณาความเพียงพอ คือ

ระดับน่าพอใจอย่างมาก	หมายถึง	$c \leq (.05SD_x)^2$
ระดับน่าพอใจ	หมายถึง	$(.05SD_x)^2 < c \leq (.10SD_x)^2$
ระดับปานกลาง	หมายถึง	$(.10SD_x)^2 < c \leq (.15SD_x)^2$
ระดับไม่น่าพอใจ	หมายถึง	$(.15SD_x)^2 < c \leq (.20SD_x)^2$
ระดับไม่น่าพอใจอย่างยิ่ง	หมายถึง	$c > (.20SD_x)^2$

5. แบบทดสอบเทียบคะแนนหมายถึง หมายถึง แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ซึ่งผู้วิจัยสร้างขึ้นและวิเคราะห์หาคุณภาพโดยใช้ทฤษฎีการตอบข้อคำถาม แล้วแบ่งออกเป็น 2 ฉบับ (ฉบับ X ฉบับ Y) โดยพิจารณาค่าโค้งลักษณะข้อสอบ (Item Characteristic Curve) ที่มีค่าความยาก (b) เท่ากันหรือเทียบเคียงกันเป็นคู่ๆ

6. กลุ่มผู้สอบร่วม (Common Person design) หมายถึง การเทียบมาตราที่ผู้สอบทั้งสองกลุ่มสอบแบบทดสอบทั้งสองฟอร์ม โดยที่ใช้แบบแผนการเทียบมาตราแบบผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (Counterbalanced Random Group Design) คือ กลุ่มแรกจะทำแบบทดสอบฉบับ X ก่อนแล้วตามด้วยฉบับ Y ส่วนกลุ่มที่สองจะทำแบบทดสอบฉบับ Y ก่อนแล้วตามด้วยฉบับ X ดังแผนภาพ



กลุ่มตัวอย่าง $S_1(\alpha)$ สอบแบบทดสอบฟอร์ม X แล้วตามด้วยฟอร์ม Y

กลุ่มตัวอย่าง $S_2(\beta)$ สอบแบบทดสอบฟอร์ม Y แล้วตามด้วยฟอร์ม X
(O_1 หมายถึงการสอบครั้งที่ 1, O_2 หมายถึงการสอบครั้งที่ 2)

7. กลุ่มเทียบคะแนน หมายถึง กลุ่มตัวอย่างที่เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ที่สอบแบบทดสอบเทียบคะแนน เพื่อนำผลการสอบไปใช้ในการแปลงคะแนนสำหรับสร้างตารางเทียบคะแนน

8. กลุ่มทานผล หมายถึง กลุ่มตัวอย่างที่เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ที่สอบแบบทดสอบเทียบคะแนน เพื่อนำผลการสอบไปใช้ในการตรวจสอบความคลาดเคลื่อนของการเทียบคะแนนในแต่ละวิธี

9. คะแนนสมมูล หมายถึง คะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 จากแบบทดสอบ ฉบับ X ฉบับ Y ซึ่งมีคะแนนเท่าหรือใกล้เคียงกันจากการเทียบคะแนนในแต่ละวิธี

10. ตารางเทียบคะแนน หมายถึง ตารางเปรียบเทียบคะแนนสมมูลจากแบบทดสอบ ฉบับ X ฉบับ Y จากวิธีการเทียบคะแนนในแต่ละวิธี

11. ความคงเส้นคงวาของการเทียบคะแนน หมายถึง ความเชื่อมั่นของแบบทดสอบที่สร้างขึ้นโดยยึดข้อตกลงเบื้องต้นของการเทียบคะแนน ซึ่งเสนอโดย เลวิน (Holland and Rubin, 1982 : 10; citing Levin) ก็คือ แบบทดสอบ 2 ฉบับใดๆ จะเทียบกันได้ ก็ต่อเมื่อแบบ ทดสอบ 2 ฉบับนั้นคู่ขนานกันในด้านต่อไปนี้

โครงสร้าง (Structure), เวลา (Timing), ชนิดของข้อสอบ (Item Type), รูปแบบ (Format) และเนื้อหาวิชา (Subject Matter)

12. ผู้เชี่ยวชาญ หมายถึง ผู้ที่ทำหน้าที่ตรวจสอบคุณภาพของแบบทดสอบกับเนื้อหาและพฤติกรรม ตามตารางวิเคราะห์รายละเอียด ได้แก่

12.1 ครูผู้สอนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่มีประสบการณ์อย่างน้อย 3 ปีจำนวน 2 ท่าน

12.2 นักวัดผลที่จบปริญญาโททางการวัดผลการศึกษาจำนวน 3 ท่าน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ ผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการเทียบคะแนนโดยผู้วิจัยได้นำเสนอตามลำดับดังนี้

ตอนที่ 1 แนวคิดเกี่ยวกับการเทียบคะแนน

- 1.1 ความหมายของการเทียบคะแนน
- 1.2 เงื่อนไขของการเทียบคะแนน
- 1.3 ประเภทของการเทียบคะแนน
- 1.4 แบบแผนการเทียบคะแนน

ตอนที่ 2 ทฤษฎีเกี่ยวข้องกับการเทียบคะแนน

- 2.1 การเทียบคะแนนตามทฤษฎีคลาสสิกอล
 - 2.1.1 วิธีอิกวิเปอร์เซนไทล์
 - 2.1.2 วิธีเทียบแบบเส้นตรง
- 2.2 การเทียบคะแนนตามทฤษฎีการตอบข้อคำถาม
 - 2.2.1 แนวคิดเกี่ยวกับทฤษฎีการตอบข้อคำถาม
 - 2.2.2 วิธีเทียบคะแนนจริง
 - 2.2.3 วิธีเทียบคะแนนจริงโดยใช้ผู้สอบร่วม
 - 2.2.4 วิธีเทียบคะแนนสังเกตโดยใช้ผู้สอบร่วม

2.3 ความคลาดเคลื่อนของการเทียบคะแนน

2.4 ความเพียงพอของการเทียบคะแนน

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนน

- 1.1 งานวิจัยต่างประเทศ
- 1.2 งานวิจัยในประเทศ

ตอนที่ 1 แนวคิดเกี่ยวกับการเทียบคะแนน

1.1 ความหมายของการเทียบคะแนน

ได้มีผู้ให้ความหมายของการเทียบคะแนนไว้ต่างๆ ดังนี้

กัลลิกเซน (Gulliksen.1950: 298 –304) ได้ให้ความหมายของการเทียบคะแนนว่า คือวิธีการทำคะแนนที่ได้จากแบบทดสอบสองฉบับของวิชาเดียวกันให้เป็นคะแนนสมมูล (Equivalent Score) ที่เปรียบเทียบกันได้โดยตรง

ฟลานาแกน (Flanagan.1951: 747 –748) ได้ให้ความหมายของการเทียบคะแนนว่าเป็นวิธีการทำคะแนนจากแบบทดสอบต่างฉบับ ให้สามารถนำมาเปรียบเทียบกันได้

แองกอฟฟ์ (Angoff.1971: 562) ให้ความหมายของการเทียบคะแนน เป็นวิธีการเปลี่ยนระบบหน่วยคะแนนของแบบทดสอบฉบับหนึ่งไปสู่ระบบหน่วยคะแนนของแบบทดสอบอีกฉบับหนึ่ง เพื่อให้ได้คะแนนที่แปลงจากแบบทดสอบทั้งสองฉบับมีความเท่าเทียมกัน

ปีเตอร์เซนและคนอื่นๆ (Petersen and others.1989: 242) กล่าวว่าวิธีการเทียบคะแนนเป็นวิธีการเชิงประจักษ์ (Empirical Procedure) ในการสร้างความสัมพันธ์ระหว่างคะแนนดิบ(Raw Score) จากแบบทดสอบสองฉบับที่สามารถใช้คะแนนจากแบบทดสอบฉบับหนึ่งไปสู่อีกฉบับหนึ่งได้

ชูศักดิ์ ชัมภลลิขิต (2527 :144) กล่าวถึงการเทียบคะแนนแบบทดสอบ (Test Equating) ว่าเป็นกระบวนการที่เกี่ยวข้องกับกิจกรรม 2 ประการ คือ

1. กระบวนการที่ทำให้แบบทดสอบ 2 ฉบับใดๆ มีความทัดเทียมกันหรือเท่ากันในเชิงโครงสร้าง
2. การใช้วิธีการทางสถิติเพื่อปรับ (Adjust) คะแนนที่ได้จากแบบทดสอบแต่ละฉบับให้อยู่ในมาตราเดียวกันและเทียบกันได้

จากความหมายของการเทียบคะแนนดังกล่าว ผู้วิจัยพอสรุปได้ว่า การเทียบคะแนนเป็นวิธีการเชิงประจักษ์ที่ใช้สถิติเพื่อเปลี่ยนหน่วยระบบของคะแนนของแบบทดสอบต่างฉบับกันให้อยู่ในมาตราเดียวกันและสามารถนำมาเทียบกันได้

1.2 เงื่อนไขของการเทียบคะแนน

การเทียบคะแนนของแบบทดสอบสองฉบับจะต้องมีข้อกำหนด (Requirement) ดังนี้ (Hambleton and Swaminathan. 1985: 199 - 200)

1. ความทัดเทียมกัน (Equity) ลอร์ด (Lord.1980 : 195) กล่าวว่า “ถ้าการเทียบคะแนนของแบบทดสอบฉบับ X และ Y มีความยุติธรรมต่อผู้สอบแล้ว จะต้องไม่มีความแตกต่างที่ทุกระดับความสามารถ θ ไม่ว่าจะสามารถด้วยแบบทดสอบ ฉบับ X หรือ Y ก็ตาม” และมีข้อกำหนดดังนี้

1.1 แบบทดสอบที่วัดคุณลักษณะ (Trait) หรือความสามารถ (Ability) ต่างกัน จะไม่สามารถเทียบคะแนนกันได้

1.2 คะแนนแบบทดสอบที่มีค่าความเชื่อมั่นไม่เท่ากันแล้ว ไม่สามารถเทียบกันได้

1.3 คะแนนของแบบทดสอบที่มีระดับความยากต่างกัน ในการเทียบคะแนนตามแนวคิด ไม่สามารถจะเทียบคะแนนกันได้ เนื่องจากคะแนนจริงของแบบทดสอบจะมีความสัมพันธ์กันแบบไม่เป็นเส้นตรง และทำให้แบบทดสอบจะมีค่าความเชื่อมั่นไม่เท่า ณ ระดับความสามารถ θ ที่แตกต่างกัน

1.4 การแจกแจงความถี่แบบมีเงื่อนไข (The Conditional Frequency Distribution) ณ ระดับความสามารถ θ ($f(x|\theta)$) ของคะแนน θ จากแบบทดสอบฉบับ X จะต้องเป็นเช่นเดียวกับการแปลงคะแนน $x(y)$ เมื่อ $f(x(y)|\theta)$ โดยที่ $x(y)$ เป็น ฟังก์ชันแบบหนึ่งต่อหนึ่งของ y

1.5 คะแนนที่อาจผิดพลาดจากแบบทดสอบฉบับ X และ Y ไม่สามารถจะเทียบกันได้นอกจากว่าแบบทดสอบฉบับ X และ Y คู่ขนานกัน

1.6 แบบทดสอบที่มีความเชื่อมั่นที่ดีสามารถเทียบกันได้

2. ความไม่แปรเปลี่ยนไปตามกลุ่ม (Invariance Across Group) คือคะแนนของกลุ่มต่างๆที่ผ่านการเทียบแล้ว จะมีค่าไม่แตกต่างกันไม่ว่าจะเทียบจากกลุ่มตัวอย่างใด

3. ความสมมาตร (Symmetry) คือผลการเทียบคะแนนจากแบบทดสอบฉบับ X ไปสู่อันดับ Y หรือจากฉบับ Y ไปสู่อันดับ X ก็ต้องเหมือนกัน

4. ความเป็นมิติเดียวกันของแบบทดสอบ (Unidimensionality of the Test) คือแบบทดสอบทั้งสองฉบับจะต้องวัดคุณลักษณะเดียวกัน (คุณลักษณะแฝง, ความสามารถ, หรือทักษะ)

1.3 ประเภทของการเทียบคะแนนแฮมเบิลตันและสวามินาธาน (Hambleton and Swaminathan, 1985 : 197 – 198) แบ่งการเทียบคะแนนออกเป็น 2 ประเภท ได้แก่

1. การเทียบคะแนนตามแนวนอน (Horizontal Equating) คือการเทียบคะแนนจากการทดสอบที่ข้อสอบมีความยากเท่าๆกัน และผู้สอบมีความสามารถใกล้เคียงกัน

2. การเทียบคะแนนตามแนวตั้ง (Vertical Equating) คือการเทียบคะแนนจากการทดสอบที่ข้อสอบมีความยากต่างกัน และผู้สอบมีความสามารถต่างกันหรือต่างระดับชั้นกัน

สงบ. ลักษณะ (2525: 23) ได้แบ่งการเทียบคะแนนออกเป็น 2 กรณีคือ

1. การเทียบคะแนนตามแนวนอน (Horizontal Equating) เป็นกรณีที่มีข้อสอบ 2 ฉบับ คือ ฉบับ ก. และฉบับ ข. มีความยากง่ายปานๆกัน นำมาทดสอบกับบุคคลประชากรเดียวกันแล้วต้องหาว่าคะแนนของแบบทดสอบชุด ก. จึงจะเทียบเท่ากับคะแนนของแบบทดสอบชุด ข.

2. การเทียบคะแนนตามแนวตั้ง (Vertical Equating) เป็นกรณีที่มีแบบทดสอบสองชุด หรือมากกว่า 2 ชุด มีความยากต่างกัน นำไปทดสอบประชากรบุคคลต่างระดับความสามารถกัน กล่าวคือ ต้องการเทียบคะแนนจากแบบทดสอบสองชุดเมื่อความยากของแบบทดสอบต่างกัน และเมื่อสอบกับบุคคลที่มีความสามารถต่างกัน การเทียบมาตราในลักษณะนี้จะกระทำโดยมีแบบทดสอบร่วมฉบับหนึ่ง

1.4 แบบแผนการเก็บรวบรวมข้อมูล

ข้อมูลจากการสอบโดยใช้แบบทดสอบหลายฉบับและเก็บรวบรวมข้อมูล โดยใช้รูปแบบต่างกันนั้น ข้อมูลที่ได้จากบางรูปแบบไม่สามารถนำมาเทียบคะแนนได้ เช่น ข้อมูลที่ได้จากการสอบโดยใช้แบบทดสอบ 2 ฉบับ กับผู้สอบที่มีความสามารถแตกต่างกัน 2 กลุ่ม จะนำคะแนนที่ได้จากแบบทดสอบต่างฉบับมาเปรียบเทียบกันไม่ได้ โดยทั่วไปคะแนนจากแบบทดสอบต่างฉบับที่จะนำมาเปรียบเทียบกันได้ ต้องได้มาจากผู้สอบกลุ่มเดียวกัน หรือผู้สอบกลุ่มเท่าเทียมกัน หรือได้จากผู้สอบที่สอบข้อสอบร่วมกันรูปแบบใดรูปแบบหนึ่ง แบบแผนการเก็บรวบรวมข้อมูล ซึ่งสามารถนำคะแนนจากแบบทดสอบต่างฉบับมาเปรียบเทียบกันได้ จำแนกได้ดังนี้ (Petersen and others, 1989: 244 - 246)

แบบแผนที่ 1 กลุ่มผู้สอบกลุ่มเดียว (Single Group Design) เป็นแบบแผนที่ง่ายที่สุดในการเก็บรวบรวมข้อมูลโดยใช้กลุ่มตัวอย่างผู้สอบกลุ่มเดียว ทำได้โดยให้ผู้สอบแต่ละคนทำแบบทดสอบทั้งสองฉบับ ผู้สอบจะทำแบบทดสอบฉบับหนึ่งแล้วตามด้วยแบบทดสอบอีกฉบับหนึ่ง การให้ทำแบบทดสอบทั้ง 2 ฉบับ ทำให้ไม่มีปัญหาในเรื่องความยากของแบบทดสอบต่างฉบับ แต่การสอบครั้งหลัง ผู้สอบอาจได้รับผลกระทบจากการสอบครั้งแรก เนื่องจากการเรียนรู้ ความเมื่อยล้า และการฝึกฝน

แบบแผนที่ 2 ผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (Counterbalanced Random Group Design) ในทางปฏิบัติ ลำดับของการทดสอบก่อนหลัง ย่อมส่งผลต่อการสอบ ดังนั้น แบบแผนนี้จึงปรับมาจากแบบแผนแรกโดยจัดผู้สอบเป็นกลุ่มสุ่ม 2 กลุ่ม แต่ละกลุ่มย่อยได้รับการสอบด้วยแบบทดสอบ 2 ฉบับ ในลักษณะการจัดลำดับก่อนหลัง สลับให้เกิดความสมดุล

การจัดแบบนี้เพื่อให้การสอบโดยใช้แบบทดสอบ 2 ฉบับ ได้รับอิทธิพลจากองค์ประกอบในเรื่องการเรียนรู้ ความเมื่อยล้าและการฝึกฝนในลักษณะที่สมดุลกัน

แบบแผนที่ 3 ผู้สอบกลุ่มเท่าเทียมกัน (Equivalent Group Design) ในทางปฏิบัติการให้ผู้สอบทำแบบทดสอบทั้ง 2 ฉบับ อาจจะทำได้ยาก การใช้ผู้สอบกลุ่มเท่าเทียมกันเป็นรูปแบบที่สามารถนำมาใช้ได้ แบบแผนนี้จำเป็นต้องจัดกลุ่มผู้สอบให้มีความคล้ายคลึงกันมากเท่าที่จะทำได้ ผู้สอบในแต่ละกลุ่มจะทำแบบ

ทดสอบฉบับเดียว ดังนั้นจึงหลีกเลี่ยงปัญหาเรื่องการเรียนรู้ ความเมื่อยล้าและการฝึกฝนจากการสอบครั้งแรกได้ แต่จะไม่มีข้อมูลที่จะนำมาใช้ปรับความแตกต่างของกลุ่มตัวอย่างทั้งสอง วิธีการที่จะลดความแตกต่างระหว่างกลุ่มทำได้โดยใช้กลุ่มตัวอย่างขนาดใหญ่

แบบแผนที่ 4 ผู้สอบกลุ่มสุ่มสองกลุ่มโดยมีแบบทดสอบร่วม (Anchor Test Random Group Design) การใช้แบบแผนนี้อาจมีข้อจำกัดในเรื่องความแตกต่างระหว่างกลุ่มตัวอย่าง การใช้แบบทดสอบร่วมด้วยทำให้สามารถปรับความแตกต่างระหว่างกลุ่มได้ตามแบบแผนนี้ กลุ่มสุ่มแต่ละกลุ่มจะได้รับการสอบแบบทดสอบกลุ่มละ 1 ฉบับ และทั้งสองกลุ่มจะได้รับการสอบด้วยแบบทดสอบร่วมด้วย ซึ่งอาจจะจัดไว้ในแบบทดสอบเรียกว่าแบบทดสอบร่วมภายใน (Internal Anchor Test) หรือจัดแยกไว้จากแบบทดสอบที่ต้องการเทียบคะแนนก็ได้ เรียกว่าแบบทดสอบร่วมภายนอก (External Anchor Test)

แบบแผนที่ 5 กลุ่มผู้สอบที่ไม่เท่าเทียมกันโดยมีแบบทดสอบร่วม (Anchor Test Non-Equivalent Group Design) ในการสอบบางครั้งจำเป็นต้องใช้แบบทดสอบฉบับใหม่กับผู้สอบกลุ่มใหม่ในโปรแกรมการสอบเดียวกัน การสอบโดยใช้แบบทดสอบต่างฉบับและเวลาสอบต่างกันั้น ผู้สอบไม่ได้มาจากกลุ่มประชากรเดียวกัน จึงจำเป็นต้องใช้แบบทดสอบร่วมในการปรับ

ความแตกต่างระหว่างกลุ่ม ตามแบบแผนนี้ผู้สอบในแต่ละกลุ่มจะได้รับแบบทดสอบ 1 ฉบับและแบบทดสอบร่วม ซึ่งต้องมีเนื้อหาและระดับความยากคล้ายกับแบบทดสอบที่ต้องการเทียบคะแนนให้มากที่สุดเท่าที่จะเป็นไปได้ คะแนนจากแบบทดสอบร่วมจะใช้เพื่อลดความลำเอียงในการเทียบคะแนนจากกลุ่มที่ต่างกัน

จากแบบแผนการเก็บรวบรวมข้อมูลที่กำลังกล่าวมาทั้งหมด สามารถอภิปรายได้โดย ภาพประกอบดังนี้

แบบแผนที่ 1 รูปแบบผู้สอบกลุ่มเดียว

กลุ่มตัวอย่าง	แบบทดสอบฉบับที่ 1	แบบทดสอบฉบับที่ 2
P1	*	*

แบบแผนที่ 2 รูปแบบผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล

กลุ่มตัวอย่าง	แบบทดสอบฉบับที่ 1		แบบทดสอบฉบับที่ 2	
	สอบครั้งที่ 1	สอบครั้งที่ 2	สอบครั้งที่ 1	สอบครั้งที่ 2
P1	*			*
P2		*	*	

แบบแผนที่ 3 รูปแบบผู้สอบกลุ่มเท่าเทียมกัน

กลุ่มตัวอย่าง	แบบทดสอบฉบับที่ 1	แบบทดสอบฉบับที่ 2
P1	*	
P2		*

แบบแผนที่ 4 รูปแบบผู้สอบกลุ่มผู้สอบกลุ่มโดยมีแบบทดสอบร่วม

กลุ่มตัวอย่าง	แบบทดสอบ		
	ฉบับที่ 1	ฉบับที่ 2	แบบทดสอบร่วม
P1	*		*
P2		*	*

แบบแผนที่ 5 รูปแบบผู้สอบกลุ่มไม่เท่าเทียมกันโดยมีแบบทดสอบร่วม

กลุ่มตัวอย่าง	แบบทดสอบ		
	ฉบับที่ 1	ฉบับที่ 2	แบบทดสอบร่วม
P1	*		*
Q1		*	*

ภาพประกอบ 1 แบบแผนการเก็บรวบรวมข้อมูลเพื่อเทียบคะแนนแบบทดสอบ

ตอนที่ 2 ทฤษฎีที่เกี่ยวข้องกับการเทียบคะแนน

2.1 การเทียบคะแนนตามทฤษฎีคลาสสิกอล

2.1.1 วิธีการเทียบคะแนนอีควิเปอร์เซ็นต์ไทล์ (Equipercntile Equating)

การเทียบคะแนนโดยวิธีการนี้อยู่บนพื้นฐานของนิยามที่ว่า คะแนนจากแบบทดสอบฉบับ X และฉบับ Y (เมื่อแบบทดสอบ X และ Y วัดสิ่งเดียวกัน และมีความเชื่อมั่นสูงพอๆกัน) จะถือว่าเป็นคะแนนสมมูลกัน (Equivalent score) เมื่อคะแนนของแต่ละฉบับนั้นมีตำแหน่งเปอร์เซ็นต์ไทล์ (Percentile ranks) เท่ากัน

นั่นคือ X_i และ Y_j เป็นคะแนนสมมูลกัน

เมื่อ $X_{PR_k} = Y_{PR_k}$

โดย X_{PR_k} คือ คะแนนจากแบบทดสอบฉบับ X ตรงตำแหน่งเปอร์เซ็นต์ไทล์ k

Y_{PR_k} คือ คะแนนจากแบบทดสอบฉบับ Y ตรงตำแหน่งเปอร์เซ็นต์ไทล์ k

$$\text{โดย } X_{PR_k} = LO + \left[\frac{\frac{kn}{100} - F_b}{f} \right] \times C$$

เมื่อ LO คือ คะแนนขีดจำกัดล่างที่แท้จริงของคะแนนในชั้นเปอร์เซ็นต์ไทล์ที่ k

n คือ จำนวนสมาชิกในกลุ่มตัวอย่าง

f คือ ความถี่ของคะแนนในชั้นเปอร์เซ็นต์ไทล์ที่ k

F_b คือ ความถี่สะสมก่อนถึงคะแนนในชั้นเปอร์เซ็นต์ที่ k

c คือ อัตรากาชั้น

วิธีการ การเทียบคะแนนระหว่างแบบทดสอบ ฉบับ X และแบบทดสอบ Y โดยวิธีอิกวิเปอร์เซ็นต์ไทล์ สามารถกระทำไดดังนี้

1. สุ่มกลุ่มผู้สอบให้เป็นตัวแทนประชากรเป้าหมายที่ต้องการนำแบบทดสอบไปใช้ โดยให้ได้กลุ่มผู้สอบที่มีความสามารถกระจาย (สูง, กลาง, ต่ำ) ครอบคลุมลักษณะประชากรเป้าหมาย แล้วทำการสุ่มแยกผู้สอบออกเป็น 2 กลุ่ม โดยให้กลุ่มหนึ่งทำแบบทดสอบฉบับ X และอีกกลุ่มหนึ่งทำแบบทดสอบฉบับ Y
2. นำคะแนนของแบบทดสอบแต่ละฉบับมาคำนวณหาค่าคะแนน ณ ตำแหน่งเปอร์เซ็นต์ไทล์ต่างๆ

3. สร้างตารางเทียบคะแนนระหว่างแบบทดสอบ ฉบับ X และแบบทดสอบฉบับ Y โดยถือว่าคะแนน ณ ตำแหน่งเปอร์เซ็นต์ไทล์เดียวกันของแบบทดสอบทั้ง 2 ฉบับ เป็นคะแนนสมมูลกัน

วิธีอิกวิเปอร์เซ็นต์ไทล์เป็นการสร้างคะแนนสมมูลระหว่างแบบทดสอบบนพื้นฐานความเท่าเทียมกันของค่าโมเมนต์ทั้ง 3 ทางสถิติ ซึ่งได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และการแจกแจงของคะแนน จึงเป็นกระบวนการแปลงคะแนนแบบไม่ใช่เส้นตรง (Nonlinear Transformation) ยังผลให้ความสัมพันธ์ระหว่างคะแนนดิบและคะแนนแปลงไม่เป็นเส้นตรง ซึ่งมีหลักการเกิดขึ้นจากความต้องการพื้นฐาน คือ $F=F^*$ โดยมีฟังก์ชันการเทียบคะแนนด้วยวิธีนี้คือ $e_X(Y) = F^{-1}(G(Y))$ (อ้างจาก ชูชีพ พงษ์สมบูรณ์ . 2528 :16)

1. ในทางทฤษฎี $F(x)$ และ $G(y)$ มาจากกลุ่มตัวอย่างที่ได้จากการสุ่ม (P_1 และ P_2) จากประชากร P
2. ในทางปฏิบัติ สถาบันบริการทดสอบทางการศึกษา (Educational Testing Service) ทำการสร้างแบบทดสอบฟอร์ม X และฟอร์ม Y แล้วสุ่มสลับกันลงในกล่องในแต่ละสนามสอบ ฉะนั้นผู้สอบครั้งหนึ่งจะสอบฟอร์ม X และอีกครั้งหนึ่งจะสอบฟอร์ม Y
3. การเทียบโดยใช้เปอร์เซ็นต์ไทล์จะเทียบ ฟอร์ม X และ Y โดยผู้สอบจากประชากร P
4. ทั้ง F และ G เป็นลักษณะต่อเนื่อง (Continuous) จาก $F^{-1}(G(Y))$
5. ความยาวของแบบทดสอบฟอร์ม X และ Y ไม่ควรแตกต่างกันมาก

ในการนำวิธีเทียบคะแนนอิกวิเปอร์เซ็นต์ไทล์ไปใช้มีข้อแนะนำพอสรุปได้ดังนี้ (Potthoff, 1982 ; Angoff, 1984)

1. ควรใช้กลุ่มผู้สอบขนาดใหญ่ ควรเป็นกลุ่มที่มีความสามารถค่อนข้างกระจายและกระจายพอๆกัน ถ้าใช้กลุ่มตัวอย่างขนาดเล็กทำให้คะแนนมีความไวต่อความแปรปรวนเชิงสุ่ม ผลการเปรียบเทียบคะแนนสามารถแปรผันไปตามกลุ่มผู้สอบได้ (Group Dependent)
2. การสร้างกราฟเปอร์เซ็นต์ไทล์ เพื่อเปรียบเทียบคะแนนระหว่างแบบทดสอบ ควรสร้างด้วยความระมัดระวัง ไม่ลำเอียง
3. แบบทดสอบที่ต้องการนำมาเปรียบเทียบคะแนนกัน ต้องมีความเชื่อมั่นสูงพอๆกัน ถ้าต่างกันมาก จะทำให้การเปรียบเทียบคะแนนขาดความคงที่
4. การเปรียบเทียบคะแนนควรอยู่ในช่วงพิสัยของคะแนนสังเกต การเปรียบเทียบคะแนนที่อยู่นอกพิสัยคะแนนสังเกตจะมีความคลาดเคลื่อนสูง

2.1.2 วิธีการเทียบคะแนนแบบเส้นตรง วิธีการนี้ตั้งอยู่ในนิยามที่ว่า คะแนนจากแบบทดสอบ 2 ฉบับใดๆ ต่างก็วัดในสิ่งเดียวกันจะพิจารณาให้เทียบกันได้ ก็ต่อเมื่อคะแนนมาตรฐานทั้งสองชุดมีค่าเท่ากัน

$$(Y - M_y) / S_y = (X - M_x) / S_x$$

เมื่อ

X คือ คะแนนจากแบบทดสอบ X

Y คือ คะแนนจากแบบทดสอบชุด Y

M_x คือ ค่าเฉลี่ยของแบบทดสอบ X

M_y คือ ค่าเฉลี่ยของแบบทดสอบ Y

S_x คือ ค่าเบี่ยงเบนมาตรฐานของแบบทดสอบ X

S_y คือ ค่าเบี่ยงเบนมาตรฐานของแบบทดสอบ Y

2.2 การเทียบคะแนนโดยใช้ทฤษฎีการตอบข้อคำถาม

2.2.1 แนวคิดเกี่ยวกับทฤษฎีการตอบข้อคำถาม ในการเทียบคะแนนตาม

ทฤษฎีการตอบข้อคำถาม อาศัยรูปแบบความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้สอบข้อสอบใดๆ ในแบบทดสอบว่าขึ้นอยู่กับความสามารถของผู้สอบซึ่งการประมาณค่าได้จากโค้งลักษณะเฉพาะของข้อสอบ (Item Characteristic Curve: ICC) ของแต่ละรูปแบบที่ใช้ ซึ่งอาจจะเป็นรูปแบบ หนึ่ง สอง หรือ สาม พารามิเตอร์ ทั้งนี้ไม่ขึ้นกับการแจกแจงของคะแนนของกลุ่มตัวอย่าง (Lord and Novick.1968 : 359) ดังนั้น การเทียบคะแนนด้วยวิธีนี้จะสามารถประมาณค่าความสามารถแท้ของผู้สอบแต่ละคนได้ (Harris and Hoover.1987 : 151) นิยามการเทียบคะแนนโดยใช้ทฤษฎีการตอบข้อคำถาม กล่าวว่า คะแนนแปลง Y และคะแนนดิบ X จะเทียบกันได้ก็ต่อเมื่อ ไม่ว่าผู้สอบจะสอบด้วยแบบทดสอบชุด X หรือชุด Y ก็จะไม่มีความแตกต่างกัน (Lord. 1977 : 128) ดังนั้นนิยามการเทียบคะแนน จึงเน้นความสำคัญของความทัดเทียมกัน (Equity) หรือความเป็นธรรมที่บุคคลจะได้รับการทดสอบอย่างเท่าเทียมกัน ถึงแม้จะสอบด้วยแบบทดสอบต่างชุดก็ตาม (Lord.1982 : 14) ผู้สอบจะไม่ได้เสียเปรียบหรือได้เปรียบแต่อย่างใดเมื่อได้รับการทดสอบด้วยแบบทดสอบชุดใดชุดหนึ่งที่มีความยากหรือง่ายกว่ากัน (Angoff.1971 : 563)ต่อมา ลอร์ด (Lord.1980: 199) ได้ตั้งข้อกำหนดที่สำคัญ 3 ประการในการเทียบคะแนนจากแบบทดสอบ 2 ชุด ที่วัดคุณลักษณะเดียวกัน ซึ่งเทียบคะแนนที่สมมูลกันที่ระดับความสามารถเดียวกัน ได้ดังนี้

1. ความทัดเทียมกัน หมายถึง ทุกๆความสามารถ θ การแจกแจงความถี่แบบมีเงื่อนไขของการแปลงคะแนน $x(y)$ หรือ y หรือ ที่ θ ที่กำหนดให้ต้องเหมือนกับการแจกแจงความถี่แบบมีเงื่อนไขของคะแนน x

2. ความไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง หมายถึง คะแนนแปลง $x(y)$ จะคงที่โดยไม่ขึ้นกับตัวแปรด้านอื่นๆของประชากรที่นำมาเทียบคะแนน

3. ความสมมาตร หมายถึง การเทียบคะแนนจะต้องเหมือนกันไม่ว่าการเทียบนั้นจะเป็นการเทียบจากแบบทดสอบฉบับ X หรือฉบับ Y ไปยังแบบทดสอบ X ก็ตาม โมเดลการเทียบคะแนนแบ่งออกได้เป็น 3 รูปแบบ ซึ่งแตกต่างกันไปตามพารามิเตอร์ที่ใช้ในแต่ละรูปแบบดังนี้

3.1 รูปแบบที่ใช้พารามิเตอร์ 1 ตัว (One – Parameter Logistic Model) ซึ่ง ราสซ (Rasch) ได้พัฒนาขึ้นมาใช้รูปแบบนี้โดยมีข้อตกลงว่า ทุกข้อไม่มีการเดา $c_i = 0$ และทุกข้อมีอำนาจจำแนกเท่ากัน $a=1$ ซึ่งมีสูตรในการคำนวณดังนี้

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

$P_i(\theta)$ คือ ความน่าจะเป็นของผู้เข้าสอบที่มีความสามารถ θ จะ ตอบข้อสอบข้อที่ i ได้ถูกต้อง

D คือ ค่าคงที่ของ scale factor เพื่อปรับค่าของ ฟังก์ชันโลจิสติก ให้ใกล้เคียงกับ ฟังก์ชันนอร์มอลลอจีว มีค่าเท่ากับ 1.70

b คือ ระดับความยากของข้อสอบข้อที่ i

e คือ ค่าคงที่ = 2.718

n คือ จำนวนข้อสอบในแบบทดสอบ

3.2 รูปแบบที่ใช้พารามิเตอร์ 2 ตัว (Two-Parameter Logistic Model)

เบียร์นบอม (Birnbaum) เสนอแบบจำลองการตอบข้อคำถามที่ใช้กันเส้นโค้งลักษณะของข้อคำถามอยู่ในรูปการแจกแจงแบบโลจิสติกที่มีพารามิเตอร์ 2 ตัว ดังนี้

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

$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบที่มีความสามารถจะตอบข้อสอบข้อที่ i ได้ถูกต้อง

b คือ ความยากของข้อสอบข้อที่ i

a คือ อำนาจจำแนกของข้อสอบข้อที่ i

n คือ จำนวนข้อสอบในแบบทดสอบ

D คือ ค่าคงที่ของ scale factor เพื่อปรับค่าของ ฟังก์ชันโลจิสติก ให้ใกล้เคียงกับ ฟังก์ชันนอร์มอลลอจีว มีค่าเท่ากับ 1.70

e คือ ค่าคงที่ = 2.718

3.3 รูปแบบที่ใช้พารามิเตอร์ 3 ตัว (Three- Parameter Logistic Model) รูปแบบ 3 พารามิเตอร์ สามารถเขียนได้จากรูปแบบ 2 พารามิเตอร์ โดยเพิ่มพารามิเตอร์ตัวที่ 3 คือ ค่าการเดา ซึ่งมีสูตรในการคำนวณดังนี้

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

$P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบที่มีความสามารถจะตอบข้อสอบข้อที่ i ได้ถูกต้อง

b คือ ความยากของข้อสอบข้อที่ i

a คือ อำนาจจำแนกของข้อสอบข้อที่ i

n คือ จำนวนข้อสอบในแบบทดสอบ

D คือ ค่าคงที่ของ scale factor เพื่อปรับค่าของ ฟังก์ชันโลจิสติก ให้ใกล้เคียงกับฟังก์ชันนอร์มอลลอจีว มีค่าเท่ากับ 1.70

e คือ ค่าคงที่ = 2.718

พารามิเตอร์ c เป็นค่ากำกับต่ำสุด (Lower Asymptote) ของฟังก์ชัน ซึ่งเป็นความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำมากๆ จะตอบข้อสอบข้อที่ i ได้ถูก การนำพารามิเตอร์ c นี้เข้ามาในแบบจำลองเพื่อใช้เป็นข้อมูลการตอบ ของผู้สอบที่มีความสามารถต่ำในการทำข้อสอบแบบเลือกตอบเนื่องจากการเดาก็เป็นปัจจัยหนึ่งที่มีอิทธิพลต่อการตอบข้อสอบ ดังนั้นจึงมักเรียกพารามิเตอร์ c นี้ว่า เป็นพารามิเตอร์การเดา (Guessing Parameter)

กระบวนการเทียบคะแนนตามทฤษฎีการตอบข้อคำถาม

ในการเทียบคะแนนตามทฤษฎีการตอบข้อคำถาม เมื่อเลือกแบบจำลองการตอบข้อคำถามที่เหมาะสมได้แล้ว และไม่ว่าการเทียบคะแนนจะเป็นการเทียบคะแนนตามแนวนอนหรือการเทียบคะแนนตามแนวตั้ง กระบวนการในการเทียบคะแนนประกอบด้วย 3 ขั้นตอน คือ

1. การเลือกแบบแผนการเก็บรวบรวมข้อมูล
2. การจัดค่าพารามิเตอร์ของแบบทดสอบต่างฉบับให้อยู่ในสเกลเดียวกัน
3. การเทียบค่าประมาณความสามารถของผู้สอบ

รายละเอียดของกระบวนการเทียบคะแนนแต่ละขั้นตอนมีดังนี้

1. การเลือกแบบแผนการเก็บรวบรวมข้อมูล แบบแผนการเก็บรวบรวมข้อมูลจากแบบทดสอบต่างฉบับเพื่อนำมาเทียบคะแนนมีหลายรูปแบบ ดังได้กล่าวไว้แล้ว การเลือกรูปแบบใดพิจารณาได้จากธรรมชาติของแบบทดสอบและกลุ่มผู้สอบ ซึ่งอาจเลือกใช้รูปแบบใดรูปแบบหนึ่งตามความเหมาะสม
2. การจัดค่าพารามิเตอร์ของแบบทดสอบต่างฉบับให้อยู่ในสเกลเดียวกัน เนื่องจากการประมาณค่าพารามิเตอร์ของข้อสอบในแบบทดสอบต่างฉบับ อาจกระทำพร้อมกันหรือกระทำแยกกัน ในกรณีที่การประมาณค่าพารามิเตอร์ของข้อสอบต่างฉบับทำพร้อมกัน ค่าประมาณที่ได้จะอยู่ในมาตราเดียวกันจึงสามารถนำมาเปรียบเทียบกันได้เลย แต่ในบางกรณีที่จำเป็นต้องประมาณค่าพารามิเตอร์แยกกัน เช่น กรณีที่ผู้สอบแบบทดสอบแต่ละฉบับมีระดับความสามารถต่างกัน การประมาณค่าพารามิเตอร์ของข้อสอบต้องกระทำแยกกัน ค่าประมาณที่ได้จึงไม่อยู่ในสเกลเดียวกัน ดังนั้นต้องมีการปรับค่าประมาณพารามิเตอร์ของข้อสอบจากแบบทดสอบต่างฉบับให้อยู่ในสเกลเดียวกันก่อนจึงจะนำค่าที่ได้ไปใช้ในการเปรียบเทียบกัน การทำให้ค่าพารามิเตอร์ของแบบทดสอบ 2 ฉบับใดๆ อยู่ในสเกลเดียวกันต้องอาศัยแบบทดสอบร่วมในแบบทดสอบในแบบทดสอบทั้ง 2 ฉบับ หรือต้องมีผู้สอบร่วมส่วนหนึ่ง ซึ่งทำแบบทดสอบทั้ง

2 ฉบับ จึงจะสามารถนำข้อมูลที่ร่วมกัน 2 ฉบับ มาใช้ปรับค่าพารามิเตอร์ของข้อสอบจากแบบทดสอบต่างฉบับนั้นให้อยู่ในสเกลเดียวกันได้

วิธีการโดยทั่วไปที่ใช้ในการจัดให้ค่าพารามิเตอร์ของข้อสอบจากแบบทดสอบต่างฉบับมาอยู่ในสเกลเดียวกันมีดังนี้ (Petersen and others. 1989 : 256)

2.1 วิธีประมาณค่าพร้อมกัน (Concurrent Calibration) ตามวิธีการนี้ พารามิเตอร์ของแบบทดสอบ 2 ฉบับ จะถูกประมาณค่าพร้อมกัน โดยใช้ ข้อมูลจากผู้สอบรวมหรือข้อสอบรวม (แล้วแต่กรณี) จะทำให้ค่าพารามิเตอร์ของข้อสอบและค่าประมาณความสามารถอยู่ในสเกลเดียวกัน

2.2 วิธีกำหนดค่าความยากให้คงที่ (The Fixed bs Method) ตามวิธีการนี้ จะประมาณค่าพารามิเตอร์ของแบบทดสอบทีละฉบับ หลังจากประมาณค่าพารามิเตอร์แบบทดสอบฉบับแรกแล้ว จะใช้ค่าความยากของข้อสอบรวมไปประมาณค่าพารามิเตอร์ของแบบทดสอบฉบับที่สองโดยไม่ต้องประมาณค่าพารามิเตอร์จากแบบทดสอบฉบับที่สองซ้ำ ดังนั้น พารามิเตอร์ของความสามารถและพารามิเตอร์ของข้อสอบจากแบบทดสอบฉบับที่สองจึงอยู่บนมาตราเดียวกับแบบทดสอบฉบับแรก

2.3 วิธีเทียบค่าความยาก (The Equated bs Method) ตามวิธีการนี้ พารามิเตอร์ของแบบทดสอบแต่ละฉบับจะถูกประมาณค่าแยกกัน ดังนั้น ในกลุ่มของแบบทดสอบรวม ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของความยากของแบบทดสอบทั้ง 2 ฉบับ จะถูกกำหนดให้เท่ากันโดยใช้การแปลงเชิงเส้นตรงเพื่อแปลงค่าพารามิเตอร์ของความสามารถและค่าพารามิเตอร์ความยาก (b) และอำนาจจำแนก (a) ของแบบทดสอบฉบับที่ 2 ไปสู่สเกลของแบบทดสอบฉบับแรก

2.4 วิธีใช้โค้งลักษณะข้อสอบ (The Characteristic Curve Method) ตาม วิธีการนี้พารามิเตอร์ของข้อสอบจากแบบทดสอบ 2 ฉบับ จะประมาณค่าแยกกัน พารามิเตอร์ที่ได้จากการประมาณค่า 2 ครั้งนี้ จะแตกต่างกันเฉพาะจุดเริ่มต้น (Origin) และหน่วยของการวัดเท่านั้น ดังนั้นจึงแปลงค่าประมาณพารามิเตอร์ที่ได้ให้อยู่ในมาตราเดียวกันโดยใช้การแปลงเชิงเส้นตรง

3. การเทียบค่าประมาณความสามารถของผู้สอบต้องพิจารณาว่าคะแนนจากแบบทดสอบจะรายงานในรูปคะแนนความสามารถหรือคะแนนจริงหรือคะแนนสังเกต

การเทียบคะแนนมีวิธีการแตกต่างกัน ดังนี้

3.1 ในกรณีที่การรายงานผลคะแนนจากแบบทดสอบรายงานในรูปคะแนนความสามารถการเทียบคะแนนสามารถทำได้เลย โดยใช้ค่าประมาณความสามารถที่ได้จากแบบทดสอบต่างฉบับ

3.2 ในกรณีที่การรายงานผลที่ทำในรูปคะแนนจริง จะต้องมีการประมาณค่าคะแนนจริงจากแบบทดสอบแต่ละฉบับโดยใช้ค่าประมาณพารามิเตอร์ที่ได้ในขั้นตอนที่ 2 คะแนนจริงจะถูกประมาณค่าหลายๆค่า เพื่อนำไปใช้ สร้างกราฟหรือตารางเพื่อเทียบคะแนนจริงที่ตรงกันจากแบบทดสอบแต่ละฉบับ

3.3 ในกรณีที่รายงานคะแนนทำในรูปคะแนนสังเกตจะต้องทำการแจกแจงคะแนนสังเกตอย่างมีเงื่อนไขไปตามทฤษฎี สำหรับความสามารถที่ตรงกันของกลุ่มผู้สอบเพื่อเทียบคะแนนที่สร้างขึ้นด้วยวิธีเทียบคะแนนอิกวิเปอร์เซนไทล์ แล้วทำการเทียบคะแนนสังเกตที่แท้จริงจากตารางหรือกราฟของการเปรียบเทียบ

2.2.2 การเทียบคะแนนจริง ในการเทียบคะแนนจริงนี้ทำได้โดยแปลงค่า θ ไปสู่คะแนนจริงที่สอดคล้องกันได้ ทั้งนี้โดยอาศัยสมการสัมพันธ์เชิงคณิตศาสตร์ระหว่างค่าความสามารถ θ และคะแนนจริงที่ได้จากแบบทดสอบแต่ละฉบับ (ฉบับ X และฉบับ Y) ดังนี้ (Hambleton and Swaminathan.1985: 212)

$$\xi_x = \sum_{i=1}^n P_i(\theta_x)$$

เมื่อ $P_i(\theta_x)$ คือ ค่าความน่าจะเป็นในการตอบข้อสอบข้อที่ i ได้ถูกต้องเมื่อผู้สอบมีความสามารถ θ

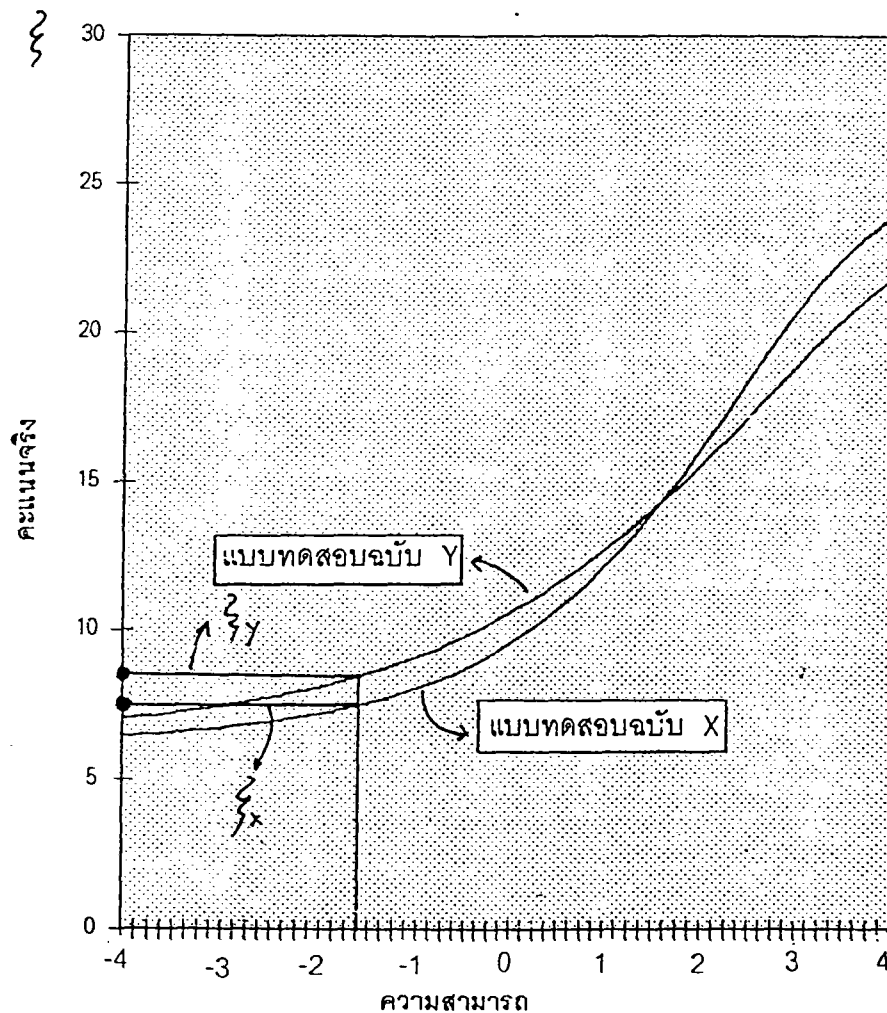
θ_x คือ ระดับความสามารถของผู้สอบในแบบทดสอบชุด X

ξ_x คือ คะแนนจริงที่ได้จากแบบทดสอบชุด X

ในทำนองเดียวกัน ถ้า θ_y คือระดับความสามารถของผู้สอบในแบบทดสอบชุด Y แล้ว คะแนนจริงที่สมมูลกันในแบบทดสอบชุด Y คือ

$$\xi_y = \sum_{j=1}^m P_j(\theta_y) \equiv \sum_{j=1}^m P_j(\alpha\theta_x + \beta)$$

เมื่อ $\theta_y = \alpha\theta_x + \beta$ คือ เส้นความสัมพันธ์ระหว่าง θ_x และ θ_y สำหรับให้ค่าคะแนนจริงที่สมมูลกันแต่ละคู่ (ξ_x, ξ_y) ดังนั้น จึงมีความเป็นไปได้ที่จะเทียบคะแนนจริงจากแบบทดสอบต่างชุด โดยคำนวณค่าคงที่ของการเทียบคะแนน α และ β หลังจากนั้นจะพล็อตกราฟเพื่อเทียบคะแนนจริงระหว่างแบบทดสอบได้ ดังภาพประกอบ (Hambleton and Swaminathan, 1985:213)



ภาพประกอบ 2 การเทียบคะแนนจริงตามทฤษฎีการตอบข้อคำถามระหว่างแบบทดสอบฉบับ X และฉบับ Y

2.2.3 การเทียบคะแนนจริงโดยใช้ผู้สอบร่วม

ในการเทียบคะแนนจริงโดยใช้ผู้สอบร่วม มีวิธีการเทียบคือถ้าแบบทดสอบฉบับ X และฉบับ Y สอบกับผู้สอบร่วม 2 กลุ่ม แต่ละกลุ่มย่อยได้รับการสอบด้วยแบบทดสอบทั้ง 2 ฉบับ ในลักษณะการจัดลำดับก่อน – หลัง สลับให้เกิดความสมดุล การจัดรูปแบบนี้เพื่อให้การสอบโดยใช้แบบทดสอบทั้ง 2 ฉบับ ได้รับอิทธิพลจากองค์ประกอบในเรื่องการเรียนรู้ ความเมื่อยล้าและการฝึกฝนในลักษณะที่สมดุลกัน ก็มีกระบวนการในการเทียบคะแนนเช่นเดียวกันกับการเทียบคะแนนจริงจากแบบทดสอบฉบับ X

$$\xi_x = \sum_{i=1}^n P_i(\theta_x)$$

เมื่อ $P_i(\theta_x)$ คือ ค่าความน่าจะเป็นในการตอบข้อสอบข้อที่ i ได้ถูกต้องเมื่อผู้สอบมีความสามารถ θ

θ_x คือ ระดับความสามารถของผู้สอบในแบบทดสอบชุด X

ξ_x คือ คะแนนจริงที่ได้จากแบบทดสอบชุด X

และสมการการเทียบคะแนนจริงจากฉบับ Y

$$\xi_y = \sum_{j=1}^m P_j(\theta_y)$$

เมื่อ ξ_y คือ คะแนนจริงที่ได้จากแบบทดสอบฉบับ Y

θ_y คือ ระดับความสามารถของผู้สอบในแบบทดสอบฉบับ Y

ดังนั้น คะแนนจริงที่สมดุลกันในแบบทดสอบฉบับ X และฉบับ Y คือ

$$\xi_x = \sum_{j=1}^m P_j(\theta_y) \equiv \sum_{j=1}^m P_j(\alpha\theta_x + \beta)$$

เมื่อ α, β คือ ค่าคงที่ของการเทียบคะแนน หาได้โดยวิธีใช้โค้งลักษณะเฉพาะของข้อสอบซึ่งจะกำหนดความสัมพันธ์ระหว่างค่าความยากของข้อสอบ จากสมการ

$$b_y = \alpha b_x + \beta$$

$$\text{เมื่อ } b_{yi} = \alpha b_{xi} + \beta$$

$$a_{yi} = a_{xi} / \alpha$$

$$c_{yi} = c_{xi}$$

ค่าคงที่ α และ β ควรเลือกให้มีความแตกต่างระหว่าง ξ_x และ ξ_y น้อยดังสมการ (Stocking and Lord. 1983:203 – 204)

$$F = \frac{1}{N} \sum_{a=1}^n (\xi_{x_a} - \xi_{y_a})^2$$

เมื่อ N คือ จำนวนผู้สอบ

และแปลงค่าพารามิเตอร์ของแบบทดสอบทั้ง 2 ฉบับ (ฉบับ X และ ฉบับ Y) ให้อยู่ในมาตราเดียวกันในสมการ

1. พารามิเตอร์ของผู้สอบจาก

$$\theta_y = \alpha \theta_x + \beta \quad \text{โดยที่ค่าคงที่ } \alpha \text{ และ } \beta \text{ หาได้จากสมการ}$$

$$\alpha = \sigma_{b_y} / \sigma_{b_x}$$

$$\beta = \mu_{\sigma_y} - \alpha \mu_{\sigma_x}$$

ดังนั้นจะได้สมการดังนี้

$$\theta_y = \frac{\sigma_{\theta_y}}{\sigma_{\theta_x}} \cdot \theta_x + \mu_{\theta_y} - \frac{\sigma_{\theta_y}}{\sigma_{\theta_x}} \cdot \mu_{\theta_x}$$

2. พารามิเตอร์ความยาก $b_y = \alpha b_x + \beta$ จะได้สมการ

$$b_y = \frac{\sigma_{b_y}}{\sigma_{b_x}} \cdot b_x + \mu_{b_y} - \frac{\sigma_{b_y}}{\sigma_{b_x}} \cdot \mu_{b_x}$$

3. พารามิเตอร์อำนาจจำแนก จาก

$$a_y = a_x / \alpha$$

จะได้สมการคือ

$$a_y = a_x \frac{\sigma_{b_x}}{\sigma_{b_y}}$$

4. พารามิเตอร์สัมประสิทธิ์การเดา

$$c_y = c_x$$

2.2.4 การเทียบคะแนนสังเกตโดยใช้ผู้สอบร่วม

ในการเทียบคะแนนโดยวิธีนี้ (Hambleton and Swaminathan.1985:214) สามารถที่จะหาได้โดยวิธีที่ง่ายตรงไปตรงมามีวิธีการเทียบคือถ้าแบบทดสอบฉบับ X และฉบับ Y สอบกับผู้สอบกลุ่ม 2 กลุ่ม แต่ละกลุ่มย่อยได้รับการสอบด้วยแบบทดสอบทั้ง 2 ฉบับ ในลักษณะการจัดลำดับก่อน – หลัง สลับให้เกิดความสมดุล การจัดรูปแบบนี้เพื่อให้การสอบโดยใช้แบบทดสอบทั้ง 2 ฉบับ ได้รับอิทธิพลจากองค์ประกอบในเรื่องการเรียนรู้ ความเมื่อยล้าและการฝึกฝนในลักษณะที่สมดุลกัน ก็มีกระบวนการในการเทียบคะแนนเช่นเดียวกันกับการเทียบคะแนนจริงโดยการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ θ หรือการประมาณคะแนนจริง (ξ) มีความเหมาะสมนั้นคือ คะแนนจริง (ξ) ของมาตราคะแนนดิบ (r) คือ

$$r = \sum_{i=1}^n U_i$$

เมื่อ U_i คือ ผลการตอบข้อสอบข้อที่ i ของผู้สอบคนที่ $1, 2, 3, \dots, n$ และหาค่า

$E(r) = \xi$ ซึ่งหาได้โดย

1. ใช้ความสัมพันธ์ระหว่างคะแนนจริง ξ_x และ ξ_y ของแบบทดสอบ 2 ฉบับ

(ดังกล่าวมา)

2. หาความสัมพันธ์ระหว่างคะแนนดิบ r_x และ r_y แล้วเทียบคะแนนดิบ

ลอร์ด (Lord.1980) ได้ให้ข้อสังเกตว่า ความสัมพันธ์ระหว่าง ξ_x และ ξ_y ไม่มีความจำเป็นจะต้องเหมือนกับคะแนนดิบ r_x และ r_y ซึ่งสามารถหาได้ใน Three-parameter Model คือ

$$\xi_x \geq \sum_{i=1}^n c_i \quad \text{และ} \quad \xi_y \geq \sum_{i=1}^n c_i$$

เมื่อคะแนนสังเกต r_x และ r_y เป็น 0 ดังนั้นคะแนนจริงจึงไม่สามารถเทียบคะแนนของผู้สอบที่ซึ่งมีคะแนนสังเกตต่ำกว่าระดับการเดาได้ ดังนั้น การแปลงคะแนนดิบโดยการใช่วิธีนี้จึงเป็นการหลีกเลี่ยงปัญหาที่ อย่างไรก็ตาม ความสัมพันธ์ระหว่างคะแนนจริง ξ_x และ ξ_y (หรือการประมาณค่าคะแนนจริง ξ_x และ ξ_y) ไม่ควรใช้การเทียบคะแนนดิบ r_x และ r_y

การเทียบคะแนนดิบโดยการใช่วิธีการใช่วิธีการตอบข้อคำถาม ซึ่งจะใช้การแจกแจงของคะแนนสังเกตของแบบทดสอบฉบับ X และฉบับ Y แล้วนำมาเทียบคะแนน (Equate) โดยใช้วิธีอีควิเปอร์เซนไทล์ ในการแจกแจงของคะแนนสังเกตตามทฤษฎี คือ $f(r/\theta)$ ของแบบทดสอบหาได้โดยสูตร

$$\sum_{r=0}^n f(r/\theta) t^r = \prod_{i=1}^n [Q_i(\theta) + t P_i(\theta)]$$

จากสมการดังกล่าว เป็นสมการที่ใช้ในการคำนวณในแต่ละความสามารถ θ จากสูตรจะแสดงวิธีการคำนวณได้ดังนี้ ตัวอย่างเช่น ถ้ามีข้อสอบจำนวน 3 ข้อ

(n = 3) แล้ว ดังนั้นจะได้สมการขวามือและทางซ้ายที่เท่ากันคือ สมการทางขวามือ

$$\prod_{i=1}^3 (Q_i + tP_i) = (Q_1 + tP_1)(Q_2 + tP_2)(Q_3 + tP_3)$$

$$= Q_1 Q_2 Q_3 + t(Q_1 Q_2 P_3 + Q_1 P_2 Q_3 + P_1 Q_2 Q_3) + t^2(Q_1 P_2 P_3 + P_1 Q_2 P_3 + P_1 P_2 Q_3) + t^3 P_1 P_2 P_3$$

สมการทางซ้ายมือ

$$\sum_{r=0}^3 f(r/\theta) t^r = f(r=0/\theta) + t f(r=1/\theta) + t^2 f(r=2/\theta) + t^3 f(r=3/\theta)$$

จากสมการที่กล่าวมาสามารถสรุปได้ดังตาราง แสดงการแจกแจงคะแนนสังเกตของผู้สอบ
ที่มีความสามารถ θ สำหรับข้อสอบ 3 ข้อ

คะแนนดิบ (Raw score)	การแจกแจงความถี่สัมพัทธ์ของคะแนนแบบมีเงื่อนไข (Conditional Relative Frequency : $f(r/\theta)$)
0	$Q_1 Q_2 Q_3$
1	$Q_1 Q_2 P_3 + Q_1 P_2 Q_3 + P_1 Q_2 Q_3$
2	$Q_1 P_2 P_3 + P_1 Q_2 P_3 + P_1 P_2 Q_3$
3	$P_1 P_2 P_3$

สัมประสิทธิ์ t^r , $f(r/\theta)$ คือผลรวมในรูปของ n^c , ซึ่งการคำนวณแบบนี้ก็คือ การคำนวณ
โดยใช้ค่า Maximum likelihood Estimation นั่นเอง เมื่อแทนค่า θ ลงไป ก็สามารถที่จะคำนวณค่า
 $f(r/\theta)$ ได้ เช่นถ้ามีผู้สอบที่มีความสามารถ $\theta_1, \theta_2, \dots, \theta_a, \dots, \theta_n$ และคำนวณหาค่า Marginal
Frequency Distribution $f(r)$ หาจากฟังก์ชัน

$$f(r) = \sum_{a=1}^N f(r/\theta_a)$$

ขั้นตอนการเทียบคะแนนทำได้ดังนี้

1. แปลงค่าพารามิเตอร์ของข้อสอบ และพารามิเตอร์ของผู้สอบให้อยู่ในสเกลเดียวกัน ของ
แบบทดสอบทั้งสองฉบับ
2. หาค่า Conditional Relative Frequency $f_x(r/\theta)$ สำหรับผู้สอบแต่ละคนในกลุ่มเทียบ
คะแนนโดยการประมาณค่าความสามารถของผู้สอบ θ และพารามิเตอร์ของแบบทดสอบฉบับ X
(a_i, b_i, c_i)
3. หาค่า Marginal Frequency Distribution $f_x(r)$ จากสมการที่กำหนดให้ดังกล่าว
4. ทำเช่นเดียวกับขั้นตอนที่ 1 และขั้นตอนที่ 2 ของแบบทดสอบฉบับ Y
5. เทียบคะแนนโดยใช้วิธีอีควิเปอร์เซนไทล์

2.3 ความคลาดเคลื่อนของการเทียบคะแนน

ความคลาดเคลื่อนของการเทียบคะแนน (Standard Error of Equating) ซึ่งหาได้จากสูตร (Macro. 1977 : 157)

$$SEE = \left[\sum_{i=1}^n (Y_{ij} - Y_{io})^2 / N \right]^{\frac{1}{2}}$$

เมื่อ SEE คือ Standard Error of Equating

Y_{ij} คือ คะแนน y (คะแนนจากการสอบ) ของผู้สอบคนที่ i จากวิธีการเทียบคะแนนที่ j ของแบบทดสอบ X

Y_{io} คือ คะแนน y (คะแนนที่เทียบได้) ของผู้สอบคนที่ i จากวิธีการเทียบคะแนนที่ j ที่ได้จากการแปลงคะแนนที่เทียบเท่ากัน

N คือจำนวนผู้สอบที่เป็นกลุ่มเทียบคะแนน

2.4 ความเพี่ยงพอของการเทียบคะแนน

ความเพี่ยงพอของการเทียบคะแนน (Discrepancy Index)

ในการหาความเพี่ยงพอของการเทียบคะแนนในแต่ละวิธีหาได้จากสูตร (Petersen and others. 1982:91)

$$c = \sum (x_i - x_i^*)^2 / nS_x^2$$

เมื่อ c คือ ค่าความเพี่ยงพอของการเทียบคะแนน

x_i คือ คะแนน x ของผู้สอบคนที่ i จากแบบทดสอบชุด X

x_i^* คือ คะแนน x ของผู้สอบคนที่ i จากการแปลงคะแนน y

n คือ จำนวนคนในกลุ่มตัวอย่างสอบทานผล

S_x^2 คือค่าความแปรปรวนของคะแนนจากแบบทดสอบ X

ซึ่งมีเกณฑ์ในการพิจารณาดังนี้ (Petersen and others. 1982 : 93-94)

ระดับน่าพอใจอย่างมาก	หมายถึง	
$c \leq (.05SD_x)^2$		
ระดับน่าพอใจ	หมายถึง	$(.05SD_x)^2 < c \leq (.10SD_x)^2$
ระดับปานกลาง	หมายถึง	$(.10SD_x)^2 < c \leq (.15SD_x)^2$
ระดับไม่น่าพอใจ	หมายถึง	$(.15SD_x)^2 < c \leq (.20SD_x)^2$
ระดับไม่น่าพอใจอย่างยิ่ง	หมายถึง	$c > (.20SD_x)^2$

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนน

งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนนตามทฤษฎีดั้งเดิมและตามทฤษฎีการตอบข้อคำถาม ได้มีการศึกษาที่หลากหลาย ในที่นี้พอจะสรุปได้ดังต่อไปนี้

งานวิจัยต่างประเทศ

การศึกษางานวิจัยเกี่ยวกับการเทียบคะแนนในต่างประเทศ ผู้วิจัยรวบรวมและสรุปได้ดังนี้
 ลอร์ด (Lord. 1977 : 117 - 138) มาร์โค (Marco. 1977 : 139 - 160) ได้เปรียบเทียบวิธีการเทียบ
 มาตรฐานระหว่างวิธีดั้งเดิม (Conventional) และวิธีการใช้ทฤษฎีการตอบข้อคำถาม
 ผลปรากฏว่าผลที่ได้แตกต่างกัน ซึ่งสอดคล้องกับการศึกษาของมาร์โค ปีเตอร์เซ็นและสตีเวอร์ท (Kolen.
 1981 : citing Marco, Petersen and Stewart.n.d.) ที่ได้ศึกษาเปรียบเทียบมาตรฐานเชิงเส้นตรง
 วิธีอีคิวเปอร์เซ็นไทล์ และวิธีการใช้โด่งลักษณะเฉพาะข้อสอบ โดยใช้แบบทดสอบ SAT แบ่งแบบทดสอบ
 โดยใช้แบบทดสอบรวม ผลการศึกษาพบว่า แบบทดสอบที่มีค่าความยากแตกต่างกัน วิธีการเทียบคะแนน
 โดยใช้โด่งลักษณะเฉพาะข้อสอบให้ผลดีที่สุด ส่วนวิธีเทียบแบบเส้นตรงได้ผลน้อยที่สุด

โคเลน และวิทนี (Kolen and Whitney.1982: 279-294) ได้เปรียบเทียบความเพียงพอของวิธี
 การเทียบคะแนน 4 วิธี คือ วิธีอีคิวเปอร์เซ็นไทล์ วิธีเชิงเส้นตรง วิธีการใช้ทฤษฎีการตอบข้อคำถามชนิด
 หนึ่งพารามิเตอร์โลจิสติกโมเดล หรือราล์ชโมเดลและชนิดสามพารามิเตอร์โลจิสติกโมเดล โดยใช้แบบ
 ทดสอบ GED (Test of General Educational Development) ซึ่งเป็นแบบทดสอบมาตรฐานวัดผลสัมฤทธิ์
 ทางเรียนจำนวน 5 วิชา แต่ละวิชามีเนื้อหาแตกต่างกันดังนี้ ทักษะการเขียน (80 ข้อ) การศึกษาสังคม
 (60 ข้อ) วิทยาศาสตร์ (60 ข้อ) ทักษะการอ่าน (40 ข้อ) และคณิตศาสตร์ (50 ข้อ) ซึ่งใช้สำหรับตัดสินให้
 ประกาศนียบัตรแก่นักเรียนระดับอุดมศึกษาทั่วประเทศที่มีความรู้เทียบเท่าระดับที่แสดงว่า ได้เรียนสำเร็จแล้ว
 แบบทดสอบแต่ละวิชามี 12 ชุด ดำเนินการสอบผู้สอบกลุ่มละ 2 ชุด และทุกกลุ่มต้องสอบแบบทดสอบชุดที่
 12 ซึ่งแยกออกมาเป็นแบบทดสอบรวม การจัดลำดับแบบทดสอบให้แต่ละกลุ่มใช้วิธีกลุ่มสมดุล
 (Counterbalance) ทำการเทียบคะแนนแต่ละคู่ของแบบทดสอบในแนวนอนด้วยวิธีการเทียบ 4 วิธีดังกล่าว
 กลุ่มตัวอย่างสุ่มจากการสอบในปี 1980 ซึ่งมีจำนวน 800,000 คน โดยสุ่มมาขนาดเล็ก กลุ่มละประมาณ
 200 คน คิดเป็นสัดส่วนโรงเรียนละ 22 คน ด้วยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling)
 ตามประเภทโรงเรียน สภาพทางภูมิศาสตร์และสถานภาพทางสังคม กลุ่มทวนผลก็ใช้วิธีการสุ่มเช่นเดียวกัน
 โดยสุ่มมาจากผู้สอบแต่ละวิชา และแต่ละชุดให้ผู้สอบ 20 คน (ประมาณร้อยละ 10 ของผู้เข้าสอบทั้งหมด)
 การเทียบคะแนนเชิงเส้นตรงและอีคิวเปอร์เซ็นไทล์ กระทำตามวิธีของแองกอฟฟ์ ส่วนวิธีการใช้ทฤษฎีการ
 ตอบข้อคำถามประมาณค่าพารามิเตอร์โดยใช้โปรแกรมโลจิสติกเฉพาะแบบทดสอบรวมก่อน จากนั้นใช้
 ค่าพารามิเตอร์ความสามารถของผู้สอบที่กำหนดให้คงที่ ในการประมาณค่าพารามิเตอร์จากแบบทดสอบ อีก
 11 ชุด แล้วสร้างตารางคะแนนสมมูลระหว่างแบบทดสอบทั้งหมดโดยใช้ประมาณค่าคะแนนจริงส่วนคะแนนที่
 อยู่ต่ำกว่าระดับการเดาตามรูปแบบสามพารามิเตอร์นั้นอาศัยวิธีการเชิงเส้นตรงมาประมาณค่าเพิ่มเติมเกณฑ์
 ในการตรวจสอบความเพียงพอของวิธีการเทียบคะแนนใช้ค่าสถิติจากกลุ่มทวนผลซึ่งผู้สอบแต่ละคนจะต้อง
 สอบแบบทดสอบทั้งสองชุดที่นำมาเทียบ ดัชนีเปรียบเทียบเปอร์เซ็นไทล์ซึ่งเป็นค่าเฉลี่ยกำลังสองของความ
 แตกต่างระหว่าง คะแนน X ที่ได้จากแบบทดสอบรวมกับคะแนน Y ที่แปลงไปสู่สเกลของคะแนน X โดยใช้วิธี

เทียบแต่ละวิธี และทดสอบความมีนัยสำคัญของค่าอันดับจากการแปลงคะแนนตามปริมาณของค่าดัชนีด้วยการทดสอบฟรายด์แมนผลการวิจัยพบว่า วิธีการเทียบคะแนนอิกวิเปอร์เซ็นไทล์ให้ผลการเทียบที่คงเส้นคงวาน้อยกว่าการเทียบคะแนนเชิงเส้นตรง ส่วนการเทียบคะแนนตามทฤษฎีการตอบข้อคำถามชนิดหนึ่งพารามิเตอร์หรือราส์โมเดลให้ผลคงที่เหมือนกับวิธีเทียบเชิงเส้นตรง ขณะที่วิธีเทียบตามทฤษฎีการตอบข้อคำถามชนิดสามพารามิเตอร์ และวิธีอิกวิเปอร์เซ็นไทล์ให้ผลน้อยกว่า ดังนั้นการศึกษานี้จึงให้ผลการเทียบแตกต่างจากการศึกษาอื่น ได้แก่การศึกษาของโคเลน (Kolen, 1981: 1-11) ที่พบว่า วิธีการเทียบคะแนนแบบอิกวิเปอร์เซ็นไทล์ให้ผลมากที่สุดเมื่อเปรียบเทียบกับวิธีเชิงเส้นตรงและทฤษฎีการตอบข้อคำถามชนิดหนึ่งพารามิเตอร์ ทั้งนี้การศึกษาของโคเลนใช้กลุ่มตัวอย่างขนาดใหญ่ซึ่งแตกต่างจากการศึกษาคั้งนี้ที่ใช้กลุ่มตัวอย่างขนาดเล็ก นอกจากนี้แบบทดสอบที่ใช้เทียบในการศึกษาคั้งนี้มีความยากใกล้เคียงกันน้อยกว่า ส่วนผลการเปรียบเทียบของคนอื่นๆ (Kolen and Whitney, 1982 : 291 ; Petersen, Cook and Stocking . 1981) ซึ่งต่างพบว่าเมื่อแบบทดสอบที่เทียบโดยใช้แบบทดสอบร่วมกระทำกับกลุ่มที่ไม่ได้มาจากการสุ่มซึ่งมีขนาดใหญ่นั้นวิธีการใช้ทฤษฎีการตอบข้อคำถามชนิดสามพารามิเตอร์ จะให้ผลการเทียบที่มีความคล้ายคลึงกันมาก ผลการศึกษาที่แตกต่างกันเหล่านี้ โคเลนและวิทนีได้สรุปว่า วิธีการเทียบที่มีความเพียงพอขึ้นกับองค์ประกอบหลายประการ ได้แก่ คุณสมบัติต่างๆของแบบทดสอบ รูปแบบที่ใช้ในการเทียบ และลักษณะของกลุ่มตัวอย่างที่ใช้เทียบ เป็นต้น

โคเลน(Kolen, 1981: 1-11) ได้เปรียบเทียบผลของการเทียบคะแนนระหว่างรูปแบบดั้งเดิมสองวิธีกับรูปแบบ IRT เจ็ดวิธี โดยใช้ข้อมูลจากผลการสอบนักเรียนในรัฐไอโอวา ปี 1978 ตามโครงการชื่อ The Iowa Test of Educational Development แบบทดสอบที่ใช้คือ แบบทดสอบที่พิมพ์ใหม่ครั้งที่ 7 เทียบไปสู่แบบทดสอบชุดเก่าที่พิมพ์ครั้งที่ 6 ซึ่งมีหนึ่งชุดคือ x_6 ใช้สอบกับนักเรียนทั้งสองระดับ คือ ระดับ 1 และระดับ 2 ซึ่งระดับ 1 เป็นนักเรียนเกรด 9 และ 10 ส่วนระดับ 2 เป็นนักเรียนเกรด 11 และ 12 ส่วนแบบทดสอบที่พิมพ์ครั้งที่ 7 มี 2 ชุด ที่คู่ขนานกันคือ x_7 และ y_7 แต่ละชุดแบ่งออกเป็น 2 ระดับ ซึ่งมีความยากต่างกันเพื่อแยกสอบนักเรียนแต่ละคน มีแบบแผนในการรวบรวมข้อมูลดังนี้ นักเรียนระดับ 1 สอบแบบทดสอบ x_6 , x_7 ระดับ 1 และ y_7 ระดับ 2 ทั้งนี้ทั้งนักเรียนแต่ละคนจะได้รับแบบทดสอบเพียง 1 ชุด คือ ชุด x_6 , x_7 หรือ y_7 โดยการสุ่ม ดังนั้นจำนวนนักเรียนที่สอบแบบทดสอบแต่ละชุดจะมี 1 ใน 3 ของจำนวนนักเรียนในแต่ละระดับ ส่วนแบบทดสอบ x_6 จะมีนักเรียน 1 ใน 3 ของจำนวนนักเรียนรวมสองระดับ สำหรับเกณฑ์ในการเปรียบเทียบผลจะใช้ค่าดัชนีจากกลุ่มทวนผลซึ่งเป็นนักเรียนทุกๆคนที่ 3 จากการสอบแบบทดสอบแต่ละชุดและ แต่ละระดับการเทียบคะแนนแบบดั้งเดิมใช้วิธีการเชิงเส้นตรงและวิธีอิกวิเปอร์เซ็นไทล์ซึ่งจัดกระทำตามวิธีของแองกอฟฟ์ ส่วนรูปแบบของทฤษฎีการตอบข้อคำถามใช้ หนึ่ง สอง และสามพารามิเตอร์โลจิสติกโมเดล โดยแต่ละโมเดลเทียบ 2 แบบ คือ เทียบโดยการประมาณค่าคะแนนจริง วิธีเทียบคะแนนสังเกต และวิธีราส์โมเดล ดัชนีการทวนผล โคเลนใช้วิธีคำนวณจากค่าเฉลี่ยกำลังสองของความแตกต่าง (สำหรับผู้สอบแบบทดสอบ x_6 ในการทวนผล) ระหว่างคะแนนจากแบบทดสอบ x_6 กับคะแนนจากแบบทดสอบ x_7 และ y_7 ที่แปลงแล้ว ณ ตำแหน่ง เปอร์เซ็นไทล์เดียวกัน ทั้งนี้โดยการใช้วิธีการเทียบคะแนนแบบต่างๆเท่าวิธี หากดัชนีจากกลุ่มทวนผลที่ได้จากวิธีการเทียบใดมีค่าน้อยกว่าแสดงว่าวิธีนั้นมีความสอดคล้องระหว่างการแจกแจงคะแนนของแบบทดสอบชุดเก่ากับคะแนนของแบบทดสอบชุดใหม่ที่แปลงไปยังสเกลของชุดเก่ามากกว่า นั่นคือ วิธีนั้นมีความคงเส้นคงวามากกว่า ทั้งนี้ โคเลนทดสอบความ

แตกต่างของดัชนีที่ได้จากการเทียบคะแนนแบบต่างๆ เพื่อเปรียบเทียบผลของแต่ละวิธี โดยใช้สถิติทดสอบของฟรายด์แมน (Friedman Test) ผลการวิจัย พบว่า วิธีการเทียบคะแนนแบบต่างๆ ให้ผลแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05 เมื่อใช้ แบบทดสอบระดับ 1 และ 2 ตามลำดับ ถ้าพิจารณาแบบทดสอบระดับที่ 1 พบว่า วิธีที่ให้ผลคงเส้นคงวามากที่สุดคือ วิธีการตามทฤษฎีการตอบข้อคำถามแบบสามพารามิเตอร์ที่ประมาณค่าคะแนนสังเกต ส่วนวิธีการที่ให้ผลคงเส้นคงวาน้อยที่สุดคือ วิธีการเชิงเส้นตรง ถัดมาคือวิธีการใช้ทฤษฎีการตอบข้อสอบชนิดหนึ่งพารามิเตอร์ สำหรับแบบทดสอบระดับที่ 2 พบว่าวิธีการที่ให้ผลคงเส้นคงวามากที่สุด คือวิธีการใช้ทฤษฎีการตอบข้อคำถามแบบสามพารามิเตอร์ที่ประมาณค่าคะแนนจริง วิธีของราล์สให้ผลคงเส้นคงวามากกว่าวิธีตามทฤษฎีการตอบข้อคำถามชนิดหนึ่งพารามิเตอร์ที่ประมาณค่าคะแนนจริง ซึ่งโคเลนอภิปรายโดยส่วนรวมว่า วิธีตามทฤษฎีการตอบข้อคำถามชนิดหนึ่งพารามิเตอร์ให้ผลการเทียบที่ไม่เพียงพอ ทั้งนี้อาจเนื่องมาจากการเดาของผู้สอบมากยกเว้นเมื่อขจัดการเดาของผู้สอบโดยใช้ วิธีของไรท์และสโตน (Wright and Stone, 1979) ซึ่งตัดข้อสอบที่พบว่าไม่เหมาะสมกับราล์สโมเดลออกจากผลการสอบ ผลการวิจัยนี้สนับสนุนงานวิจัยของสไลด์และลินน์ (Slinde and Linn, 1978: 23-25; Slinde and Linn, 1979 437-452) ที่พบว่า การเทียบคะแนนในแนวตั้งโดยใช้แบบทดสอบร่วมที่มีความยากต่างกันสอบกับนักเรียนต่างระดับชั้นตามวิธีของราล์สโมเดลให้ผลที่ไม่เพียงพอ

ลอร์ด (Lord, 1980: 203) ได้ศึกษาทฤษฎีการเทียบคะแนนของแบบทดสอบสองชุด ด้วยวิธีการเทียบคะแนนจริงและคะแนนสังเกต พบว่ามีข้อได้เปรียบและเสียเปรียบ วิธีการเทียบคะแนนจริงให้ความหมายของคะแนนสมมูลที่อยู่เหนือค่าเฉลี่ยการเดา ส่วนวิธีการประมาณค่าคะแนนสังเกต ก็มีจุดอ่อนที่เป็นเพียงการเทียบมาตราอย่างประมาณเท่านั้น ต่อข้อคำถามที่ว่า วิธีใดจะให้ผลการเทียบที่ดีกว่านั้น ลอร์ดกล่าวว่า ทั้งสองวิธีมีความสอดคล้องกันมาก แต่การสรุปอ้างอิงทั่วไปต้องทำอย่างพิถีพิถัน

ลอร์ด และวิงเกอร์สกี (Lord and Wingersky, 1984: 453-461) ได้เปรียบเทียบการเทียบคะแนนระหว่างรูปแบบการเทียบคะแนนจริงโดยการใช้ทฤษฎีการตอบข้อคำถามจากแบบทดสอบ SAT ด้านภาษา จำนวน 6 ชุด เทียบเข้าสู่สเกลเดียวกันแบบสายโซ่โดยใช้แบบทดสอบร่วม ทั้งนี้แบบทดสอบชุดแรกและชุดสุดท้ายเป็นชุดเดียวกัน ขั้นตอนการเทียบคะแนนเริ่มจาก แบบทดสอบ v_1 เทียบไปสู่แบบทดสอบ x_2 โดยใช้แบบทดสอบร่วม fe ส่วนแบบทดสอบ x_2 เทียบไปสู่ y_3 โดยใช้แบบทดสอบร่วม fm ซึ่งจะใช้ผลการเทียบจาก v_4 ไป y_3 กระทำการเทียบคะแนนในลักษณะเดียวกันไปยังแบบทดสอบอื่น สุดท้ายจะได้ผลการเทียบ v_4 กับ v_1 ซึ่งเป็นการเทียบแบบทดสอบกับตัวเอง และใช้เป็นเกณฑ์ในการเปรียบเทียบความคงเส้นคงวาของวิธีการเทียบคะแนนทั้งสองวิธี ทั้งนี้เกณฑ์นี้แตกต่างจากที่ โคเลน (Kolen, 1981 : 1-11) ได้เคยศึกษาความคงเส้นคงวาของวิธีการเทียบคะแนนแบบต่างๆ โคเลนใช้เกณฑ์จากกลุ่มทานผลซึ่งเป็นกลุ่มอิสระอีกกลุ่มหนึ่ง โดยผู้วิจัยกล่าวไว้ว่าเกณฑ์ที่โคเลนใช้ไม่ใช่เกณฑ์ที่มีคุณสมบัติสำหรับการเลือกวิธีการเทียบคะแนนที่ดีที่สุด เพราะวิธีการเทียบคะแนนที่ไม่ถูกต้องอาจให้ผลการเทียบคงเส้นคงวามากกว่าวิธีการที่ถูกต้อง เกณฑ์การเทียบแบบทดสอบกับตัวเองที่ใช้ในการศึกษานี้ อาจไม่เที่ยงตรงในทางทฤษฎี และไม่น่าพอใจในการปฏิบัติสำหรับการเทียบแบบทดสอบสองฉบับที่แตกต่างกัน แต่เป็นเกณฑ์ที่มีประโยชน์ต่อการเปลี่ยนวิธีการเทียบคะแนนใหม่หากพบว่าคะแนนที่แปลงแล้วนั้นในขั้นสุดท้ายไม่เป็นอย่างเดียวกันกับคะแนนเดิมของตัวเอง ผลการวิจัยยังไม่มีมติเห็นชอบทางทฤษฎีที่ชัดเจนในการเลือกใช้การเทียบคะแนนแบบ IRT ที่ใช้คะแนนจริงกับคะแนนสังเกตโดยวิธีอควิเปอร์เซ็นไทล์ เนื่องจากทั้งสองวิธีมีความสอดคล้องกัน

ฮาน ,โคเลน และ โพลแมนน์ (Han, Kolen and Pohlmann, 1997:105 - 121) ได้ศึกษาเปรียบเทียบผลของการเทียบคะแนน 3 วิธี คือการเทียบคะแนนจริง คะแนนสังเกตตามทฤษฎีการตอบข้อคำถาม

และวิธีอีคิวเปอร์เซนไทล์ โดยใช้แบบทดสอบของ ACT (American College Testing) และโปรแกรมการทดสอบและการประเมินผล (ACT Program.1989) ประกอบด้วยวิชาคณิตศาสตร์และวิชาภาษาอังกฤษ จำนวน 7 ฟอรม์ โดยแบบทดสอบคณิตศาสตร์เป็นข้อสอบแบบเลือกตอบจำนวน 60 ข้อๆละ 5 ตัวเลือก ส่วนข้อสอบวิชาภาษาอังกฤษแบบเลือกตอบ จำนวน 75 ข้อๆ ละ 4 ตัวเลือก โดยมีการเรียงสลับกันดังนี้ คือ ฟอรม์ A,B,C,D,E,F และG โดยการสุ่ม กลุ่มตัวอย่างที่ใช้ในการศึกษาค้างนี้เป็นักเรียนระดับมัธยมศึกษา จำนวน 22,848 คน ซึ่งในแต่ละฟอรม์จะใช้นักเรียนประมาณ 3000 คน ในการศึกษาครั้งนี้ได้แบ่งครึ่งกลุ่ม โดยการสุ่มแบ่งครึ่งในแต่ละกลุ่ม กลุ่มต่างๆเหล่านี้จะทำข้อสอบฟอรม์ที่เหมือนกัน และทั้งหมดจะเทียบคะแนน 14 รูปแบบในแต่ละวิธี ใช้การวิเคราะห์โดยเปรียบเทียบความสอดคล้องของผลการเทียบคะแนนทั้ง 3 วิธี มีการคำนวณดัชนีความยากในข้อสอบทั้ง 7 ฟอรม์ ในแต่ละฟอรม์จะมีรูปแบบความยาก 3 รูปแบบ คือ รูปแบบที่มีความยากมาก,ค่าความยากปานกลาง และค่าความยากน้อย จากการเปรียบเทียบความเพียงพอของการเทียบคะแนนและค่าความยากแล้ว ผลการวิจัยพบว่า ผลของวิธีการเทียบคะแนนสังเกตตามทฤษฎีการตอบข้อคำถาม ในการแจกแจงคะแนนสังเกตกับค่าความเบี่ยงเบนมาตรฐาน คะแนนเฉลี่ยที่มากจากการแจกแจงคะแนนดิบมีลักษณะเบ้ทางบวก และคะแนนเฉลี่ยน้อยกว่าของการแจกแจงคะแนนดิบมีลักษณะเบ้ทางลบ ความคลาดเคลื่อนจากการเทียบคะแนนเกิดขึ้นจากการผันแปรของการเทียบคะแนนวิธีอีคิวเปอร์เซนไทล์ที่ยังไม่ปรับให้เรียบ ผลจากการเทียบคะแนนจริงตามทฤษฎีการตอบข้อคำถามมีความคงเส้นคงวามากกว่าอีก 2 วิธี แต่ทว่า ผลต่างของค่าเฉลี่ยของความคงเส้นคงวาในการเทียบคะแนนระหว่าง 3 วิธีมีค่าน้อยมาก และโดยเฉพาะผลต่างของค่าเฉลี่ยของความคงเส้นคงวาในการเทียบคะแนนระหว่างคะแนนจริงและคะแนนสังเกตตามทฤษฎีการตอบข้อคำถามจะแตกต่างกันอย่างไม่มีนัยสำคัญ วิธีการเทียบคะแนนสังเกตมีความคงเส้นคงวามากกว่าวิธีอีคิวเปอร์เซนไทล์ และดูเหมือนว่าความแตกต่างในการเทียบคะแนนจะมีค่ามากกว่าระหว่าง 2 ใน 3 วิธีที่เกิดขึ้นเมื่อความแตกต่างของความยากระหว่างแบบทดสอบทั้ง 2 ฟอรม์มีค่าแตกต่างกันมาก

หยาง (Yang ,Wen-Ling. 1997) ได้ศึกษาผลของเนื้อหาเดียวกันและความแม่นยำของวิธีการเทียบคะแนนที่ใช้แบบทดสอบรวม เครื่องมือที่ใช้เป็นแบบทดสอบ 2 ฟอรม์ และมีชุดของข้อสอบรวมด้วยในแต่ละฟอรม์ ใช้วิธีการเทียบแบบเส้นตรง,อีคิวเปอร์เซนไทล์ และวิธีจาก IRT ทั้งสองแบบ เกณฑ์ที่ใช้ในการประเมินความแม่นยำของผลการเทียบเหล่านี้เป็นเกณฑ์อื่น 2 เกณฑ์ และใช้ความพิถีพิถันในการทำให้เกิดความเที่ยงตรงและมีติของข้อสอบที่สัมพันธ์กับการเทียบคะแนน พบว่า วิธีเทียบคะแนนทั้งหมดให้ผลที่มีความเที่ยงตรงในระดับปานกลาง ทุกวิธีให้ความเที่ยงตรงเมื่อมีข้อสอบรวมเป็นตัวแทนของข้อสอบทั้งหมดหรือเมื่อข้อสอบในแบบทดสอบมีเนื้อหาเดียวกัน เพราะฉะนั้น เพื่อเพิ่มความเที่ยงตรงในการเทียบคะแนนขอเสนอแนะว่า รวมข้อสอบรวมที่ให้ผลอย่างเต็มที่ในส่วนของเนื้อหาแบบทดสอบ และพบอีกว่าผลการเทียบโดยใช้ IRT มีความเที่ยงตรงกว่าวิธีอื่นๆ

สวนทอง (Suantong ,Surintorn .1998) ได้ศึกษาสำรวจปัจจัยที่มีผลต่อการเทียบคะแนนแบบทดสอบใน IRT โดยสำรวจปัจจัยทั้ง 5 ที่สามารถส่งผลต่อการเทียบคะแนนจากแบบทดสอบ 2 ฉบับ ไปยังมาตราคะแนนรวม ปัจจัยทั้ง 5 ได้แก่ 1) ชนิดที่มีการแจกแจง(ได้แก่ ชนิดธรรมดาเปรียบเทียบกันเอง) 2) ค่าเบี่ยงเบนมาตรฐานของค่าความยากของข้อสอบ (ได้แก่ .68 ,.95 ,.99) 3)ความยาวของแบบทดสอบหรือจำนวนข้อสอบ (ได้แก่ 50,100,200) 4)จำนวนข้อสอบรวม (ได้แก่ 10,20,30) และ 5) จำนวนกลุ่มตัวอย่าง (100,300,500) โดยใช้ โปรแกรมวิเคราะห์ SIMTESTและ BIGSTEPS สำหรับการจำลองข้อมูลในการเทียบของชุดข้อมูลของข้อสอบ จำนวน 4,860 ข้อ และใช้โปรแกรม QUICKBASIC ในการปฏิบัติการโดยอัตโนมัติ

ผลที่เกิดขึ้น พบว่า ปัจจัยจำเพาะ 5 ทางของความแปรปรวน แสดงให้เห็นว่า มี 3 คู่ที่มีผลการปฏิสัมพันธ์ 2 ทางที่มีนัยสำคัญ ได้แก่ ความสัมพันธ์ของความยาวของข้อสอบรวมและความยาวของข้อสอบ , ค่าความเบี่ยงเบนมาตรฐานของความยากของข้อสอบและชนิดของการแจกแจง , และค่าความเบี่ยงเบนมาตรฐานของค่าความยากและขนาดของกลุ่มตัวอย่าง สรุปว่า จำนวนข้อสอบรวมมีส่งผลมากที่สุด ในการเทียบคะแนนเมื่อปฏิสัมพันธ์กับความยาวข้อสอบ การก่อให้เกิดปัจจัยที่เหมาะสมที่สุด ที่ควรพิจารณา เมื่อเทียบคะแนน 2 ฉบับ ปรากฏว่าเป็นการแจกแจงที่มีรูปแบบและค่าความเบี่ยงเบนมาตรฐานของค่าความยาก .99 และความยาวข้อสอบ 100 ข้อ ข้อสอบรวม 20 ข้อ และกลุ่มตัวอย่าง 100 คน

คาสโควิตซ์(Kaskowitz ,Gary Scott .1998) ศึกษาผลของความคลาดเคลื่อนของการประมาณค่าพารามิเตอร์และการเทียบคะแนนด้วยทฤษฎีการตอบข้อคำถาม โดยใช้ โค้งลักษณะเฉพาะของข้อสอบ เป็นการศึกษาจำลองในการเทียบคะแนนในแนวนอนที่มีข้อสอบรวมภายใน จำนวน (5,15 หรือ 25) โมเดล IRT (2PL,3PL) และความสัมพันธ์ของความคลาดเคลื่อนมาตรฐานกับการกะประมาณ (ชนิดสัมพันธ์กับความคลาดเคลื่อนและไม่สัมพันธ์) ศึกษาทั้งหมด 37 กลุ่มย่อย ผลการศึกษาพบว่า ทั้งพารามิเตอร์ที่สัมพันธ์และไม่สัมพันธ์ ใช้ข้อสอบรวม 15 หรือ 25 ข้อ ให้ผลดีกว่า 5 ข้อ ซึ่งแสดงให้เห็นการก่อให้เกิดผลดีด้วยผลต่างจากเดิมมากระหว่าง TCC ซึ่งส่งผลต่อค่า IR ที่แสดงความสำเร็จที่เพิ่มดังที่เงื่อนไขของความคลาดเคลื่อนในการเทียบคะแนนจริงคงที่ ในการใช้ ฟังก์ชัน ICC และชุดพารามิเตอร์(ชนิดสัมพันธ์กับความคลาดเคลื่อนและไม่สัมพันธ์) ซึ่งแสดงให้เห็นว่าเมื่อมีความคลาดเคลื่อนในการกะประมาณค่าพารามิเตอร์ วิธีเทียบด้วย ICC ก็ยังคงมีความแข็งแกร่ง

จากงานวิจัยที่กล่าวมาพอสรุปประเด็นสำคัญได้ดังนี้

1. การเทียบคะแนนโดยใช้ทฤษฎีการตอบข้อคำถาม (Item Response Theory) มีผลการเทียบคะแนนดีกว่าวิธีการเทียบคะแนนตามทฤษฎีคลาสสิกอล (Classical Test Theory) แต่ก็ไม่ทุกสถานการณ์
2. การเทียบคะแนนโดยโลจิสติก 3 พารามิเตอร์ จะต้องใช้จำนวนกลุ่มตัวอย่างที่มีจำนวนมาก ถ้าในกรณีของกลุ่มตัวอย่างขนาดเล็ก จะทำให้เกิดความผันแปรเชิงสุ่มจนทำให้ผลการเทียบคะแนนขาดความแน่นอนและแม่นยำไป
3. ไม่มีวิธีการใดที่สามารถให้ผลในการเทียบคะแนนได้ดีทุกสถานการณ์

งานวิจัยในประเทศ

สำหรับงานวิจัยในประเทศที่เกี่ยวกับการเทียบคะแนน ผู้วิจัยได้สรุปรวบรวมไว้ดังนี้

ทิพาภรณ์ ทุงฤทธิ์ (2544: บทคัดย่อ) ได้เปรียบเทียบผลการเทียบคะแนนโดยวิธีเชิงเส้นตรงเมื่อใช้วิธีการแปลงคะแนนที่แตกต่างกันในแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2542 สังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี จำนวน 820 คน ซึ่งได้มาจากการสุ่มแบบแบ่งชั้น ผลการศึกษาพบว่า 1) คะแนนที่เทียบได้จากการเทียบคะแนน 3 วิธี แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 2) ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน 3 วิธีแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ 3) ความเพียงพอของการเทียบคะแนน 3 วิธีแตกต่างกัน

พรโสภา พินิจผล (2539 : 64-67) ได้เปรียบเทียบการเทียบคะแนนวิชาการพยาบาลสูติศาสตร์ โดยใช้วิธีการเทียบคะแนนตามทฤษฎีการตอบข้อคำถาม 2 วิธี คือการเทียบคะแนนดิบโดยใช้แบบทดสอบร่วมและเทียบคะแนนจริงโดยใช้แบบทดสอบร่วม กลุ่มตัวอย่างเป็นนักศึกษาพยาบาลศาสตร์ (ระดับต้น) ชั้นปีที่ 2 ปีการศึกษา 2538 ของวิทยาลัยพยาบาลสังกัดสถาบันพัฒนากำลังคนด้านสาธารณสุข จำนวน 980 คน ผลการศึกษาพบว่า ความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้งสองวิธีไม่แตกต่างกัน ส่วนความเพียงพอของการเทียบคะแนนพบว่า วิธีการเทียบคะแนนจริงมีความเพียงพอสูงกว่าการเทียบคะแนนดิบ แต่เมื่อเปรียบเทียบกับเกณฑ์ของการประเมินความเพียงพอ พบว่า ทั้งสองวิธีอยู่ในเกณฑ์ที่น่าพอใจอย่างยิ่ง

พรพิมล นาคเวช (2537 : 99 - 101) ศึกษาคุณภาพของการเทียบคะแนนในแนวตั้งโดยใช้ทฤษฎีการตอบข้อคำถาม สำหรับผู้มีความสามารถต่างกัน 3 ระดับ ซึ่งได้แก่ นักเรียนชั้น ม.1 ม.2 และ ม.3 โดยวิเคราะห์ความคลาดเคลื่อนมาตรฐานและความเพียงพอของการเทียบคะแนน กลุ่มตัวอย่างเป็นนักเรียนชั้น ม.1 ม.2 และ ม.3 สังกัดกองการมัธยมศึกษา กรมสามัญศึกษา จำนวน 6,429 คน ที่ได้จากการสุ่มกลุ่มตัวอย่างเทียบคะแนนเพื่อนำผลไปเทียบคะแนนระหว่างแบบทดสอบ จำนวน 3,279 คน และกลุ่มตัวอย่างสอบทานผลเพื่อนำผลไปตรวจสอบคุณภาพของการเทียบคะแนน จำนวน 3,105 คน เครื่องมือที่ใช้เป็นแบบทดสอบคณิตศาสตร์ เรื่องสมการและอสมการ สำหรับระดับชั้น ม.1 ม.2 และ ม.3 รวม 3 ฉบับ ๆ ละ 25 ข้อ โดยมีข้อสอบร่วมระหว่างระดับชั้นที่ติดกันระดับชั้นละ 5 ข้อ ผลการศึกษาพบว่า ความคลาดเคลื่อนมาตรฐานที่ค่อนข้างต่ำและผลการวิเคราะห์ความเพียงพอของการเทียบคะแนนมีค่าอยู่ระหว่าง .031 - .051 ซึ่งเป็นความคลาดเคลื่อนมาตรฐานที่ค่อนข้างต่ำและผลการวิเคราะห์ความเพียงพอของการเทียบคะแนนตามแนวตั้งพบว่า ค่าดัชนีความแตกต่างของการเทียบคะแนนอยู่ในระดับที่น่าพอใจ

วรรณดี แสงประทีปทอง (2538 : 113 - 123) ศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทย ของนักเรียนชั้นมัธยมศึกษาตอนต้น โดยใช้วิธีการเทียบคะแนนตามแนวตั้งตามทฤษฎีการตอบข้อคำถาม แบบสามพารามิเตอร์โดยศึกษาความก้าวหน้าทั้งในกลุ่มรวมและในกลุ่มย่อยตามระดับความสามารถของนักเรียน กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาตอนต้นของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษาในจังหวัดสิงห์บุรี ภาคเรียนที่ 1 ปีการศึกษา 2537 กลุ่มตัวอย่างแบ่งออกเป็น 2 กลุ่ม คือกลุ่มตัวอย่างที่ใช้ในการเทียบคะแนนเป็นนักเรียนชั้น ม.1 จำนวน 1,245 คน ม.2 จำนวน 1,280 คน และ ม.3 จำนวน 1,171 คน กลุ่มทานผล คือนักเรียนชั้นม.1 จำนวน 373 คน ม.2 จำนวน 384 คน และม.3 จำนวน 352 คน เครื่องมือที่ใช้ในการวิจัยคือแบบทดสอบเทียบคะแนน 3 ฉบับ คือ ท101 ท203 และ ท305 และแบบทดสอบร่วมซึ่งเป็นแบบทดสอบวัดความสามารถทั่วไป 1 ฉบับ เก็บรวบรวมข้อมูลจากการสอบจริงปลายภาคเรียนที่ 1 ปีการศึกษา 2537 พร้อมกันทุกระดับชั้น การวิเคราะห์ข้อมูลใช้แบบจำลองสามพารามิเตอร์ทฤษฎีการตอบข้อคำถาม เพื่อเปรียบเทียบการเทียบคะแนนต่างระดับชั้น และเปรียบเทียบผลสัมฤทธิ์ทางการเรียนแต่ละระดับชั้นภายหลังจากการเทียบคะแนน โดยเปรียบเทียบในกลุ่มรวมและกลุ่มย่อยตามระดับความสามารถ ผลการศึกษา พบว่า ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่า ชั้น ม.1 และ ม.2 ตามลำดับ แต่ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.1 และ ม.2 ไม่แตกต่างกันในกลุ่มนักเรียนที่มีความสามารถสูง ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้น ม.3 สูงกว่าชั้น ม.1 แต่ผลสัมฤทธิ์ทางการเรียนชั้น ม.3 และชั้น ม.2 ไม่แตกต่างกัน และผลสัมฤทธิ์ทางการเรียนชั้น ม.2 และชั้น ม.1 ไม่แตกต่างกัน ในกลุ่มที่มีความสามารถปานกลาง ผลสัมฤทธิ์ทางการเรียน ชั้น ม.1 ม.2 และ ม.3 ก้าวหน้าขึ้นตามลำดับ โดยที่ผลสัมฤทธิ์ทางการเรียนชั้น ม.3 สูงกว่า ชั้น ม.2 และผลสัมฤทธิ์ทางการเรียนของชั้น ม.2 สูง

กว่าชั้น ม.1 ส่วนในกลุ่มที่มีความสามารถต่ำผลสัมฤทธิ์ทางการเรียนของชั้น ม.3 สูงกว่าชั้น ม.2 และ ม.1 ตามลำดับ แต่ผลสัมฤทธิ์ทางการเรียนของชั้น ม.1 และ ม.2 ไม่แตกต่างกัน

ชูชีพ พงษ์สมบุรณ์ (2528 : 112 – 118) ได้เปรียบเทียบการเทียบมาตราระหว่างรูปแบบที่ใช้แบบทดสอบร่วมกับรูปแบบที่ใช้ผู้สอบร่วมโดยใช้วิธีการเทียบ 3 วิธี คือ การเทียบเชิงเส้นตรง การเทียบโดยใช้เปอร์เซ็นต์ไทล์และการเทียบโดยใช้โด่งลักษณะข้อสอบเพื่อต้องการเปรียบเทียบประสิทธิภาพและความคงที่ของการเทียบมาตรา ทั้ง 3 วิธี ใช้กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2526 จำนวน 3,721 คน ที่สอบจากแบบทดสอบของโครงการตรวจสอบคุณภาพการศึกษาทั่วประเทศในวิชาคณิตศาสตร์ ของกรมวิชาการ กระทรวงศึกษาธิการ ใช้ข้อสอบ จำนวน 60 ข้อ แบ่งออกเป็น 2 ฉบับๆ ละ 38 ข้อ ซึ่งรวมด้วยแบบทดสอบร่วม 20 ข้อ ด้วยผลการศึกษาพบว่า ผลของการเปรียบเทียบประสิทธิภาพของการเทียบมาตราระหว่างรูปแบบของแบบทดสอบร่วมกับรูปแบบที่ใช้ผู้สอบร่วมในแต่ละวิธีไม่แตกต่างกัน การเปรียบเทียบความคงที่การเทียบมาตราที่ใช้ผู้สอบร่วม พบว่าการเทียบโดยใช้โด่งลักษณะข้อสอบคงที่พอๆกับการเทียบโดยใช้เปอร์เซ็นต์ไทล์ และการเทียบเชิงเส้นตรงมีความคงที่ต่ำที่สุด ส่วนการเปรียบเทียบความคงที่ของการเทียบมาตราที่ใช้แบบทดสอบร่วม ได้ผลเช่นเดียวกับแบบใช้ผู้สอบร่วม

ภาวิณี ศรีสุขวัฒนานนท์ (2529 : 155 – 170) ศึกษาเปรียบเทียบผลการใช้รูปแบบการเทียบมาตราที่ต่างกัน 3 รูปแบบ คือ รูปแบบอิกวิเปอร์เซ็นต์ไทล์ รูปแบบเชิงเส้นตรง และรูปแบบอิงทฤษฎีการตอบข้อคำถามแบบโลจิสติกส์สามพารามิเตอร์ ด้วยการออกแบบเทียบมาตราใช้แบบทดสอบร่วมภายในที่มีความยาวต่างกันสามขนาดคิดเป็นร้อยละ คือร้อยละ 60 (21 ข้อ) ร้อยละ 40 (14 ข้อ) และร้อยละ 20 (7 ข้อ) โดยใช้ข้อสอบสอบทั้งหมด 35 ข้อ การประเมินผลของการเทียบมาตรา โดยการวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการเทียบมาตราใช้เป็นเกณฑ์ภายใน เพื่อเปรียบเทียบความแม่นยำระหว่างวิธีการเทียบภายในรูปแบบเดียวกัน และวิเคราะห์กลุ่มทวนผลใช้เป็นเกณฑ์ภายนอก เพื่อเปรียบเทียบความเพียงพอของการเทียบมาตราโดยตรงกับผลสุดท้ายของการเทียบมาตรา โดยศึกษากับประชากรสองกรณี คือ การสอบคัดเลือกจากประชากร 15,875 คน และแบบทดสอบวัดผลสัมฤทธิ์ จากประชากร 12,383 คน ในแต่ละกรณีใช้กลุ่มตัวอย่าง กลุ่มละ 1,500 คน แบ่งแบบทดสอบกรณีละ 2 ชุด โดยใช้ชุดละ 35 ข้อ มีแบบทดสอบร่วมขนาด 21 , 14 และ 7 ข้อ โดยยึดหลักแบบทดสอบร่วมค่าความเชื่อมั่นเท่าเทียมกัน ในแต่ละคู่ใช้รูปแบบการเทียบมาตรา 3 รูปแบบ สร้างตารางแปลงคะแนนสมมูลจากคะแนนแบบทดสอบชุดที่ 1 ไปยังแบบทดสอบชุดที่ 2 แล้ววิเคราะห์ความคลาดเคลื่อนมาตรฐานการเทียบมาตรา ค่าดัชนีประสิทธิภาพสัมพัทธ์ (RE) การวิเคราะห์กลุ่มทวนผล โดยคำนวณค่าดัชนีเปรียบเทียบความแตกต่าง (C) ผลการศึกษาพบว่า การเทียบมาตราทั้ง 9 รูปแบบ ได้ผลเป็นที่ยอมรับได้ความยาวแบบทดสอบร่วมมีผลต่อความแม่นยำ และความเพียงพอของการเทียบมาตรา โดยที่แบบทดสอบร่วมมีความยาวมากและแบบทดสอบร่วมทั้งสองฉบับ มีความเชื่อมั่นสูงและใกล้เคียงกัน จะทำให้การเทียบมาตราแม่นยำมากขึ้น และในแต่ละสถานการณ์การสอบให้ผลการเทียบมาตราในรูปแบบต่างกัน คือ กรณีการสอบคัดเลือก รูปแบบที่ให้ผลดีตามลำดับ คือ อิกวิเปอร์เซ็นต์ไทล์ รูปแบบอิงทฤษฎีการตอบข้อคำถามและรูปแบบอิงเส้นตรง ส่วนกรณีวัดผลสัมฤทธิ์ รูปแบบที่ให้ผลดีตามลำดับคือ รูปแบบเชิงเส้นตรง รูปแบบอิกวิเปอร์เซ็นต์ไทล์ และรูปแบบอิงทฤษฎีการตอบข้อคำถาม

เรวดี อินทะสระ (2530 : 64 – 67) ศึกษาเปรียบเทียบความคลาดเคลื่อนมาตรฐานและความเพียงพอของการเทียบคะแนนตามทฤษฎีการตอบข้อคำถามแบบสามพารามิเตอร์กับวิธีการวิเคราะห์องค์ประกอบ กลุ่มตัวอย่างในการศึกษาเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2529 จำนวน 2,823 คน แบบทดสอบที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์ วิชาคณิตศาสตร์ 2 ฉบับๆ ละ 45 ข้อ โดยมีแบบทดสอบรวมภายใน ฉบับละ 15 ข้อ ผลการวิจัยพบว่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนโดยวิธีสามพารามิเตอร์น้อยกว่าวิธีการใช้เทคนิคการวิเคราะห์องค์ประกอบ ส่วนความเพียงพอของการเทียบคะแนน โดยวิธีสามพารามิเตอร์อยู่ในระดับที่น่าพอใจ ส่วนวิธีการวิเคราะห์องค์ประกอบอยู่ในระดับกลาง

สรุปแล้วงานวิจัยในประเทศส่วนใหญ่จะศึกษาเทคนิคและวิธีการเทียบคะแนน เพื่อให้ได้ผลการเทียบที่ถูกต้องแม่นยำ แต่ไม่มีวิธีการใดที่ให้ผลดีในทุกสถานการณ์ต้องอาศัยปัจจัยอื่นๆหลายอย่าง

จากเอกสารและงานวิจัยที่เกี่ยวข้องทั้งต่างประเทศและในประเทศดังกล่าว จะเห็นว่า การวิจัยเกี่ยวกับการเทียบคะแนนส่วนใหญ่ศึกษาวิจัยเทคนิคและวิธีการเทียบคะแนน เพื่อให้ได้ผลการเทียบที่ถูกต้องแม่นยำ ผลจากการวิจัยสรุปได้ว่า ไม่มีวิธีการใดเพียงวิธีเดียวที่ให้ผลดีในทุกสถานการณ์ วิธีการเทียบคะแนนที่ใช้ได้ผลดีที่สุดดูเหมือนจะขึ้นอยู่กับองค์ประกอบหลายอย่างได้แก่ ความยาวของแบบทดสอบ เนื้อหาของแบบทดสอบ วิธีการประมาณค่าพารามิเตอร์ วิธีปรับค่าพารามิเตอร์ การกระจายของค่าความยากของข้อสอบ และค่าพารามิเตอร์ความสามารถของกลุ่มตัวอย่างที่ใช้ วิธีการประเมินความเพียงพอของการเทียบคะแนน ซึ่งอาจมีอิทธิพลต่อการแปลความหมายของคะแนน ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาเปรียบเทียบผลการเทียบคะแนนระหว่างการใช้คะแนนจริง , คะแนนจากการสังเกตตามทฤษฎีการตอบข้อคำถาม และวิธีอิกวิเปอร์เซนไทล์ ว่าวิธีใดจะมีความเหมาะสมมากกว่ากัน

สมมติฐานของการศึกษาค้นคว้า

การเทียบคะแนนวิธีอิกวิเปอร์เซนไทล์เป็นวิธีการนำทฤษฎีคลาสสิคอลล (Classical Test Theory) มาใช้ มักมีปัญหาในการวัด ดังที่ แฮมเบิลตัน (Hambleton .1978) ได้เสนอไว้หลายประการ ที่เห็นได้ชัดประการหนึ่งคือ ค่าพารามิเตอร์ของข้อสอบ คือค่าความยาก ค่าอำนาจจำแนกมีค่าไม่คงที่ เปลี่ยนไปตามความสามารถของผู้สอบ ดังนั้นการเทียบด้วยวิธีอิกวิเปอร์เซนไทล์จากกลุ่มที่มีความสามารถต่างกันก็จะให้ผลที่ต่างกันไป แต่เพื่อเป็นการลดความแตกต่างในข้อบกพร่องนี้จึงนำวิธีการใช้ผู้สอบร่วมเข้ามามีใช้ในการเทียบคะแนน ส่วนการเทียบคะแนนโดยใช้โด่งลักษณะข้อสอบซึ่งเป็นการใช้ทฤษฎีการตอบข้อคำถาม (Item Response Theory) โดยใช้การเทียบด้วยคะแนนจริง และคะแนนสังเกต ซึ่งเป็นการประมาณค่าพารามิเตอร์ของข้อสอบ จะไม่เปลี่ยนไปตามกลุ่มตัวอย่าง (Sample free) และจากงานวิจัยของฮาน ,โคเลนและโพลแมนน์ (Han,Kolen and Pohlmann .1997:105 - 121) ศึกษาการเทียบคะแนนระหว่างทฤษฎีการตอบข้อคำถามที่ใช้คะแนนจริง,คะแนนสังเกตและวิธีอิกวิเปอร์เซนไทล์ นั้น พบว่า วิธีเทียบคะแนนจริงให้ความคงเส้นคงวามากกว่าการเทียบคะแนนสังเกต, การเทียบคะแนนสังเกตให้ความคงเส้นคงวามากกว่าวิธีอิกวิเปอร์เซนไทล์ และค่าเฉลี่ยของความคลาดเคลื่อนนั้นทั้งสามวิธีแตกต่างกัน จากผลที่เกิดขึ้นนี้ก็ยังไม่เป็นข้อสรุปที่ชัดเจนแน่นอนจึงยังจำเป็นต้องมีการตรวจสอบอีกต่อไป ดังนั้นผู้วิจัยจึงตั้งสมมติฐานเกี่ยวกับการเทียบคะแนนทั้งสามวิธี ดังนี้

1. คะแนนที่เทียบได้จากวิธีการเทียบจากคะแนนจริงและการเทียบคะแนนการสังเกตตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์มีค่าแตกต่างกัน

2. ความคลาดเคลื่อนของการเทียบจากคะแนนจริงและการเทียบคะแนนการสังเกตตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์มีค่าแตกต่างกัน

3. ความเพียงพอของการเทียบจากคะแนนจริงและการเทียบคะแนนการสังเกตตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์มีค่าแตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2542 สังกัดสำนักงานการประถมศึกษาจังหวัดหนองคาย จำนวน 485 โรงเรียน จำนวนนักเรียน 12,707 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2542 สังกัด สำนักงานการประถมศึกษาจังหวัดหนองคาย ได้มาโดยวิธีการสุ่มแบบ สองขั้นตอน (Two – Stage Random Sampling)ซึ่งมีวิธีการสุ่มดังนี้

ขั้นที่ 1 สุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีขนาดโรงเรียนเป็นชั้น (Strata) และมีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) โดยสุ่มมา 10 % ของจำนวนโรงเรียนทั้งหมด ได้โรงเรียน 47 โรงเรียน มีจำนวนห้องเรียน 60 ห้องเรียน จากการศึกษาพบว่าโรงเรียนขนาดใหญ่มี 52 โรงเรียน สุ่มมา 5 โรงเรียน จำนวนห้องเรียน 18 ห้องเรียน โรงเรียนขนาดกลางจำนวน 206 โรงเรียน สุ่มมา 20 โรงเรียน จำนวนห้องเรียน 20 ห้องเรียน และโรงเรียนขนาดเล็กจำนวน 227 โรงเรียน สุ่มมา 22 โรงเรียน จำนวนห้องเรียน 22 ห้องเรียน

ขั้นที่ 2 วิธีการสุ่มอย่างง่าย (Simple Random Sampling) โดยสุ่มห้องเรียนจากโรงเรียนที่สุ่มได้ในขั้นที่ 1 โดยใช้ห้องเรียนเป็นหน่วยการสุ่ม (Sampling Unit) สุ่มกลุ่มตัวอย่างมาสำหรับเทียบคะแนน จำนวน 1,062 คน และสำหรับทานผล 523 คน รวมทั้งหมด 1,585 คน ซึ่งเชื่อมั่นได้ 95 % (ล้วน สายยศ และอังคณา สายยศ.2540 :401.อ้างจาก Yamane.1967:886 - 887) ดังตาราง 1

ตาราง 1 จำนวนนักเรียนในกลุ่มตัวอย่างที่ใช้ในการเทียบคะแนนและการทานผลการเทียบคะแนน

โรงเรียน	กลุ่มเทียบคะแนน (จำนวนห้องเรียน)	กลุ่มทานผล (จำนวนห้องเรียน)
ขนาดใหญ่		
1. อนุบาลหนองคาย	125 (3)	76 (2)
2. อนุบาลเซกา	104 (3)	48 (1)
3. อนุบาลวิศิษฐ์อำเภวนายศิลป์	60 (2)	30 (1)
4. อนุบาลปากคาด	61 (2)	36 (1)
5. โกมลวิทยาคาร	90 (2)	47 (1)

ตาราง 1 (ต่อ)

โรงเรียน	กลุ่มเทียบคะแนน (จำนวนห้องเรียน)	กลุ่มทานผล (จำนวนห้องเรียน)
<u>ขนาดกลาง</u>		
1. บ้านบึงกาฬ	36 (1)	
2. บ้านหนองยาว		21 (1)
3. บ้านท่าโพธิ์	17 (1)	
4. บ้านเหล่าคาม	30 (1)	
5. บ้านโนนสูงสุขสมบูรณ์	26 (1)	
6. บ้านก่องนาง		34 (1)
7. บ้านแสนสำราญ	17 (1)	
8. หอคำพิทยาสรรค์	20 (1)	
9. บ้านทองสาย		30 (1)
10. ประสานมิตรวิทยา	42 (1)	
11. บ้านโคกนิคม	23 (1)	
12. บ้านโคกหนองลาด		34 (1)
13. บ้านห้วยไม้ซอด	22 (1)	
14. บ้านดงบัง	32 (1)	
15. บ้านนาคำแคน	45 (1)	
16. บ้านหนองแก่งทราย	39 (1)	
17. ท่าไร่วิทยา		26 (1)
18. บ้านคำไผ่	48 (1)	
19. บ้านหนองพันทา		23 (1)
20. บ้านน้ำโมง		29 (1)
<u>ขนาดเล็ก</u>		
1. บ้านจอมมณี	18 (1)	
2. บ้านท่าไคร้	11 (1)	
3. บ้านหนองตอ	18 (1)	
4. บ้านเล็ดสิน		10 (1)
5. บ้านโคกสว่าง		12 (1)

ตาราง 1 (ต่อ)

โรงเรียน	กลุ่มเทียบคะแนน (จำนวนห้องเรียน)	กลุ่มทานผล (จำนวนห้องเรียน)
6. นาเจริญวิทยา	16 (1)	
7. บ้านนาคำ	18 (1)	
8. บ้านสุขสำราญ		11 (1)
9. บ้านอาสง	11 (1)	
10. บ้านท่าสะอาด	7 (1)	
11. บ้านกำแพงเพชร		8 (1)
12. ชนตรีวิทยา	18 (1)	
13.บ้านสรรเสริญ		8 (1)
14. บ้านโนนยาง	8 (1)	
15. บ้านคำภู	9 (1)	
16. บ้านโนนจำปา	8 (1)	
17. บ้านห้วยเชือกเหนือ	17 (1)	
18. บ้านคำหมื่น "ต.ช.ด.อนุสรณ์"	17 (1)	
19. บ้านหนองคอน	19 (1)	
20. บ้านดงกำฟ้า		16 (1)
21. ชุมชนบ้านหม้อ		24 (1)
22.บ้านดอนหญ้านาง	23 (1)	
รวม	1,062 (40)	523 (20)

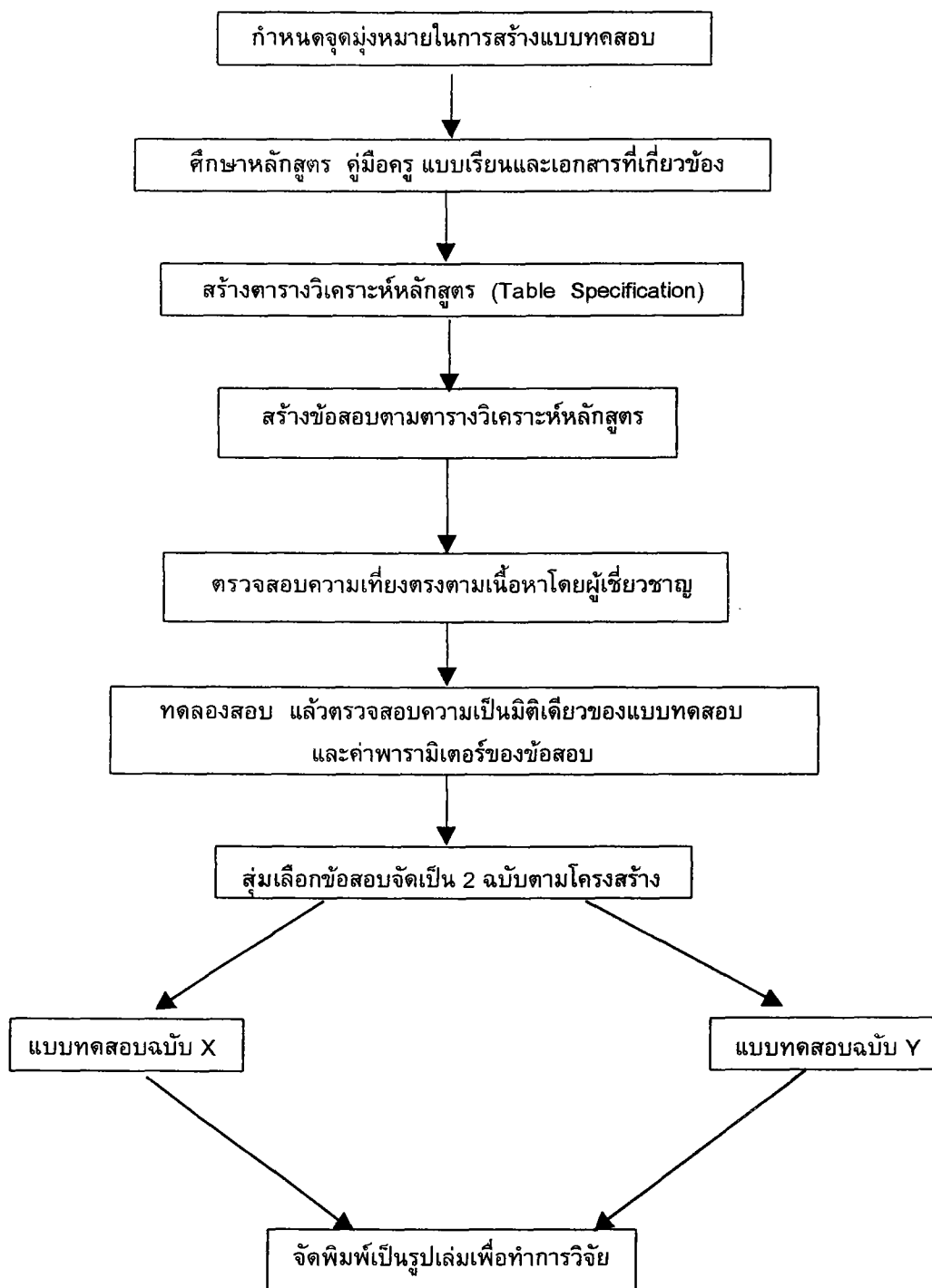
เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบในวิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 เรื่อง สมการ ปีการศึกษา 2542 ที่ผู้วิจัยสร้างขึ้นตามหลักสูตรประถมศึกษา พ.ศ. 2521 (ฉบับปรับปรุง พ.ศ. 2533) จำนวน 100 ข้อ เป็นแบบทดสอบ เลือกตอบ ชนิด 5 ตัวเลือก และแบ่งข้อสอบออกเป็น 2 ฉบับ โดยพิจารณาจาก ค่าความยาก ที่เท่ากันหรือเทียบเคียงกันจะได้แบบทดสอบสองฉบับ ดังนี้

1. แบบทดสอบวิชาคณิตศาสตร์ ฉบับ X จำนวน 30 ข้อ
2. แบบทดสอบวิชาคณิตศาสตร์ ฉบับ Y จำนวน 30 ข้อ

วิธีดำเนินการสร้างแบบทดสอบ

ในการสร้างแบบทดสอบผู้วิจัยได้ดำเนินการสร้างแบบทดสอบตามแผนผังดังนี้



ภาพประกอบ 3 แผนผังการสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

รายละเอียดในการดำเนินการสร้างแบบทดสอบมีดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ
 - 1.1 เพื่อสร้างแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เพื่อจะนำไปใช้ในการวิจัย
 - 1.2 เพื่อเปรียบเทียบผลการเทียบคะแนนของแบบทดสอบ
 2. ศึกษาหลักสูตร คู่มือครู แบบเรียน คู่มือการประเมินผลวิชาคณิตศาสตร์และเอกสารที่เกี่ยวข้องในเนื้อหาวิชาคณิตศาสตร์ระดับชั้นประถมศึกษาปีที่ 6 เรื่อง สมการ
 3. สร้างตารางวิเคราะห์หลักสูตรโดยผู้เชี่ยวชาญจำนวน 5 คน เป็นผู้พิจารณาว่าแต่ละเนื้อหาวัดพฤติกรรมใดและมีน้ำหนักเท่าใด
 4. สร้างข้อสอบซึ่งมีลักษณะเป็นแบบทดสอบเลือกตอบจำนวน 100 ข้อตามเนื้อหาและพฤติกรรมที่กำหนดไว้ในตารางวิเคราะห์หลักสูตร(Table Specification)โดยยึดข้อตกลงเบื้องต้นของการเทียบคะแนนซึ่งเสนอโดย เลวิน (Holland and Rubin. 1982 : 10; citing Levin) ก็คือแบบทดสอบ 2 ฉบับใดๆ จะเทียบกันได้ ก็ต่อเมื่อแบบทดสอบ 2 ฉบับนั้นคู่ขนานกันในด้านต่อไปนี้
 1. โครงสร้าง (Structure)
 2. เวลา (Timing)
 3. ชนิดของข้อสอบ (Item Type)
 4. รูปแบบ (Format)
 5. เนื้อหาวิชา (Subject Matter)
- เนื้อหาที่ใช้ในการออกข้อสอบ คือ เนื้อหาวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรประถมศึกษา 2521 (ฉบับปรับปรุง พ.ศ. 2533) เรื่อง สมการ ซึ่งมีเนื้อหาย่อยดังนี้
1. สมการ
 2. สมการที่เป็นจริง
 3. สมการที่มีตัวไม่ทราบค่า
 4. คำตอบของสมการ
 5. สมบัติการเท่ากันเกี่ยวกับการบวกและลบ
 6. การแก้สมการโดยใช้การเท่ากันเกี่ยวกับการบวกและลบ
 7. สมบัติการเท่ากันเกี่ยวกับการคูณและหาร
 8. การแก้สมการโดยใช้สมบัติการเท่ากันเกี่ยวกับการคูณและการหาร
 9. การแก้โจทย์ปัญหาด้วยสมการ
 10. การเขียนสมการจากข้อความที่กำหนดให้

ซึ่งแบบทดสอบที่สร้างขึ้นทั้ง 2 ฉบับ ดังตัวอย่าง

ฉบับ X (เนื้อหาการแก้โจทย์ปัญหาด้วยสมการ)

พฤติกรรมที่ต้องการวัด การนำไปใช้

ตัวอย่างแบบทดสอบ

1. จออ่านหนังสือได้ M วัน ละคร 6 หน้า

ถ้าอ่าน ได้ 108 หน้า อยากทราบว่า

จออ่านหนังสือกี่วัน

- ก. 8 วัน
- ข. 9 วัน
- ค. 16 วัน
- ง. 18 วัน**
- จ. 20 วัน

ฉบับ Y (เนื้อหาการแก้โจทย์ปัญหาด้วยสมการ)

พฤติกรรมที่ต้องการวัด การนำไปใช้

ตัวอย่างแบบทดสอบ

1. แม่ขายขนมได้วันละ K กล่อง ละคร 15 บาท

ถ้าแม่ขายได้ 150บาท อยากทราบว่า

ว่าแม่ขายขนมไปกี่กล่อง

- ก. 5 กล่อง
- ข. 10 กล่อง**
- ค. 15 กล่อง
- ง. 20 กล่อง
- จ. 25 กล่อง

5 .ตรวจสอบความเที่ยงตรงตามเนื้อหา (Content Validity)โดยนำข้อสอบที่สร้างขึ้นไปให้ผู้เชี่ยวชาญตรวจสอบว่าข้อสอบที่สร้างขึ้นวัดเนื้อหาและพฤติกรรมที่ต้องการหรือไม่โดยพิจารณาค่า IOC ที่มีค่าตั้งแต่ 0.5 ขึ้นไป ประเมินได้ว่าเป็นข้อสอบที่มีความเที่ยงตรงตามเนื้อหา

ตัวอย่าง แบบประเมินความสอดคล้องระหว่างแบบข้อสอบกับพฤติกรรมที่ต้องการวัด

คำชี้แจง ให้ท่านพิจารณาว่าข้อสอบแต่ละข้อต่อไปนี้ วัดได้ตรงตามเนื้อหาและพฤติกรรมที่กำหนดไว้หรือไม่ โดยเขียนเครื่องหมาย $\checkmark$ ลงในช่องความคิดเห็น พิจารณาให้คะแนนตามเกณฑ์ดังนี้

ให้คะแนน +1 เมื่อท่านพิจารณาแล้วเห็นว่าข้อสอบนั้นวัดได้ตรงตามพฤติกรรมนั้นจริง

ให้คะแนน 0 เมื่อท่านพิจารณาแล้วไม่แน่ใจว่าข้อสอบข้อนั้นวัดได้ตรงตามพฤติกรรมนั้นจริง

ให้คะแนน -1 เมื่อท่านพิจารณาแล้วเห็นว่าข้อสอบข้อนั้นไม่ได้วัดตรงตามพฤติกรรมนั้นจริง

เนื้อหา	พฤติกรรมที่ต้องการวัด	ข้อสอบ	คะแนนการพิจารณา		
			+1	0	-1
โจทย์ปัญหา สมการ	การนำไปใช้	00. นิดมีเงินอยู่ 70 บาท แบ่งให้น้อง X บาท นิดเหลือเงินอยู่ 39 บาท อยากทราบว่านิดแบ่งเงินให้น้องกี่บาท	$\checkmark$		
		ก. 21 บาท			
		ข. 31 บาท			
		ค. 41 บาท			
		ง. 45 บาท			
		จ. 46 บาท			

หลังจากให้ผู้เชี่ยวชาญพิจารณาความสอดคล้อง ระหว่างข้อสอบกับพฤติกรรมที่ต้องการวัด นำคะแนนไปคำนวณหาค่าดัชนีความสอดคล้อง (ล่วน สายยศ และอังคณา สายยศ.2536) ถ้ามีค่ามากกว่า 0.5 ถือว่าข้อสอบข้อนั้นๆ วัดตรงตามพฤติกรรมนั้นจริง ผลปรากฏว่าได้ข้อสอบที่มีค่าดัชนีความสอดคล้อง ตั้งแต่ 0.5 - 1.0 จำนวน 100 ข้อ ดังแสดงในตาราง 21 ภาคผนวก ญ

6. นำแบบทดสอบที่สร้างขึ้นทั้งหมดไปทดลองสอบเองกับนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2542 จำนวน 500 คน ได้แก่นักเรียนจากโรงเรียนบ้านนาเจริญ จำนวน 28 คน ,บ้านไร่สุขสันต์ จำนวน 19 คน,บ้านโนนสว่างจำนวน 28 คน ,บ้านหนองจันทร์ จำนวน 41 คน, บ้านศรีวิไล จำนวน 44 คน ,บ้านห้วยสามยอดเทวกุล จำนวน 21 คน, บ้านโป่งเปือย จำนวน 24 คน ,บ้านเหล่าถาวร จำนวน 25 คน, บ้านหนองนาแซง จำนวน 31 คน, บ้านนาป่า จำนวน 13 คน ,บ้านอาสง จำนวน 26 คน , บ้านห้วยเชื่อมเหนือ จำนวน 28 คน, บ้านท่าสะอาด จำนวน 31 คน,บ้านไคสี จำนวน 27คน และบ้านห้วยเชื่อมใต้ จำนวน 35 คน ในช่วงเดือน สิงหาคม ถึง กันยายน พ.ศ. 2542 เพื่อนำมาวิเคราะห์ดังนี้

6.1 ตรวจสอบความเป็นมิติเดียวกันของแบบทดสอบ (Unidimensionality)โดยวิธีการวิเคราะห์องค์ประกอบ (Factor Analysis) โดยใช้โปรแกรม SPSS/FW ซึ่งเลือกการสกัดองค์ประกอบแบบแกนหลัก (Principal Component : PC) แล้วหมุนแกนด้วยวิธีแวนริแมกซ์ (Varimax) ได้ผลดังตาราง 12 ในภาคผนวก ข ซึ่งปรากฏว่า ได้องค์ประกอบที่มีค่าไอเกน (Eigen Value)ขององค์ประกอบที่ 1 ของแบบทดสอบ มีค่าเท่ากับ 16.60 และค่าไอเกนขององค์ประกอบที่ 2 ถึงองค์ประกอบที่ 30 มีค่าอยู่ระหว่าง 3.720 ถึง 1.013 แสดงให้เห็นว่าองค์ประกอบที่ 1 มีค่าไอเกนสูงกว่าองค์ประกอบอื่นๆเป็นอย่างมาก ส่วนองค์ประกอบที่ 2 ถึง 30 มีค่าไอเกนที่ใกล้เคียงกัน แสดงว่าแบบทดสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่องสมการ มีความเป็นมิติเดียว (Warm. 1978 :104 ; Lord. 1980 :19) นั่นคือ สามารถนำแบบทดสอบไปวิเคราะห์ตามทฤษฎีการตอบข้อสอบได้

6.2 ประเมินค่าพารามิเตอร์ของข้อสอบจากแบบทดสอบ คือ ค่าความยาก (b) อำนาจจำแนก (a) และค่าความสัมประสิทธิ์การเดา (c) โดยการวิเคราะห์ด้วยโปรแกรม BILLOG version 3.11 (Mislevy and Bock.1990) ผลปรากฏว่า มีค่าความยากอยู่ระหว่าง -0.004 ถึง 2.427 ค่าอำนาจจำแนกมีค่าอยู่ระหว่าง 0.043 ถึง 1.730 และค่าสัมประสิทธิ์การเดาอยู่ระหว่าง 0.111 ถึง 1.053 (รายละเอียดดังตาราง 13 ในภาคผนวก ค) ผู้วิจัยจึงคัดเลือกข้อสอบที่มีคุณภาพถึงเกณฑ์ คือค่าความยากตั้งแต่ -2.00 ถึง 2.00

7. คัดเลือกข้อคำถามโดยพิจารณาค่าพารามิเตอร์ในข้อ 6.2 มาจับคู่ 30 คู่ ตามเนื้อหา พฤติกรรมและค่าพารามิเตอร์ จากนั้นแบ่งออกเป็น 2 ฉบับ ได้ข้อคำถามฉบับละ 30 ข้อ คือ ฉบับ X จำนวน 30 ข้อ และ ฉบับ Y จำนวน 30 ข้อ ซึ่งมีคุณภาพดังนี้

แบบทดสอบฉบับ X มีค่าพารามิเตอร์อำนาจจำแนกอยู่ระหว่าง 0.569 ถึง 1.744 ค่าพารามิเตอร์ความยากอยู่ระหว่าง -1.849 ถึง 1.922 และค่าสัมประสิทธิ์การเดามีค่าอยู่ระหว่าง 0.105 ถึง 0.376 ส่วนค่าอินฟอร์เมชันของแบบทดสอบมีค่าอยู่ระหว่าง 0.225 ถึง 2.745 และ ณ ระดับความสามารถที่เพิ่มขึ้น ค่าอินฟอร์เมชันของแบบทดสอบจะเพิ่มขึ้น ซึ่งจะให้ค่าสูงสุดที่ระดับความสามารถ 1.4

แบบทดสอบฉบับ Y มีค่าพารามิเตอร์อำนาจจำแนกอยู่ระหว่าง 0.559 ถึง 1.768 ค่าพารามิเตอร์ความยากมีค่าอยู่ระหว่าง -1.235 ถึง 1.845 ค่าสัมประสิทธิ์การเดามีค่าอยู่ระหว่าง 0.074 ถึง 0.262 ส่วนค่าอินฟอร์เมชันของแบบทดสอบมีค่าอยู่ระหว่าง 0.154 ถึง 3.375 และ ณ ที่ระดับความสามารถที่เพิ่มขึ้น ค่าอินฟอร์เมชันของแบบทดสอบจะเพิ่มขึ้น ซึ่งจะให้ค่าที่สูงสุดที่ระดับความสามารถ 1.6

จากแบบทดสอบที่ได้จำนวน 2 ฉบับ (ฉบับ X และฉบับ Y) มีคุณภาพดังรายละเอียดในภาคผนวก จ

7. นำแบบทดสอบที่สร้างได้ในข้อ 7 ไปใช้ในการวิจัยต่อไป

วิธีดำเนินการเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูลในครั้งนี้ ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1. วางแผนในการดำเนินการสอบ โดยติดต่อขอความร่วมมือในการนำแบบทดสอบไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่าง กำหนดวัน เวลาและสถานที่สอบ ในช่วงเดือนพฤศจิกายน พ.ศ. 2542 ถึงเดือนกุมภาพันธ์ พ.ศ. 2543

2. นำแบบทดสอบทั้ง 2 ฉบับ คือฉบับ X และฉบับ Y ไปทดสอบกับกลุ่มตัวอย่าง ฉบับ X จำนวน 531 คน ฉบับ Y จำนวน 531คน จากการสุ่มกลุ่มตัวอย่างที่ใช้ห้องเรียนเป็นหน่วยการสุ่มแล้วใช้การสุ่มจัดกลุ่มให้สมดุล (Counterbalanced Random Group Design) โดยแจกแบบทดสอบแก่คนที่ 1 ด้วยฉบับ X แล้วตามด้วยฉบับ Y และคนที่ 2 ด้วยฉบับ Y แล้วตามด้วยฉบับ X ดังเช่น XY, YX, XY, YX,... และกลุ่มทวนผล 523 คน

3. นำผลการสอบที่ได้ในข้อ 2 มาเตรียมข้อมูลสำหรับการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบและพารามิเตอร์ของข้อสอบ แล้วนำผลการประมาณค่าไปวิเคราะห์ทางสถิติและทดสอบสมมติฐานการวิจัยต่อไป

การวิเคราะห์ข้อมูลและสถิติที่ใช้

การวิเคราะห์ข้อมูลดำเนินการโดยแบ่งออกเป็น 4 ตอน ดังนี้

1. การวิเคราะห์แบบทดสอบ
2. การสร้างตารางเทียบคะแนนแบบทดสอบ
3. การศึกษาผลการเทียบคะแนน
4. การทดสอบความแตกต่างของผลการเทียบคะแนน

ตอนที่ 1 การวิเคราะห์แบบทดสอบ

1.1 ตรวจสอบความเป็นมิติเดียว (Unidimensionality) ของแบบทดสอบแต่ละฉบับ ตามข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อคำถาม โดยนำแบบทดสอบแต่ละฉบับไปทดลองสอบกับนักเรียนชั้นประถมศึกษาปีที่ 6 ที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 500 คน แล้วนำผลมาวิเคราะห์องค์ประกอบ (Factor Analysis) โดยใช้โปรแกรมสำเร็จรูป SPSS/FW ซึ่งเลือกสกัดองค์ประกอบด้วยวิธีการวิเคราะห์องค์ประกอบหลัก (Principal Component :PC) แล้วหมุนแกนด้วยวิธีแวนแมกซ์ (Varimax)

1.2 วิเคราะห์ข้อสอบในแบบทดสอบเทียบคะแนน เพื่อประมาณค่าพารามิเตอร์ของข้อสอบคือ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าสัมประสิทธิ์การเดา (c) ด้วยโปรแกรมสำเร็จรูป BILOG version 3.11 (Mislevy and Bock.1990) ได้ผลดังตาราง 14 ในภาคผนวก ง

1.3 หาค่าสถิติพื้นฐานของแบบทดสอบแต่ละฉบับก่อนการเทียบคะแนน คือค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน ด้วยโปรแกรมสำเร็จรูป SPSS/FW

1.4 ตรวจสอบสถิติเพื่อทดสอบการกระจายคะแนนของแบบทดสอบทั้งสองฉบับ คือ พิจารณาจากค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน และการแจกแจงของคะแนนจากแบบทดสอบทั้งสองฉบับว่ามีการกระจายใกล้เคียงที่ใกล้เคียงกัน

ตอนที่ 2 การสร้างตารางเทียบคะแนนแบบทดสอบ

ผู้วิจัยมีขั้นตอนการสร้างตารางเทียบคะแนนในแต่ละวิธีดังนี้

2.1 การเทียบคะแนนจริง

มีขั้นตอนการสร้างตารางการเทียบคะแนนดังนี้

ขั้นที่ 1 แปลงค่าพารามิเตอร์อำนาจจำแนก (a_i) ความยาก (b_i) และพารามิเตอร์ความสามารถ θ ที่ได้จากการประมาณค่าโดยใช้โปรแกรมสำเร็จรูป BILOG version 3.11 (Mislevy and Bock, 1990) ให้อยู่ในสเกลเดียวกัน จากสูตร (Warm, 1978: 116)

$$\theta_1 = \left(\frac{SD_{b_1}}{SD_{b_2}} \right) \theta_2 + \left(\bar{b}_1 - \left(\frac{SD_{b_1}}{SD_{b_2}} \right) \bar{b}_2 \right)$$

$$b_1 = \left(\frac{SD_{b_1}}{SD_{b_2}} \right) b_2 + \left(\bar{b}_1 - \left(\frac{SD_{b_1}}{SD_{b_2}} \right) \bar{b}_2 \right)$$

$$a_1 = a_2 \left(\frac{SD_{b_2}}{SD_{b_1}} \right)$$

$$C_{yi} = C_{xi}$$

โดยที่

SD_{b_1} คือ ค่าความเบี่ยงเบนมาตรฐานของพารามิเตอร์ความยากของแบบทดสอบฉบับ

X

SD_{b_2} คือ ค่าความเบี่ยงเบนมาตรฐานของพารามิเตอร์ความยากของแบบทดสอบฉบับ Y

θ_1 คือ ค่าพารามิเตอร์ความสามารถของผู้สอบที่ทำแบบทดสอบฉบับ X

θ_2 คือ ค่าพารามิเตอร์ความสามารถของผู้สอบที่ทำแบบทดสอบฉบับ Y

b_1 คือ พารามิเตอร์ความยากของข้อสอบฉบับ X

b_2 คือ พารามิเตอร์ความยากของข้อสอบ Y

a_1 คือ พารามิเตอร์อำนาจจำแนกของข้อสอบฉบับ X

a_2 คือ พารามิเตอร์อำนาจจำแนกของข้อสอบฉบับ Y

C_{xi} คือ พารามิเตอร์การเดาของฉบับ X ข้อที่ i

C_{yi} คือ พารามิเตอร์การเดาของฉบับ Y ข้อที่ i

ขั้นที่ 2 นำค่าพารามิเตอร์ที่แปลงแล้วมาคำนวณโอกาสในการตอบถูกทั้งสองกลุ่มจากสูตร

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

โดยที่ $P_i(\theta)$ คือ โอกาสในการตอบข้อสอบข้อที่ i ถูกของผู้สอบที่มีความสามารถ θ
 a_i คือ พารามิเตอร์ของค่าอำนาจจำแนกที่แปลงให้อยู่ในมาตราเดียวกันแล้ว
 b_i คือ พารามิเตอร์ของค่าความยากที่แปลงให้อยู่ในมาตราเดียวกันแล้ว
 c_i คือ พารามิเตอร์การเดาของข้อสอบข้อที่ i
 D คือ ค่าปรับสเกลมีค่าประมาณ 1.7
 e คือ ค่าคงที่ของ $\log_e$ มีค่าประมาณ 2.718

ขั้นที่ 3 คำนวณค่าคะแนนจริง (True Score) จากสูตรที่เสนอโดยลอร์ด (Lord . 1980 : 199)

$$\xi_x = \sum_{i=1}^n P_i(\theta_x)$$

$$\xi_y = \sum_{j=1}^m P_j(\theta_y)$$

โดยที่ ξ_x คือคะแนนจริงของผู้สอบที่มีความสามารถ θ_x ได้จากการตอบแบบทดสอบฉบับ X
 ξ_y คือคะแนนจริงของผู้สอบที่มีความสามารถ θ_y ได้จากการตอบแบบทดสอบฉบับ Y

ขั้นที่ 4 นำคะแนนจริงของผลการสอบที่หาได้ในขั้นที่ 3 มาหาความสัมพันธ์โดย พิจารณาค่า ξ_x และ ξ_y ที่สมนัยกัน ซึ่งเป็นการเทียบมาตราของแบบทดสอบทั้งสองชุด ณ ที่จุด แต่ละจุดของ θ เดียวกัน เป็นคะแนนจริงของแบบทดสอบทั้งสองฉบับ

ขั้นที่ 5 ทำตารางเทียบคะแนนทั้งสองฉบับโดยยึดค่า θ ระดับเดียวกันเป็นหลัก ปรากฏผลดังตาราง 15 ในภาคผนวก จ

หมายเหตุ การคำนวณหาค่า $P_i(\theta_x), P_j(\theta_y), \xi_x, \xi_y$ โดยใช้โปรแกรมคอมพิวเตอร์ ในการแยกคำนวณในแต่ละกรณีของแบบทดสอบ โดยใช้การกำหนดค่า θ ที่ครอบคลุมพิสัยกลุ่มตัวอย่าง

2.2 การเทียบคะแนนสังเกต

มีขั้นตอนการสร้างตารางเทียบคะแนนดังนี้คือ

ขั้นที่ 1 ประเมินค่าพารามิเตอร์ของความสามารถของผู้สอบแต่ละคน ในแบบทดสอบฉบับ X และฉบับ Y โดยการใช้โปรแกรมสำเร็จรูป BILOG version 3.11(Mislevy and Bock.1990)

ขั้นที่ 2 ประเมินค่าพารามิเตอร์ของข้อสอบในแบบทดสอบฉบับ X และฉบับ Y โดยการใช้โปรแกรมสำเร็จรูป BILOG version 3.11(Mislevy and Bock.1990)

ขั้นที่ 3 หาค่าการแจกแจงความถี่อย่างมีเงื่อนไข (Conditional Frequency Distribution $f(x/\theta)$) หรือ $E(X)$ ของคะแนนจากแบบทดสอบฉบับ X และ ฉบับ Y ด้วยสมการ (Lord. 1980 : 196 – 197 ; Hambleton and Swaminathan. 1985 : 215 - 216)

$$\sum_{x=0}^n f(x/\theta)t^x = \prod_{i=1}^n [Q_i(\theta) + tP_i(\theta)]$$

$$\sum_{y=0}^n f(y/\theta)t^y = \prod_{i=1}^n [Q_i(\theta) + tP_i(\theta)]$$

โดยที่ $f(x/\theta)$ คือ ค่า Conditional Frequency Distribution ของแบบทดสอบฉบับ X
 $f(y/\theta)$ คือ ค่า Conditional Frequency Distribution ของแบบทดสอบฉบับ Y
 x คือ จำนวนข้อที่ตอบถูกในแบบทดสอบฉบับ X
 y คือ จำนวนข้อที่ตอบถูกในแบบทดสอบฉบับ Y
 θ_i คือ พารามิเตอร์ความสามารถของผู้สอบคนที่ i โดยที่ $i = 1, 2, 3, \dots, n$
 t คือ ความถี่ หรือจำนวนคนที่ได้รับคะแนน x โดยที่ $x = 0, 1, 2, \dots, n$ และ y โดยที่ $y = 0, 1, 2, \dots, m$
 $P_i(\theta)$ คือ โอกาสหรือความน่าจะเป็นในการตอบข้อสอบข้อที่ i ถูกของผู้สอบที่มี
 ความสามารถ θ
 $Q_i(\theta)$ คือ ค่าที่หาได้จากสมการ $Q \equiv 1 - P$

ขั้นที่ 4 นำคะแนนที่แจกแจงจากขั้นที่ 3 มาหาความสัมพันธ์ของการเทียบคะแนนสังเกตของแบบทดสอบฉบับ X และฉบับ Y โดยพิจารณาจากคะแนนโอกาสในการตอบถูก กับคะแนนสังเกต (Observed Score) ณ ที่ระดับ θ เดียวกันเป็นหลัก

ขั้นที่ 5 สร้างตารางเทียบคะแนนของแบบทดสอบทั้ง 2 ฉบับที่ ระดับ θ เดียวกันเป็นหลัก ปรากฏผลดังตาราง 16 ในภาคผนวก ข

หมายเหตุ การคิดคำนวณโดยการเขียนโปรแกรมคำสั่งลงในเครื่องคอมพิวเตอร์

2.3 การเทียบคะแนนวิธีอีควิเปอร์เซ็นต์ไทล์ (Equipercentile Equating)

การเทียบคะแนนโดยวิธีการนี้อยู่บนพื้นฐานของนิยามที่ว่า คะแนนจากแบบทดสอบฉบับ X และ ฉบับ Y (เมื่อแบบทดสอบ X และ Y วัดสิ่งเดียวกัน และมีความเชื่อมั่นสูงพอๆกัน) จะถือว่าเป็นคะแนนสมมูลกัน (Equivalent score) เมื่อคะแนนของแต่ละฉบับนั้นมีตำแหน่งเปอร์เซ็นต์ไทล์ (Percentile Ranks) เท่ากัน

นั่นคือ X_i และ Y_i เป็นคะแนนสมมูลกัน

เมื่อ $X_{PR_k} = Y_{PR_k}$

โดย X_{PR_k} คือ คะแนนจากแบบทดสอบฉบับ X ตรงตำแหน่งเปอร์เซ็นต์ไทล์ k

Y_{PR_k} คือ คะแนนจากแบบทดสอบฉบับ Y ตรงตำแหน่งเปอร์เซ็นต์ไทล์ k

$$\text{โดย } X_{PR_k} = L_0 + \left[\frac{\frac{kn}{100} - F_b}{f} \right] \times C$$

เมื่อ L_0 คือ คะแนนขีดจำกัดล่างที่แท้จริงของคะแนนในชั้นเปอร์เซ็นต์ไทล์ที่ k

n คือ จำนวนสมาชิกในกลุ่มตัวอย่าง

f คือ ความถี่ของคะแนนในชั้นเปอร์เซ็นต์ไทล์ที่ k

F_b คือ ความถี่สะสมก่อนถึงคะแนนในชั้นเปอร์เซ็นต์ไทล์ที่ k

c คือ อัตรภาคชั้น

วิธีการ การเทียบคะแนนระหว่างแบบทดสอบ ฉบับ X และแบบทดสอบ Y โดยวิธีอิกวิเปอร์เซ็นต์ไทล์ สามารถกระทำได้ดังนี้

1. สุ่มกลุ่มผู้สอบให้เป็นตัวแทนประชากรเป้าหมายที่ต้องการนำแบบทดสอบไปใช้ โดยให้ได้กลุ่มผู้สอบที่มีความสามารถกระจาย (สูง, กลาง, ต่ำ) ครอบคลุมลักษณะประชากรเป้าหมาย แล้วทำการสุ่มแยกผู้สอบออกเป็น 2 กลุ่ม โดยให้กลุ่มหนึ่งทำแบบทดสอบฉบับ X และอีกกลุ่มหนึ่งทำแบบทดสอบฉบับ Y
2. นำคะแนนของแบบทดสอบแต่ละฉบับมาคำนวณหาค่าคะแนน ณ ตำแหน่งเปอร์เซ็นต์ไทล์ต่างๆ
3. สร้างตารางเทียบคะแนนระหว่างแบบทดสอบ ฉบับ X และแบบทดสอบฉบับ Y โดยถือว่าคะแนน ณ ตำแหน่งเปอร์เซ็นต์ไทล์เดียวกันของแบบทดสอบทั้ง 2 ฉบับ เป็นคะแนนสมมูลกัน

วิธีอิกวิเปอร์เซ็นต์ไทล์เป็นการสร้างคะแนนสมมูลระหว่างแบบทดสอบบนพื้นฐานความเท่าเทียมกันของค่าโมเมนต์ทั้ง 3 ทางสถิติ ซึ่งได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และการแจกแจงของคะแนน จึงเป็นกระบวนการแปลงคะแนนแบบไม่ใช่เส้นตรง (Nonlinear Transformation) ยังผลให้ความสัมพันธ์ระหว่างคะแนนดิบและคะแนนแปลงไม่เป็นเส้นตรง ซึ่งมีหลักการเกิดขึ้นจากความต้องการพื้นฐาน คือ $F=F^*$ โดยมีฟังก์ชันการเทียบคะแนนด้วยวิธีนี้คือ $e_X(Y) = F^{-1}(G(Y))$ (อ้างจาก ชูชีพ พงษ์สมบุญ . 2528 :16)

1. ในทางทฤษฎี $F(x)$ และ $G(y)$ มาจากกลุ่มตัวอย่างที่ได้จากการสุ่ม (P_1 และ P_2)

จากประชากร P

2. ในทางปฏิบัติ สถาบันบริการทดสอบทางการศึกษา (Educational Testing Service) ทำการสร้างแบบทดสอบฟอร์ม X และฟอร์ม Y แล้วสุ่มสลับกันลงในกล่องในแต่ละสนามสอบ ฉะนั้นผู้สอบครั้งหนึ่งจะสอบฟอร์ม X และอีกครั้งหนึ่งจะสอบฟอร์ม Y

3. การเทียบโดยใช้เปอร์เซ็นต์ไทล์จะเทียบ ฟอร์ม X และ Y โดยผู้สอบจากประชากร P

4. ทั้ง F และ G เป็นลักษณะต่อเนื่อง (Continuous) จาก $F^{-1}(G(Y))$
5. ความยาวของแบบทดสอบฟอร์ม X และ Y ไม่ควรแตกต่างกันมาก

ในการนำวิธีเทียบคะแนนอิกวิเปอร์เซนไทล์ไปใช้มีข้อแนะนำพอสรุปได้ดังนี้ (Potthoff.1982 ; Angoff.1984)

1. ควรใช้กลุ่มผู้สอบขนาดใหญ่ ควรเป็นกลุ่มที่มีความสามารถค่อนข้างกระจายและกระจายพอๆกัน ถ้าใช้กลุ่มตัวอย่างขนาดเล็กทำให้คะแนนมีความไวต่อความแปรปรวนเชิงสุ่ม ผลการเปรียบเทียบคะแนนสามารถแปรผันไปตามกลุ่มผู้สอบได้ (Group dependent)
2. การสร้างกราฟเปอร์เซนไทล์ เพื่อเปรียบเทียบคะแนนระหว่างแบบทดสอบ ควรสร้างด้วยความระมัดระวัง ไม่ลำเอียง
3. แบบทดสอบที่ต้องการนำมาเปรียบเทียบคะแนนกัน ต้องมีความเชื่อมั่นสูงพอๆกัน ถ้าต่างกันมาก จะทำให้การเปรียบเทียบคะแนนขาดความคงที่
4. การเปรียบเทียบคะแนนควรอยู่ในช่วงพิสัยของคะแนนสังเกต การเปรียบเทียบคะแนนที่อยู่นอกพิสัยคะแนนสังเกตจะมีความคลาดเคลื่อนสูงปรากฏผลดังตาราง 17,18,19 ในภาคผนวก ข

ตอนที่ 3 การศึกษาผลการเทียบคะแนน

3.1 คะแนนที่ได้จากการเทียบคะแนนตามข้อ 2.1 2.2 และ 2.3

3.2 ความคลาดเคลื่อนของการเทียบคะแนน (Standard Error of Equating) ซึ่งหาได้จากสูตร (Macro. 1977 : 157)

$$SEE = \left[\sum_{i=1}^N (Y_{ij} - Y_{io})^2 / N \right]^{1/2}$$

เมื่อ SEE คือ Standard Error Of Equating

Y_{ij} คือ คะแนน y ของผู้สอบคนที่ i จากวิธีการเทียบคะแนนที่ j ของแบบทดสอบ X

Y_{io} คือ คะแนน y ของผู้สอบคนที่ i จากวิธีการเทียบคะแนนที่ j ที่ได้จากการแปลงคะแนนที่สมมูลกัน

N คือจำนวนผู้สอบที่เป็นกลุ่มเทียบคะแนน

3.3 ความเพี้ยนพองของการเทียบคะแนน (Discrepancy Index)

ในการหาความเพี้ยนพองของการเทียบคะแนนในแต่ละวิธีหาได้จากสูตร (Petersen and others. 1982:91)

$$c = \sum (x_i - x_i^*)^2 / nS_x^2$$

เมื่อ c คือ ค่าความเพี้ยนพองของการเทียบคะแนน

x_i คือ คะแนน x ของผู้สอบคนที่ i จากแบบทดสอบชุด X

x_i^* คือ คะแนน x ของผู้สอบคนที่ i จากการแปลงคะแนน Y

n คือ จำนวนคนในกลุ่มตัวอย่างสอบทานผล

S_x^2 คือค่าความแปรปรวนของคะแนนจากแบบทดสอบ X

ซึ่งมีเกณฑ์ในการพิจารณาดังนี้ (Petersen and others. 1982 : 93-94)

ระดับน่าพอใจอย่างมาก	หมายถึง	$c \leq (.05SD_x)^2$
ระดับน่าพอใจ	หมายถึง	$(.05SD_x)^2 < c \leq (.10SD_x)^2$
ระดับปานกลาง	หมายถึง	$(.10SD_x)^2 < c \leq (.15SD_x)^2$
ระดับไม่น่าพอใจ	หมายถึง	$(.15SD_x)^2 < c \leq (.20SD_x)^2$
ระดับไม่น่าพอใจอย่างยิ่ง	หมายถึง	$c > (.20SD_x)^2$

ตอนที่ 4 การทดสอบความแตกต่างของผลการเทียบคะแนน

ในการทดสอบความแตกต่างของผลการเทียบคะแนนทั้ง 3 วิธี ได้แก่ คะแนนที่ได้จากการเทียบความคลาดเคลื่อนและความพึงพอใจของการเทียบคะแนน โดยใช้สถิติวิเคราะห์ คือ

t – test แบบ t - Dependent (ล้วน สายยศ และอังคณา สายยศ.2540 : 291) ซึ่งมีสูตรดังนี้

$$t = \frac{\sum D}{\sqrt{\frac{\sum D^2 - (\sum D)^2}{n - 1}}} \quad \text{เมื่อ } df = n - 1$$

กำหนดให้ D เป็นความแตกต่างของคะแนนแต่ละคู่

n เป็นจำนวนคู่

$\sum D^2$ เป็นผลรวมของ D แต่ละตัวยกกำลังสอง

$(\sum D)^2$ เป็นการเอาผลรวมของ D ทั้งหมดมายกกำลังสอง

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการเสนอผลการวิเคราะห์ข้อมูล ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูลดังนี้

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

$\bar{X}$	แทน	คะแนนเฉลี่ย
N	แทน	จำนวนนักเรียนในกลุ่มตัวอย่าง
n	แทน	จำนวนข้อในแบบทดสอบ
k	แทน	จำนวนคู่ที่ใช้ในการทดสอบความแตกต่าง
SD	แทน	ค่าเบี่ยงเบนมาตรฐาน
θ	แทน	ค่าความสามารถของผู้สอบ
θ_1	แทน	ค่าความสามารถของผู้สอบที่ทำแบบทดสอบฉบับ X
θ_2	แทน	ค่าความสามารถของผู้สอบที่ทำแบบทดสอบฉบับ Y
a	แทน	ค่าพารามิเตอร์อำนาจจำแนกของแบบทดสอบ
b	แทน	ค่าพารามิเตอร์ความยากของแบบทดสอบ
c	แทน	ค่าพารามิเตอร์สัมประสิทธิ์การเดาของแบบทดสอบ
SEE	แทน	ค่าความคลาดเคลื่อน
t-test	แทน	ค่าการทดสอบค่าที่ แบบ t - dependent
df	แทน	ชั้นของความเป็นอิสระ
C	แทน	ดัชนีความเพียงพอของการเทียบคะแนน
Y^*	แทน	คะแนนที่ผ่านการแปลงให้อยู่ในมาตราเดียวกัน
$\bar{X}_T$	แทน	ค่าเฉลี่ยของคะแนน จากวิธีเทียบคะแนนจริง
$\bar{X}_R$	แทน	ค่าเฉลี่ยของคะแนน จากวิธีเทียบคะแนนสังเกต
$\bar{X}_{Eq}$	แทน	ค่าเฉลี่ยของคะแนน จากวิธีเทียบอิกวิเปอร์เซนไทล์
CEEB	แทน	คะแนนมาตรฐาน CEEB = 500+100z

การเสนอผลการวิเคราะห์ข้อมูล

ในการศึกษาครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับดังนี้

1. ค่าสถิติและคุณภาพของแบบทดสอบ
2. การสร้างตารางเทียบคะแนนของแบบทดสอบ
3. คะแนนที่ได้จากการเทียบคะแนน
4. ความคลาดเคลื่อนของการเทียบคะแนน
5. ความเพียงพอของการเทียบคะแนน

ผลการวิจัย

1. ค่าสถิติพื้นฐานของแบบทดสอบ

จากการนำแบบทดสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่องสมการ จำนวน 2 ฉบับ คือ ฉบับ X และฉบับ Y ฉบับละ 30 ข้อ ไปทดสอบกับนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 1,062 คน ได้ค่าสถิติพื้นฐานของคะแนนดังแสดงในตาราง 2

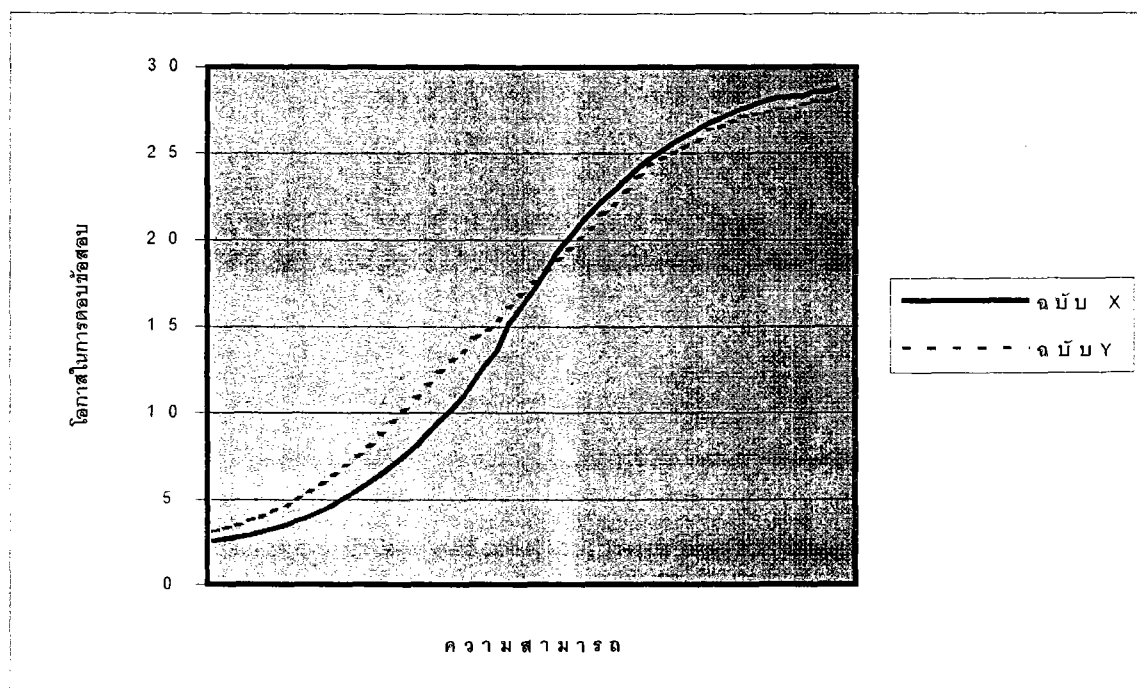
ตาราง 2 แสดงค่าสถิติพื้นฐานของแบบทดสอบวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 เรื่องสมการ

แบบทดสอบ	คะแนน เต็ม	$\bar{X}$	SD
X	30	17.812	4.429
Y	30	17.504	4.741

จากตาราง 2 แสดงให้เห็นว่า แบบทดสอบฉบับ X และฉบับ Y มีค่าเฉลี่ยเท่ากับ 17.812 กับ 17.504 ตามลำดับ ซึ่งมีค่าเกินครึ่งหนึ่งของคะแนนเต็ม แสดงให้เห็นว่า แบบทดสอบค่อนข้างง่าย เมื่อพิจารณาการกระจายของคะแนน พบว่า มีค่าความเบี่ยงเบนมาตรฐาน เท่ากับ 4.429 กับ 4.741 ตามลำดับ แสดงให้เห็นว่า แบบทดสอบทั้งสองฉบับมีการกระจายของคะแนนใกล้เคียงกัน

ในการหาคุณภาพแบบทดสอบ ผู้วิจัยนำค่าพารามิเตอร์ประจำแต่ละข้อของแบบทดสอบทั้งสองฉบับ มาหาค่าโอกาสในการตอบถูก และค่าอินฟอร์เมชันของแบบทดสอบที่ระดับความสามารถต่างๆ ดังนี้

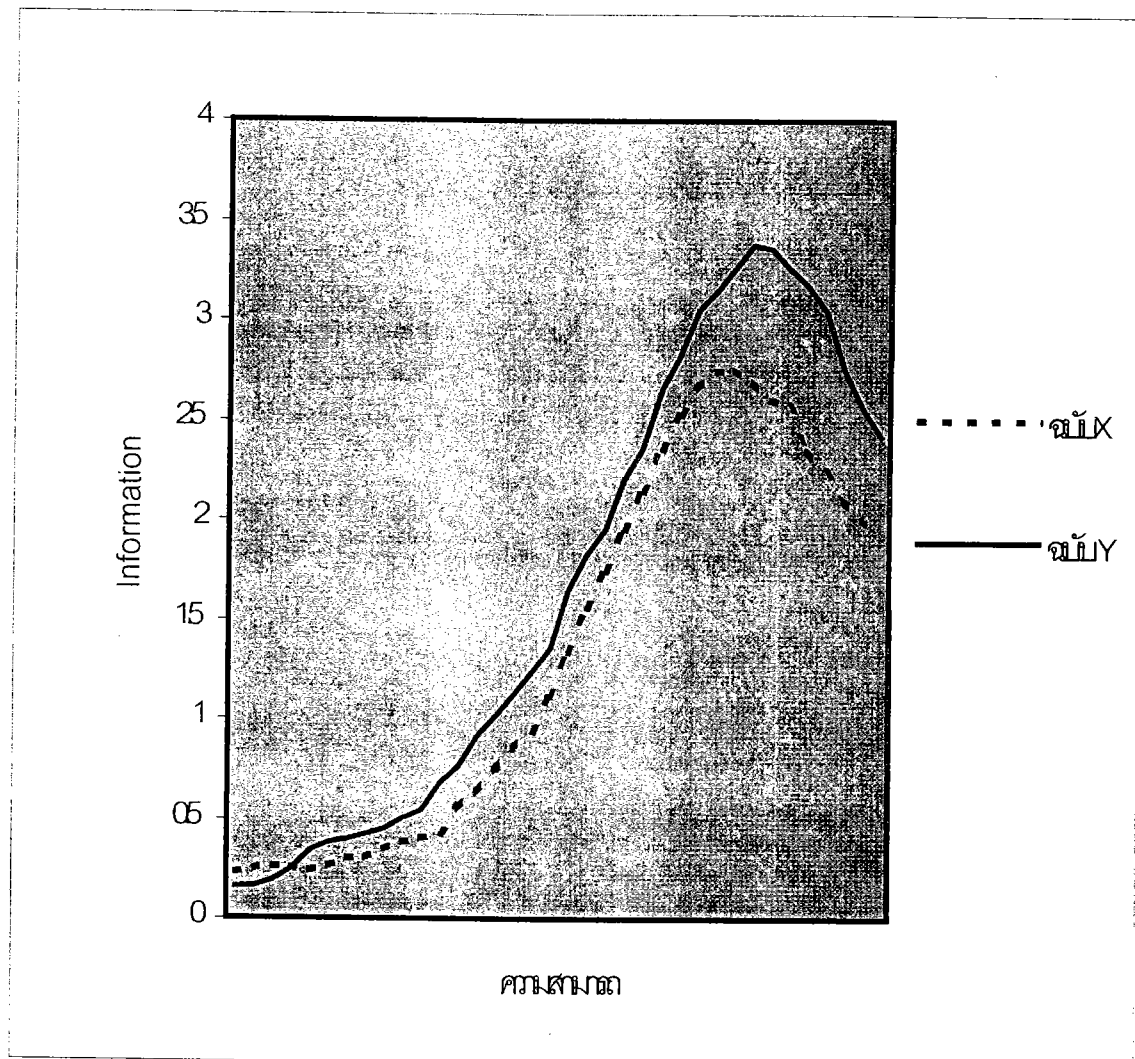
1. โอกาสในการตอบข้อสอบถูก ณ ระดับความสามารถต่างๆ แล้วนำมาเขียนโค้งลักษณะเฉพาะของแบบทดสอบ ซึ่งแสดงได้ในภาพประกอบ 7



ภาพประกอบ 7 แสดงโค้งลักษณะเฉพาะของแบบทดสอบทั้งสองฉบับ

จากภาพประกอบ 7 แสดงให้เห็นว่า ที่ระดับความสามารถ -3.0 ถึง 0.0 โอกาสในการตอบข้อสอบถูกของแบบทดสอบ ฉบับ Y สูงกว่าฉบับ X และที่ระดับความสามารถ 1.0 ถึง 3.0 แบบทดสอบทั้งสองฉบับมีโอกาสในการตอบข้อสอบถูกที่ใกล้เคียงกัน

2. ค่าสูงสุดของอินฟอร์เมชันของแบบทดสอบ ณ ระดับความสามารถต่างๆ แล้วนำมาเขียนโค้งอินฟอร์เมชันของแบบทดสอบที่ระดับความสามารถต่างๆ จะได้ดังภาพประกอบ 8



ภาพประกอบ 8 โค้งอินฟอร์เมชันของแบบทดสอบทั้งสองฉบับที่ระดับความสามารถต่างๆ

จากภาพประกอบ 8 แสดงให้เห็นว่า ที่ระดับความสามารถ -3.0 ถึง -2.0 แบบทดสอบทั้งสองฉบับให้ค่าอินฟอร์เมชันของแบบทดสอบที่ใกล้เคียงกัน แต่เมื่อพิจารณาที่ระดับความสามารถตั้งแต่ -2.0 ถึง 3.0 แล้วพบว่าฉบับ Y ให้ค่าอินฟอร์เมชันของแบบทดสอบสูงกว่าฉบับ X

2. การสร้างตารางการเทียบคะแนนแบบทดสอบ

2.1 ผลการสร้างตารางเทียบคะแนนจริง

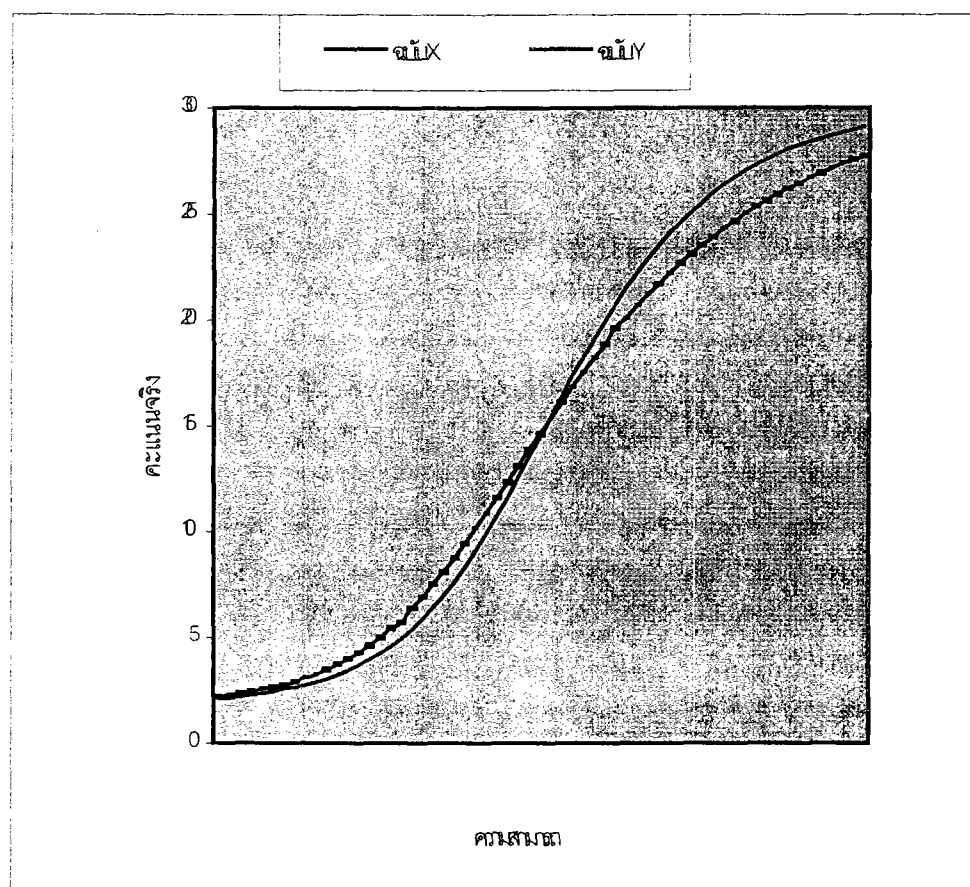
ในการสร้างตารางเทียบคะแนนของแบบทดสอบได้ดำเนินการ โดยนำผลการวิเคราะห์ค่าพารามิเตอร์ต่างๆ ไปใช้กับสูตรของวอร์ม (Warm.1978 : 114-115) ตามหลักการแปลงคะแนนในเชิงเส้นตรง และหาค่าคงที่ในการแปลงคะแนนดังสมการ

$$\theta_1 = 2.436\theta_2 - 0.1520$$

$$b_1 = 2.436b_2 - 0.1520$$

$$a_1 = 0.4109a_2$$

จากนั้นผู้วิจัยได้นำค่าพารามิเตอร์ที่แปลงให้อยู่ในสเกลเดียวกันไปคำนวณหาค่าคะแนนจริงของผู้สอบที่ระดับความสามารถต่างๆ แล้วนำมาเขียนกราฟ ดังแสดงในภาพประกอบ 4



ภาพประกอบ 4 คะแนนจริงของผู้สอบแบบทดสอบทั้งสองฉบับที่ระดับความสามารถต่างๆ

จากกราฟ แสดงคะแนนจริงของผู้สอบแบบทดสอบทั้งสองฉบับที่ระดับความสามารถต่างๆ สามารถอ่านค่าได้เป็นคะแนนสมมูลของการเทียบคะแนนจากแบบทดสอบ X ไปสู่แบบทดสอบ Y ที่ระดับความสามารถเดียวกัน และเนื่องจากคะแนนที่ได้มีค่าใกล้เคียงกัน จึงต้องขยายคะแนนด้วยวิธีการแปลงเชิงเส้นตรง (Linear Transformation) ให้เป็นคะแนนมาตรฐาน CEEB เพื่อที่จะให้เห็นคะแนนที่มีค่าแตกต่างกัน จะได้ตารางเทียบคะแนนดังนี้

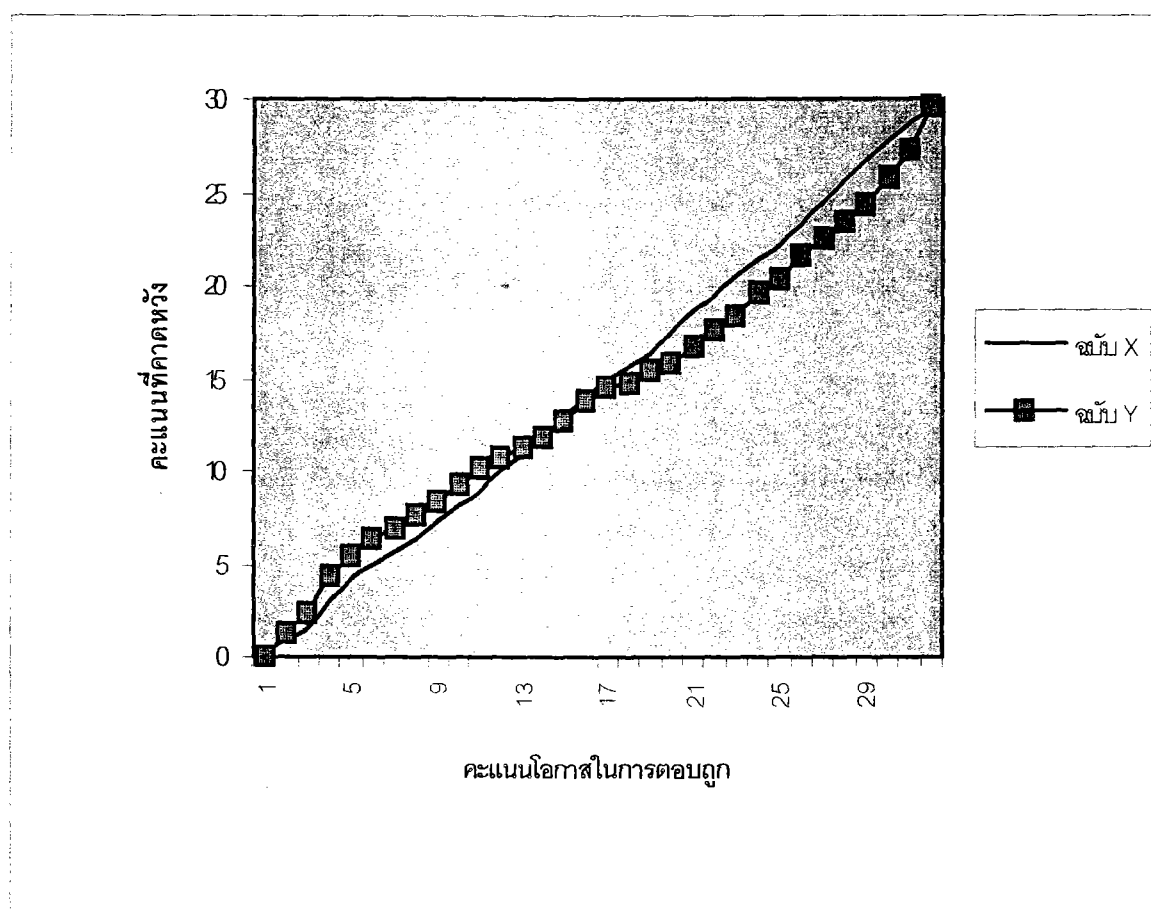
ตาราง 3 คะแนนสมมูลของการเทียบคะแนน จากแบบทดสอบฉบับ X ไปสู่ฉบับ Y ที่ระดับคะแนนและความสามารถเดียวกันตามวิธีเทียบคะแนนจริง

คะแนน	θ	X	Y*	CEEB _X	CEEB _Y *
4	-3	2.1016	2.1865	383	378
5	-2.8	2.2053	2.3237	384	380
6	-2.1	2.8892	3.2427	390	388
7	-2.0	3.0492	3.4578	391	391
8	-1.9	3.2316	3.7015	393	393
9	-1.8	3.4391	3.9766	395	396
10	-1.7	3.6748	4.2857	397	399
11	-1.5	4.2447	5.0146	402	406
12	-1.2	5.4016	6.4026	412	419
13	-0.7	8.3563	9.4556	439	449
14	-0.5	9.9233	10.8753	453	463
15	0.2	16.402	16.0956	511	513
16	0.3	17.3332	16.8109	519	520
17	0.4	18.2395	17.5071	528	527
18	0.5	19.1145	18.1812	535	533
19	0.6	19.9530	18.8314	543	540
20	0.8	21.5063	20.0544	557	552
21	1.4	25.1075	23.0924	589	581
22	1.9	27.0296	24.9927	606	600
23	2.0	27.3204	25.3162	609	603
24	2.3	28.0434	26.1886	615	611
25	2.4	28.2409	26.4489	617	614
26	2.5	28.5810	26.6950	619	616
27	2.7	28.7267	27.1465	622	620
28	2.8	28.8581	27.3529	623	622
29	2.9	28.9764	27.547	624	624
30	3.0	29.0829	27.7293	625	626

จากตาราง 3 แสดงให้เห็นว่า ที่ระดับความสามารถตั้งแต่ -3.0 ถึง 3.0 คะแนนจริงของแบบทดสอบ X มีค่าอยู่ระหว่าง 2.1016 ถึง 29.0829 และเมื่อแปลงคะแนนเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 383 ถึง 625 ส่วนคะแนนจริงของแบบทดสอบ Y มีค่าอยู่ระหว่าง 2.1865 ถึง 27.7293 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 378 ถึง 626

2.2 ผลการสร้างตารางเทียบคะแนนสังเกต

ในการสร้างตารางเทียบคะแนนของแบบทดสอบ ได้ดำเนินการโดยนำผลการวิเคราะห์ค่าพารามิเตอร์ต่างๆ ไปใช้กับสูตรของลอร์ด (Lord.1980:196-197; Hambieton and Swaminathan . 1985 : 215-216) เพื่อที่จะคำนวณหาค่าคะแนนสังเกตของผู้สอบที่ระดับคะแนนโอกาสในการตอบข้อสอบถูก และนำมาเขียนกราฟ ดังแสดงในภาพประกอบ 5



ภาพประกอบ 5 คะแนนที่คาดหวังของผู้สอบแบบทดสอบทั้งสองฉบับที่คะแนนโอกาสในการตอบถูก

จากกราฟ แสดงคะแนนสังเกตที่ได้จากผู้สอบแบบทดสอบทั้งสองฉบับที่ระดับคะแนนต่างๆ สามารถอ่านค่าได้ คะแนนสมมูลของการเทียบคะแนน โดยเทียบจากแบบทดสอบฉบับ X ไปสู่ฉบับ Y ที่ระดับคะแนนเดียวกัน ค่าความสามารถ เดียวกัน และเนื่องจากคะแนนที่ได้มีค่าคะแนนใกล้เคียงกัน จึงต้องขยายคะแนนด้วยวิธีการแปลงคะแนนเชิงเส้นตรง (Linear Transformation) ให้เป็นคะแนนมาตรฐาน CEEB เพื่อที่จะให้เห็นคะแนนที่มีค่าแตกต่างกันจะได้ตารางเทียบคะแนนดังนี้

ตาราง 4 คะแนนสมมูลของการเทียบคะแนน จากแบบทดสอบฉบับ X ไปสู่แบบทดสอบฉบับ Y ที่ระดับคะแนนและความสามารถเดียวกันตามวิธีเทียบคะแนนสังเกต

คะแนน	θ	X	Y*	CEEBS _x	CEEBS _y *
4	-3	3.1274	3.6431	370	365
5	-2.8	3.6453	3.9874	375	360
6	-2.1	5.3121	6.1532	392	394
7	-2.0	6.5942	7.9363	406	412
8	-1.9	7.4891	8.5310	415	419
9	-1.8	9.6436	10.4677	437	440
10	-1.7	10.6471	11.5931	447	453
11	-1.5	11.8707	12.7330	460	465
12	-1.2	12.6973	13.9584	468	479
13	-0.7	13.8644	14.5335	480	485
14	-0.5	14.3540	15.5617	485	496
15	0.2	15.7461	16.7978	500	510
16	0.3	16.1459	17.4538	504	517
17	0.4	17.3491	17.4659	516	517
18	0.5	18.8211	18.6762	531	531
19	0.6	19.1479	18.1611	534	525
20	0.8	20.6544	19.2643	550	537
21	1.4	21.7382	19.3887	561	538
22	1.9	22.6120	21.6740	570	564
23	2.0	23.2124	23.1073	576	579
24	2.3	24.5781	23.6675	590	586
25	2.4	25.8973	24.3381	604	593
26	2.5	26.8800	25.4433	614	605
27	2.7	27.6752	27.7664	622	631
28	2.8	28.2375	28.4003	628	638
29	2.9	29.6774	29.7466	643	656
30	3.0	29.8366	29.9639	644	655

จากตาราง 4 แสดงให้เห็นว่า ที่ระดับความสามารถตั้งแต่ - 3 ถึง 3 คะแนนสังเกตของแบบทดสอบฉบับ X มีค่าอยู่ระหว่าง 3.1271 ถึง 29.8366 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 334 ถึง 644 ส่วนคะแนนสังเกตของแบบทดสอบ Y มีค่าอยู่ระหว่าง 3.6431 ถึง 29.9639 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่ามีค่าอยู่ระหว่าง 326 ถึง 655

2.3 ผลการสร้างตารางเทียบคะแนนวิธีอีควิเปอร์เซนไทล์

ตาราง 5 คะแนนสมมูลของการเทียบคะแนน จากแบบทดสอบฉบับ X ไปสู่ฉบับ Y ที่ระดับคะแนนและตำแหน่งเปอร์เซนไทล์ที่เท่ากัน

เปอร์เซนไทล์	คะแนนจากแบบทดสอบ					
	X	Z _X -Score	CEEBS _X	Y*	Z _Y -Score	CEEBS _Y *
0.38	4	-1.4933	351	5	-1.4501	355
0.5	5	-1.3720	363	5.1	-1.4378	357
1	6	-1.2509	375	6	-1.3260	368
2	7	-1.1296	388	7	-1.2017	380
3	8	-1.0084	400	8	-1.0775	393
5	9	-0.8872	412	9.6	-0.8788	413
7	10	-0.7660	424	10.3	-0.7918	421
9.5	10.8	-0.6690	434	10.8	-0.7269	428
10.5	11	-0.6448	436	11.2	-0.6800	432
14.5	12	-0.5236	448	12.2	-0.5557	445
16.5	13	-0.4020	460	12.6	-0.5060	450
20	14	-0.2812	472	13	-0.4563	455
25	15	0.1600	484	14	0.1103	489
32	16	0.0387	503	15	0.0430	496
42	17	0.0824	508	16.2	0.0035	500
52	18	0.2036	520	18	0.1647	516
61	19	0.3248	532	19.2	0.3138	531
71	20	0.4460	544	20.8	0.5125	551
80.8	21	0.5673	556	22.1	0.6740	567
88	22	0.6884	568	23.2	0.8107	581
93	23	0.8096	580	25.2	0.7762	577
95	23.8	0.9.66	590	26.0	1.1585	615
96	24	0.9309	593	26.4	1.2081	620
97	25	1.0521	605	27.2	1.3076	630
99.4	28	1.4157	641	28	1.4069	640
99.8	29	1.5369	653	29.3	1.5684	656
100	30	1.6582	665	30	1.6554	665

จากตาราง 5 แสดงให้เห็นว่า ที่ตำแหน่งเปอร์เซ็นต์ที่ 0.38 ถึง 100 คะแนนจากแบบทดสอบฉบับ X มีคะแนนอยู่ระหว่าง 4 ถึง 30 คะแนน และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 351 ถึง 665 ส่วนคะแนนจากแบบทดสอบฉบับ Y มีค่าอยู่ระหว่าง 5 ถึง 30 และเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้วพบว่า มีค่าอยู่ระหว่าง 355 ถึง 665

3. คะแนนที่ได้จากการเทียบคะแนน

3.1 จากคะแนนสมมูลที่ได้จากการสร้างตารางเทียบคะแนนทั้งสามวิธี ผู้วิจัยได้นำเสนอในระดับคะแนนและ ความสามารถเดียวกัน

ตาราง 6 คะแนนสมมูลของการเทียบคะแนนทั้งสามวิธี ที่ระดับคะแนน และความสามารถเดียวกัน

คะแนน	θ	การเทียบคะแนนจริง		การเทียบคะแนน สังเกต		การเทียบคะแนนอิกวิ เปอร์เซ็นต์ไทล์	
		Y*	CEEB	Y*	CEEB	Y*	CEEB
4	-3	2.1865	383	3.6431	370	5	355
5	-2.8	2.3237	384	3.9874	375	5.1	357
6	-2.1	3.2427	390	6.1532	392	6	368
7	-2.0	3.4578	391	7.9363	406	7	380
8	-1.9	3.7015	393	8.5310	415	8	393
9	-1.8	3.9766	395	10.4677	437	9.6	413
10	-1.7	4.2857	397	11.5931	447	10.3	421
11	-1.5	5.0146	402	12.7330	460	10.8	428
12	-1.2	6.4026	412	13.9584	468	11.2	432
13	-0.7	10.8753	439	14.5335	480	12.2	445
14	-0.5	16.0956	453	15.5617	485	12.6	450
15	0.2	16.0956	511	16.7978	500	13	455
16	0.3	16.8109	519	17.4538	504	14	489
17	0.4	17.5071	528	17.4659	516	15	496
18	0.5	18.1812	535	18.6762	531	16.2	500
19	0.6	18.8314	543	18.1611	534	18	516
20	0.8	20.0544	557	19.2643	550	19.2	531
21	1.4	23.0924	589	19.3887	561	20.8	551
22	1.9	24.9927	606	21.6740	570	22.1	567
23	2.0	25.3162	609	23.1073	576	23.2	581
24	2.3	26.1886	615	23.6675	590	25.2	577
25	2.4	26.4489	617	24.3381	604	26.0	615
26	2.5	26.6950	619	25.4433	614	26.4	620

ตาราง 6 (ต่อ)

คะแนน	θ	การเทียบคะแนนจริง		การเทียบคะแนน สังเกต		การเทียบคะแนนอิกวิ เปอร์เซนไทล์	
		Y*	CEEB	Y*	CEEB	Y*	CEEB
27	2.7	27.1465	622	27.7664	622	27.2	630
28	2.8	27.3529	623	28.4003	628	28	640
29	2.9	27.5470	624	29.7446	643	29.3	656
30	3.0	27.7293	625	29.9639	644	30	665

จากตาราง 6 แสดงให้เห็นว่าคะแนนสมมูลที่ได้จากการเทียบคะแนนจากแบบทดสอบ X ไปสู่แบบทดสอบ Y ช่วงคะแนน 4 ถึง 13 วิธีเทียบคะแนนสังเกตให้ความสมมูลสูงกว่าวิธีการเทียบคะแนนจริงและอิกวิเปอร์เซนไทล์ , ในช่วงคะแนน 20 ถึง 24 วิธีเทียบคะแนนจริงให้คะแนนสมมูลสูงกว่าคะแนนสังเกต และ อิกวิเปอร์เซนไทล์ ส่วนช่วงคะแนน 14 ถึง 19 และ 25 ถึง 30 ทั้งสามวิธีมีคะแนนสมมูลใกล้เคียงกัน

3.2 การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบคะแนนทั้งสามวิธี

จากคะแนนที่ได้จากการเทียบคะแนนทั้งสามวิธี สามารถนำมาทดสอบความแตกต่างของคะแนนสมมูล ดังแสดงในตาราง 7

ตาราง 7 การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบคะแนนจริงและคะแนนสังเกต , และวิธีอิกวิเปอร์เซนไทล์

	$\bar{X}_T = 15.734$	$\bar{X}_R = 17.423$	$\bar{X}_{Eq} = 16.70$
$\bar{X}_T = 15.734$	-	t = 2.29*	t = 0.99
$\bar{X}_R = 17.423$		-	t = 0.98
$\bar{X}_{Eq} = 16.70$			-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 7 แสดงให้เห็นว่า ความแตกต่างของคะแนนที่เทียบได้จากวิธีเทียบคะแนนจริงกับคะแนนสังเกต มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนวิธีเทียบคะแนนจริงกับวิธีอิกวิเปอร์เซนไทล์ และวิธีเทียบคะแนนสังเกตกับวิธีอิกวิเปอร์เซนไทล์ มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

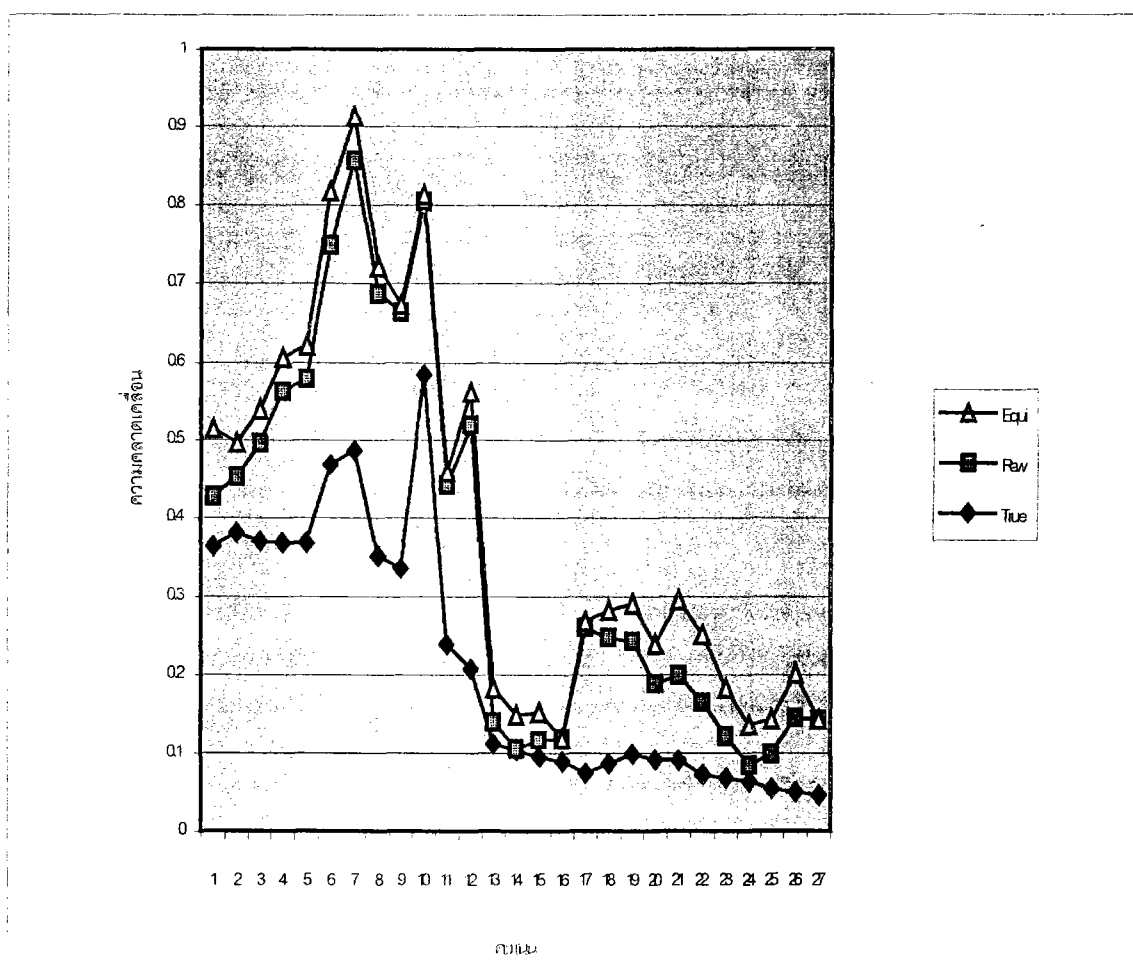
4. ความคลาดเคลื่อนจากการเทียบคะแนน

4.1 จากการเทียบคะแนน เมื่อนำมาเปรียบเทียบกับคะแนนที่ได้จากการสอบ แล้วนำมาคำนวณหาค่าความคลาดเคลื่อนในแต่ละระดับคะแนน ดังตาราง 8

ตาราง 8 ความคลาดเคลื่อนของการเทียบคะแนนทั้งสามวิธี

คะแนน	θ	SEE คะแนนจริง	SEE คะแนนสังเกต	SEE อิกวิเปอร์เซนไทล์
4	-3.0	0.364	0.063	0.087
5	-2.8	0.381	0.072	0.043
6	-2.1	0.370	0.126	0.043
7	-2.0	0.368	0.194	0.043
8	-1.9	0.368	0.210	0.043
9	-1.8	0.468	0.28	0.069
10	-1.7	0.486	0.371	0.056
11	-1.5	0.351	0.335	0.035
12	-1.2	0.335	0.328	0.009
13	-0.7	0.584	0.220	0.009
14	-0.5	0.239	0.203	0.017
15	0.2	0.208	0.31	0.043
16	0.3	0.111	0.028	0.043
17	0.4	0.103	0.002	0.043
18	0.5	0.094	0.022	0.035
19	0.6	0.089	0.029	0.000
20	0.8	0.074	0.185	0.009
21	1.4	0.086	0.161	0.035
22	1.9	0.098	0.144	0.048
23	2.0	0.091	0.096	0.052
24	2.3	0.090	0.109	0.095
25	2.4	0.072	0.092	0.087
26	2.5	0.067	0.054	0.061
27	2.7	0.062	0.021	0.052
28	2.8	0.053	0.046	0.043
29	2.9	0.049	0.096	0.056
30	3.0	0.045	0.097	0.000
$\bar{X}$		0.0425	0.1326	0.1954
SD		0.1856	0.1283	3.8266

จากตาราง 8 แสดงให้เห็นว่า ค่าความคลาดเคลื่อนรายข้อของการเทียบคะแนนจริง มีค่าอยู่ระหว่าง 0.045 ถึง 0.584 , วิธีการเทียบคะแนนสังเกตมีค่าอยู่ระหว่าง 0.002 ถึง 0.371 ส่วนวิธีเทียบอิกวิเปอร์เซนไทล์ มีค่าอยู่ระหว่าง 0.000 ถึง 0.095 และเมื่อพิจารณาค่าเฉลี่ยของความคลาดเคลื่อน พบว่า วิธีเทียบคะแนนจริง, วิธีเทียบคะแนนสังเกตและวิธีอิกวิเปอร์เซนไทล์ มีค่าเฉลี่ย เท่ากับ 0.19542 , 0.1326 และ 0.0425 ตามลำดับ ส่วนค่าความเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนของวิธีเทียบคะแนนจริงมีค่าเท่ากับ 3.8266 รองลงมาคือวิธีเทียบอิกวิเปอร์เซนไทล์มีค่าเท่ากับ 0.1856 และวิธีเทียบคะแนนสังเกตมีค่าเท่ากับ 0.1283 ซึ่ง วิธีการเทียบคะแนนจริงมีการกระจายของคะแนนมากกว่าวิธีอิกวิเปอร์เซนไทล์ และ วิธีเทียบคะแนนสังเกตตามลำดับ และเพื่อให้เห็นภาพที่ชัดเจน ผู้วิจัยได้นำค่าความคลาดเคลื่อนที่ได้จากการเทียบคะแนนไปเขียนกราฟ ได้ดังภาพประกอบ 6



ภาพประกอบ 6 ความคลาดเคลื่อนของการเทียบคะแนนทั้งสามวิธีที่ระดับคะแนนเดียวกัน

จากภาพประกอบ 6 แสดงให้เห็นว่า ที่ระดับคะแนนตั้งแต่ 1 ถึง 13 และ 15 ถึง 30 วิธีเปรียบเทียบคะแนนจริงให้ค่าความคลาดเคลื่อนน้อยกว่าวิธีการเทียบคะแนนสังเกตและวิธีอีควิเปอร์เซนไทล์ ส่วนระดับคะแนนที่ 1 ถึง 30 วิธีเทียบคะแนนสังเกตให้ค่าความคลาดเคลื่อนน้อยกว่าวิธีเทียบคะแนนวิธีอีควิเปอร์เซนไทล์

4.2 การเปรียบเทียบความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้งสามวิธี

จากความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้งสามวิธี สามารถนำมาทดสอบความแตกต่างของคะแนนความคลาดเคลื่อน ดังแสดงในตาราง 9

ตาราง 9 การเปรียบเทียบความแตกต่างของคะแนนความคลาดเคลื่อนที่ได้จากวิธีเทียบคะแนนจริง,คะแนนสังเกต และวิธีอีควิเปอร์เซนไทล์

	$\bar{X}_T=0.0428$	$\bar{X}_R=0.1326$	$\bar{X}_{Eq}=0.1954$
$\bar{X}_T=0.0428$	-	t =0.098	t =3.36**
$\bar{X}_R=0.1326$		-	t =0.79
$\bar{X}_{Eq}=0.1954$			-

** มีนัยสำคัญทางสถิติที่ระดับ . 01

จากตาราง 9 แสดงให้เห็นว่า คะแนนความคลาดเคลื่อนของวิธีเทียบคะแนนจริงกับวิธีอีควิเปอร์เซนไทล์มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ . 01 ส่วนวิธีเทียบคะแนนจริงกับคะแนนสังเกต และวิธีเทียบคะแนนสังเกตกับวิธีอีควิเปอร์เซนไทล์มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

5. ความเพียงพอและการเปรียบเทียบความเพียงพอทั้งสามวิธีของการเทียบคะแนน

ความเพียงพอของการเทียบคะแนนทั้งสามวิธี ในการหาความเพียงพอของการเทียบคะแนน ผู้วิจัยได้นำแบบทดสอบทั้งสองฉบับ ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการหาผลการเทียบคะแนน แล้วนำมาคำนวณหาความคลาดเคลื่อนที่ได้จากการเทียบคะแนน เพื่อที่จะสามารถนำไปประเมินความเพียงพอของการเทียบคะแนน ดังแสดงในตาราง 10

ตาราง 10 ความเพียงพอและการเปรียบเทียบความเพียงพอของการเทียบคะแนนทั้งสามวิธี
ที่ระดับคะแนนเดียวกัน

คะแนน	θ	$(x-x^*)^2$ คะแนนจริง	$(x-x^*)^2$ คะแนนสังเกต	$(x-x^*)^2$ อคติเปอร์เซ็นต์ไทล์
4	-3.0	0.132	0.147	0.056
5	-2.8	0.145	0.341	0.135
6	-2.1	0.136	0.331	0.214
7	-2.0	0.135	0.233	0.137
8	-1.9	0.132	0.002	0.114
9	-1.8	0.236	0.070	0.107
10	-1.7	0.123	0.272	0.312
11	-1.5	0.112	0.435	0.316
12	-1.2	0.341	0.367	0.221
13	-0.7	0.057	0.174	0.619
14	-0.5	0.043	0.009	0.473
15	0.2	0.212	0.346	0.419
16	0.3	0.010	0.255	0.316
17	0.4	1.009	0.164	0.254
18	0.5	0.038	0.313	0.281
19	0.6	0.343	0.052	0.238
20	0.8	0.441	0.477	0.466
21	1.4	0.534	0.331	0.381
22	1.9	0.067	0.067	0.007
23	2.0	0.641	0.462	0.124
24	2.3	0.664	0.440	0.201
25	2.4	0.686	0.217	0.202
26	2.5	0.773	0.368	0.415
27	2.7	0.886	0.346	0.227
28	2.8	0.891	0.414	0.387
29	2.9	0.763	0.083	0.289
30	3.0	0.674	0.147	0.316
K		27	27	27
SD		4.5870	5.4582	5.0771
c		0.0191	0.0086	0.0112
$(.05SD)^2$		0.0998	0.7448	0.0644

จากตาราง 10 แสดงค่าความเพียงพอและการเปรียบเทียบความเพียงพอของการเทียบคะแนนทั้งสามวิธี พบว่า วิธีเทียบคะแนนคะแนนจริง,วิธีเทียบคะแนนสังเกต,วิธีเทียบอิกิเปอร์เซนไทล์มีความเพียงพอเท่ากับ 0.0191 , 0.0086 และ 0.0112 ตามลำดับ โดยวิธีเทียบคะแนนจริงมีความเพียงพอสูงสุด รองลงมาคือวิธีการเทียบอิกิเปอร์เซนไทล์ และต่ำที่สุดคือ วิธีเทียบคะแนนสังเกต แต่เมื่อเทียบกับเกณฑ์แล้ว ทั้งสามวิธีให้ค่าความเพียงพอของการเทียบคะแนนอยู่ในระดับที่น่าพอใจยิ่ง

บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยมีความมุ่งหมายเพื่อศึกษาเปรียบเทียบผลการเทียบคะแนนแบบทดสอบวิชาคณิตศาสตร์ จำนวน 2 ฉบับด้วยวิธีการเทียบคะแนนจริง และการเทียบคะแนนสังเกตตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์ โดยมีจุดมุ่งหมายเฉพาะดังนี้

1. เพื่อเปรียบเทียบคะแนนที่เทียบได้จากแบบทดสอบวิชาคณิตศาสตร์ จากวิธีการเทียบคะแนนจริง คะแนนสังเกต ตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์
2. เพื่อเปรียบเทียบความคลาดเคลื่อนของการเทียบคะแนนจริง คะแนนสังเกตตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์
3. เพื่อเปรียบเทียบความเพียงพอของการเทียบคะแนนจริง คะแนนสังเกตตามทฤษฎีการตอบข้อคำถาม และการเทียบคะแนนตามวิธีอิกวิเปอร์เซนไทล์

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2542 สังกัด สำนักงานการประถมศึกษาจังหวัดหนองคาย โดยวิธีการสุ่มแบบสองขั้นตอน (Two – Stage Random Samplig) ได้แก่การสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) และการสุ่มอย่างง่าย (Simple Random Sampling) มีห้องเรียนเป็นหน่วยการสุ่ม (Sampling Unit) ได้กลุ่มตัวอย่างจำนวน 1,062 คน และกลุ่มทวนผล 523 คน จาก 47 โรงเรียน

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นแบบทดสอบแบบเลือกตอบ วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่อง สมการ ชนิด 5 ตัวเลือก ตามหลักสูตรประถมศึกษา 2521(ฉบับปรับปรุง พ.ศ. 2533) จำนวน 2 ฉบับ ๆ ละ 30 ข้อโดยพิจารณาค่าความยากที่เท่าหรือเทียบเคียงกันได้แบบทดสอบทั้งสองฉบับ ดังนี้

1. แบบทดสอบวิชาคณิตศาสตร์ ฉบับ X มี 30 ข้อ
2. แบบทดสอบวิชาคณิตศาสตร์ ฉบับ Y มี 30 ข้อ

วิธีดำเนินการเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูลในครั้งนี้ ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1. วางแผนในการดำเนินการสอบ โดยติดต่อขอความร่วมมือในการนำแบบทดสอบไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่าง กำหนดวัน เวลาและสถานที่สอบ ในช่วงเดือน พฤศจิกายน พ.ศ.2542 ถึง เดือน กุมภาพันธ์ พ.ศ. 2543

2. นำแบบทดสอบทั้ง 2 ฉบับ คือฉบับ X และฉบับ Y ไปทดสอบกับกลุ่มตัวอย่าง ฉบับ X จำนวน 631 คน ฉบับ Y จำนวน 631 คน จากการสุ่มกลุ่มตัวอย่างที่ใช้ห้องเรียนเป็นหน่วยการสุ่มแล้วใช้การสุ่มจัดกลุ่มให้สมดุล (Counterbalanced Random Group Design) โดยแจกแบบทดสอบแก่คนที่ 1 ด้วยฉบับ X แล้วตามด้วยฉบับ Y และคนที่ 2 ด้วยฉบับ Y แล้วตามด้วยฉบับ X ดังเช่น XY, YX, XY, YX,... และกลุ่มทานผล 523 คน

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลดำเนินการโดยแบ่งออกเป็น 4 ตอน ได้แก่

1. ค่าสถิติพื้นฐานและคุณภาพของแบบทดสอบ
2. การสร้างตารางเทียบคะแนนแบบทดสอบ
3. คะแนนที่ได้จากการเทียบ
4. ความคลาดเคลื่อนจากการเทียบคะแนน
5. ความเพียงพอของการเทียบคะแนน

สรุปผลการวิจัย

1. คะแนนเฉลี่ยของแบบทดสอบ ฉบับ X และ แบบทดสอบฉบับ Y มีค่าเฉลี่ยเท่ากับ 17.812 กับ 17.504 ตามลำดับ ซึ่งมีค่าเกินครึ่งหนึ่งของคะแนนเต็ม แสดงให้เห็นว่า แบบทดสอบทั้งสองฉบับค่อนข้างง่าย เมื่อพิจารณาการกระจายของคะแนน พบว่า มีค่าความเบี่ยงเบนมาตรฐานเท่ากับ 4.429 กับ 4.741 ตามลำดับ แสดงว่ามีการกระจายของคะแนนที่ใกล้เคียงกัน

แบบทดสอบฉบับ X มีค่าพารามิเตอร์อำนาจจำแนกอยู่ระหว่าง 0.569 ถึง 1.744 ค่าพารามิเตอร์ความยากอยู่ระหว่าง -1.849 ถึง 1.922 และค่าสัมประสิทธิ์การเดามีค่าอยู่ระหว่าง 0.105 ถึง 0.376 ส่วนค่าอินฟอร์เมชันของแบบทดสอบมีค่าอยู่ระหว่าง 0.225 ถึง 2.745 และ ณ ระดับความสามารถที่เพิ่มขึ้น ค่าอินฟอร์เมชันของแบบทดสอบจะเพิ่มขึ้น ซึ่งจะให้ค่าสูงสุดที่ระดับความสามารถ 1.4

แบบทดสอบฉบับ Y มีค่าพารามิเตอร์อำนาจจำแนกอยู่ระหว่าง 0.559 ถึง 1.768 ค่าพารามิเตอร์ความยากมีค่าอยู่ระหว่าง -1.235 ถึง 1.845 ค่าสัมประสิทธิ์การเดามีค่าอยู่ระหว่าง 0.074 ถึง 0.262 ส่วนค่าอินฟอร์เมชันของแบบทดสอบมีค่าอยู่ระหว่าง 0.154 ถึง 3.375 และ ณ ที่ระดับความสามารถที่เพิ่มขึ้น ค่าอินฟอร์เมชันของแบบทดสอบจะเพิ่มขึ้น ซึ่งจะให้ค่าสูงสุดที่ระดับความสามารถ 1.6

2. ความแตกต่างของคะแนนที่เทียบได้ จากวิธีเทียบคะแนนจริงกับวิธีเทียบคะแนนสังเกต แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนวิธีเทียบคะแนนสังเกตและวิธีอิกวิเปอร์เซนไทล์ วิธีเทียบคะแนนจริงกับวิธีอิกวิเปอร์เซนไทล์ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. ความแตกต่างของคะแนนความคลาดเคลื่อนของวิธีเทียบคะแนนจริงกับวิธีอิกวิเปอร์เซนไทล์ พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนวิธีเทียบคะแนนจริงกับวิธีเทียบคะแนนสังเกต , วิธีเทียบคะแนนสังเกตกับวิธีอิกวิเปอร์เซนไทล์ พบว่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4. วิธีเทียบคะแนนจริงให้ความเพียงพอสูงที่สุด รองลงมาคือ วิธีเทียบคะแนนสังเกตและวิธีอิกวิเปอร์เซนไทล์ ตามลำดับ แต่เมื่อเปรียบเทียบกับเกณฑ์ของความเพียงพอแล้ว ทุกวิธีการเทียบคะแนนให้ผลเป็นที่น่าพอใจอย่างยิ่ง

อภิปรายผล

1. การเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบตามวิธีเทียบคะแนนจริง กับ คะแนนสังเกตมีความแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะการคิดคำนวณของวิธีเทียบคะแนนสังเกตซับซ้อนมากกว่าวิธีเทียบคะแนนจริงและมีจุดอ่อนที่เป็นเพียงการเทียบมาตรงอย่างประมาณเท่านั้นและวิธีเทียบคะแนนจริงนั้นให้ความหมายของคะแนนสมมูลที่อยู่เหนือค่าเฉลี่ยการเดาดังนั้นคะแนนที่เทียบได้จึงมีความแตกต่างกัน สอดคล้องกับผลการศึกษาของ ลอร์ด (Lord.1980) และ พรโสภา (พรโสภา พินิจผล. 2539) ส่วนวิธีเทียบคะแนนจริงกับวิธีอิกวิเปอร์เซนไทล์ และวิธีเทียบคะแนนสังเกตกับวิธีอิกวิเปอร์เซนไทล์ มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แต่เมื่อเปรียบเทียบความแตกต่างของคะแนนที่ได้จากการเทียบเมื่อแปลงเป็นคะแนน CEEB พบว่า วิธีการเทียบคะแนนจริงกับวิธีอิกวิเปอร์เซนไทล์ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ เพราะเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้ว ซึ่งเป็นการขยายคะแนนด้วยวิธีเชิงเส้นตรง (Linear Transformation) ทำให้เห็นความแตกต่างของคะแนนชัดเจนขึ้น ส่วนวิธีเทียบคะแนนจริงกับคะแนนสังเกต, วิธีเทียบคะแนนสังเกตกับวิธีอิกวิเปอร์เซนไทล์ มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ เพราะเมื่อแปลงเป็นคะแนนมาตรฐาน CEEB แล้ว ทำให้ช่วงการกระจายคะแนน ที่ได้จากการเทียบ มีค่าเข้าสู่คะแนนเฉลี่ยมาตรฐานซึ่งเข้าใกล้ 500 คะแนน ดังนั้นคะแนนที่ได้จากการเทียบมีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนที่ได้จากการเทียบคะแนนตามวิธีการเทียบคะแนนจริงกับวิธีอิกวิเปอร์เซนไทล์ พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ และสอดคล้องกับการศึกษาของ ฮาน ,โคเลน และโพลแมนน์ แมน (Han,Kolen and Pohlmann.1997:105 - 121) ที่พบว่าความคลาดเคลื่อนของวิธีเทียบอิกวิเปอร์เซนไทล์ที่ยังไม่ปรับให้เรียบ และวิธีเทียบคะแนนจริงให้ผลคลาดเคลื่อนมากกว่า วิธีเทียบคะแนนสังเกต และวิธีอิกวิเปอร์เซนไทล์ และอีกเหตุผลข้อหนึ่งคือ คะแนนจริงเป็นวิธีการเทียบทฤษฎีการตอบข้อสอบให้ค่าความคลาดเคลื่อนน้อยกว่าวิธีอิกวิเปอร์เซนไทล์เป็นวิธีเทียบแบบดั้งเดิมซึ่งสอดคล้องกับงานวิจัยของโคเลน (Kolen.1981:1-11) ส่วนตามวิธีเทียบคะแนนจริง กับวิธีเทียบคะแนนสังเกต มีความคลาดเคลื่อนแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะ เป็นวิธีเทียบคะแนนตามทฤษฎีการตอบข้อคำถามเหมือนกัน ที่มีการคิดคำนวณค่อนข้างละเอียดเหมือนกัน จึงมีความคลาดเคลื่อนไม่แตกต่างกัน สอดคล้องกับผลการศึกษาของ ลอร์ด(Lord.1980) ว่าทั้งสองวิธีมีความสอดคล้องกันมาก ถ้าจะสรุปอ้างอิงต้องทำอย่างพิถีพิถัน ส่วนวิธีเทียบคะแนนสังเกตกับวิธีอิกวิเปอร์เซนไทล์ มีความคลาดเคลื่อนแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้ อาจเป็นเพราะในการวิจัยครั้งนี้ใช้กลุ่มตัวอย่างที่มีขนาดใหญ่พอ จึงทำให้มีความคลาดเคลื่อนไม่แตกต่างกัน ซึ่งสอดคล้องกับ ผลการศึกษาของ โคเลนและวิทนีและคนอื่น ๆ (Kolen and Whitney.1982 : 291 ; Petersen,Cook and Stocking .1981)

3. การเปรียบเทียบความแตกต่างของความเพียงพอที่ได้จากการเทียบคะแนน ทั้งสามวิธี นั้นวิธีเทียบคะแนนจริง ให้ความเพียงพอในการเทียบคะแนนสูงกว่าวิธีเทียบคะแนนสังเกต และวิธีอิกวิเปอร์เซนไทล์ ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ และการศึกษาครั้งนี้ ให้ความเพียงพอของการเทียบคะแนนให้ผลอยู่ในเกณฑ์ที่น่าพอใจอย่างยิ่ง สอดคล้องกับผลการศึกษาของซูชีพ พงษ์สมบุรณ์(ซูชีพ พงษ์สมบุรณ์ : 2528)ที่ศึกษารูปแบบที่ใช้ผู้สอบร่วมในแนวทฤษฎีการตอบข้อคำถามว่ามีความคงที่พอๆ กันกับวิธีอิกวิเปอร์

เซนไทล์ซึ่งเป็นไปตามทฤษฎีการตอบข้อสอบที่กล่าวไว้ว่าการเทียบคะแนนตามทฤษฎีการตอบข้อสอบที่ให้ความเพียงพอของการเทียบคะแนนอยู่ในระดับที่น่าพอใจ และสอดคล้องกับการศึกษาของลอร์ด (Lord.1980 : 203) , โคลเลนและวิทนี (Kolen and Whitney 1982 : 279 - 289), เรวดี อินทะสระระ(2530) , พรโสภา พิณิจผล (2539) ศึกษาการเทียบคะแนนตามทฤษฎีการตอบข้อสอบ ที่ให้ค่าความเพียงพออยู่ในระดับที่น่าพอใจ

จะเห็นว่า วิธีการเทียบคะแนนทั้งสามวิธีให้ผลการเทียบคะแนนความคลาดเคลื่อนระหว่างวิธีเทียบคะแนนจริงกับวิธีอิกวิเปอร์เซนไทล์ต่างกันอย่างน้อยมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนวิธีเทียบคะแนนจริงกับวิธีเทียบคะแนนสังเกต, วิธีเทียบคะแนนสังเกตกับวิธีอิกวิเปอร์เซนไทล์มีความแตกต่างกันอย่างไม่มีนัยสำคัญ แต่ทั้งสามวิธีความเพียงพอของการเทียบคะแนนให้ผลเป็นที่น่าพอใจอย่างยิ่ง

ข้อเสนอแนะ

จากการศึกษาค้นคว้าครั้งนี้ ทำให้ผู้วิจัยได้แนวคิดจากการศึกษาที่จะเสนอแนะดังนี้

1. ผลการเทียบคะแนนทั้งสามวิธี พบว่า วิธีเทียบคะแนนจริงมีความคลาดเคลื่อนน้อยที่สุด รองลงมาคือวิธีเทียบคะแนนสังเกตและวิธีอิกวิเปอร์เซนไทล์ แต่ทุกวิธีต่างก็ให้ความเพียงพอในการเทียบคะแนนอยู่ในเกณฑ์ที่น่าพอใจอย่างยิ่ง ดังนั้นควรนำวิธีการเทียบคะแนนจริงไปใช้เพราะมีการคิดคำนวณที่ยุ่งยากซับซ้อนน้อยกว่าวิธีเทียบคะแนนสังเกต ส่วนวิธีอิกวิเปอร์เซนไทล์เป็นวิธีเทียบคะแนนแบบดั้งเดิม ซึ่งมีความคลาดเคลื่อนในการเทียบคะแนนพอสมควรยังคงต้องมีการตรวจสอบกันต่อไป
2. ควรนำวิธีการเทียบคะแนนจริงไปใช้ในการเทียบคะแนนจากแบบทดสอบวิชาคณิตศาสตร์ในชั้นอื่น ๆ และวิชาอื่นๆ
3. ในการเก็บรวบรวมข้อมูล ต้องควบคุมตัวแปรแทรกซ้อนอย่างพิถีพิถัน เช่น ระยะเวลาการศึกษา ครูผู้สอน รวมถึงการประมาณค่าพารามิเตอร์
4. ควรนำวิธีการเทียบคะแนนไปใช้กับการเปรียบเทียบคะแนนสอบในระดับต่างๆ เช่น การสอบคัดเลือกในระดับที่สำคัญๆ ที่ต้องการความแม่นยำในการเทียบสูง เพื่อเป็นมาตรฐานในการตัดสินผลสอบและก่อให้เกิดความยุติธรรมแก่ผู้สอบมากยิ่งขึ้น

บรรณานุกรม

บรรณานุกรม

- ชูชีพ พงษ์สมบุญ. (2528). การเปรียบเทียบการเทียบมาตรฐานระหว่างรูปแบบที่ใช้ผู้สอบร่วมกับรูปแบบที่ใช้แบบสอบร่วม. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- ชูศักดิ์ ชัมภลลิขิต. (2527). *Test Equating Procedures*. หน้า 143-150. เอกสารประกอบการบรรยายในการสัมมนาระดับชาติครั้งที่ 4 โครงการพัฒนาอาเซียน. กรุงเทพฯ : ม.ป.พ.
- ทิพาภรณ์ ทุงฤทธิ์. (2544). การเปรียบเทียบผลการเทียบคะแนนโดยใช้วิธีเชิงเส้นตรง เมื่อใช้วิธีการแปลงคะแนนที่แตกต่างกันในแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- พรพิมล นาคเวช. (2537). การศึกษาคุณภาพของการเทียบคะแนนในแนวดิ่งโดยใช้ทฤษฎีการตอบสนองข้อสอบ. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- พรโสภา พินิจผล. (2539). ผลการเทียบคะแนนแบบทดสอบวิชาการพยาบาลสูติศาสตร์ตามทฤษฎีการตอบสนองที่ใช้วิธีการวิเคราะห์ต่างกัน. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ภาวณี ศรีสุขวัฒนานนท์. (2529). การเปรียบเทียบผลจากการใช้รูปแบบการเทียบมาตรฐานที่ต่างกันเมื่อแบบสอบร่วมมีความยาวต่างกัน. วิทยานิพนธ์ ค.ด. กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- เรวดี อินทสระ. (2530). การเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตรฐานระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ล้วน สายยศ และอังคณา สายยศ. (2540). สถิติวิทยาทางการวิจัย. พิมพ์ครั้งที่ 3 . กรุงเทพฯ : สุวีริยาสาส์น.
- .(2527). หลักการสร้างแบบทดสอบความถนัดทางการเรียน. กรุงเทพฯ: วัฒนาพานิช.
- วรรณดี แสงประทีปทอง. (2538ก). การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนวิชาภาษาไทยของนักเรียนชั้นมัธยมศึกษาตอนต้น โดยการเทียบคะแนนตามแนวดิ่งตามทฤษฎีการตอบข้อสอบ แบบจำลองโลจิสติก สามพารามิเตอร์. วิทยานิพนธ์ กศ.ด. กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- .(2538ข). "การเทียบคะแนนระหว่างแบบทดสอบ (Test Equating),"วารสารการวัดผลการศึกษา.48 (16) : 49-59.
- สงบ ลักษณะ. (2525). "การเทียบระดับคะแนนระหว่างแบบทดสอบ(Equating),"วารสารการวัดผลการศึกษา 4(2) : 21-32.
- สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ. (2532). ความก้าวหน้าของคุณภาพนักเรียนชั้นประถมศึกษาปีที่ 6 ระดับประเทศ.สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ.กระทรวงศึกษาธิการ.

- สุจินดา ผ่องอักษร. (2533). การศึกษาความก้าวหน้าของผลสัมฤทธิ์ทางการเรียนในกลุ่มทักษะของนักเรียน
ชั้นประถมศึกษาปีที่ 6 ที่เรียนจบตามหลักสูตรประถมศึกษา พุทธศักราช 2521 ในช่วงระยะเวลา 3
ปีการศึกษา (ปีการศึกษา 2529 ถึง 2531) โดยใช้การเทียบคะแนนแบบบราลส์. ปรินูญนิพนธ์
กศ.ด. กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สุรัตน์ ขวัญบุญจันทร์. (2531). การสร้างแบบทดสอบและตารางเทียบคะแนนตามแนวนอนใน
วิชาวิทยาศาสตร์ที่วิเคราะห์ด้วยบราลส์โมเดล. ปรินูญนิพนธ์ กศ.ม. กรุงเทพฯ : บัณฑิตวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สำราญ มีแจ้ง. (2534). การประเมินผลและการสร้างแบบทดสอบ. เลย : ภาควิชาการทดสอบและวิจัยการ
ศึกษา วิทยาลัยครูเลย.
- อาภรณ์ กาญจนกิจโสภณ. (2531). การสร้างแบบทดสอบและตารางเทียบคะแนนตามแนวตั้ง
ในวิชาคณิตศาสตร์. ปรินูญนิพนธ์ กศ.ม. กรุงเทพฯ : บัณฑิตวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- Angoff, W.H. (1971). "Scale, Norm and Equivalent Score," In R.L. Thorndike (Ed.). *Educational
Measurement*. p.508-600. 2nd ed. Washington: D.C. American Council on Educational.
- Flanagan, T.C. (1951). "Unit, Score and Norm," In E.F. Linn (Ed). *Educational Measurement*.
p.665 – 673 .2nd ed. Washington: D.C. American Council on Educational .
- Gulliksen, Harold. (1950). *Theory of Mental Tests*. New York : John Wiley and Sons, Inc.
- Han, T., Kolen, M. & Pohlmann, J. (1997). "A comparison Among IRT True-and Observed-score
Equating and Traditional Equipercentile Equating," *Applied Measurement in Education*. 10
(2):105 – 121.
- Hambleton, R.K. and Swaminathan, H (1985). *Item Response Theory*. P.198 - 223 .Boston :
Kluwer, Nijhoff .
- Harris, R.K. and Hoover. (1987, June). "An Application of The Three – Parameter IRT Model to
Vertical Equating ," *Journal of Applied Psychological Measurement*. 11(2) : 151 – 159.
- Kaskowitz, Gary Scott. (1998). *The effect of error in item parameter estimates on linking and equating
with the IRT test characteristic curve method*, Dissertation abstracts. Ph.D. University of
Maryland College Park.
- Kolen, M.J. (1981, Spring). "Comparison of Traditional and Item Response Theory Methods for
Equating Test," *Journal of Educational Measurement*. 18(1) : 1-11.
- Kolen, M.J. and Whiney. (1982). "Comparison four Procedure for Equating The Test of General
Educational Development." *Journal of Educational Measurement*. 19:279 – 294.
- Lord, F.M. (1977a). "The Standard Error of Equipercentile Equating," *Journal of Educational Statistics*. 14
(2):117 – 138.
- (1977b). "Practical Applications of Item Characteristic Curve Theory," *Journal of Education
Measurement*. 14(2):117 - 138.
- (1980). *Application of Item Response Theory to Practical Testing Problem*. Hillsdale, N.J. :
Erlbaum.

- Lord, F.M. (1982). *"Item Response Theory and Equating-A Technical Summary"*. In P.W. Holland and D.B. Rubin (Eds). *Test Equating*. New York : Academic press.
- Lord, F.M. and Novick, M.R. (1968). *Statistical Theory of Mental Test Score*. MA: Annison – Wesley.
- Lord, F.M. and Wingersky, M.S. (1984). "Comparison of IRT True Score and Equipercetile Observed-Score Equating," *Journal of Applied Psychological Measurement*. 4(8) 453 – 461.
- Marco, Gary L. (1977). *"Item Characteristic Curve Solution to Three Intractable Testing Problem"*, *Journal of Educational Measurement*. 14:139 – 160.
- Petersen, N.S., Cook, L. and Stocking, L. (1981). *IRT Versus Conventional Equating Method : A Comparative Study of Scale Stability*. Los Angeles.
- Petersen, N.S. and others. (1982). "A test of A dequacy of linear Score Equating Method". In P.W. Holland and D.B. Rubin (Eds). *Test Equating*. New York : Academic press.
- Petersen, N.S., Kolen, M.J. and Hoover, H.D. (1989). "Scaling, Norming, and Equating." In R.L. Linn (Ed.). *Educational Measurement*. New York: Macmillan : 3rd 221 – 262.
- Rubin, Donald, B and Holland, D.W. (1982). *Test Equating*. Princeton, New Jersey : Education Testing Serving.
- Slinde, J.A. and Linn, R.L. (1978. Spring). "Vertical Equating test Factor Phantom," *Journal of Educational Measurement*. 14:23 – 35.
- (1979). "The Rasch Model. Objective Measurement Equating and Robustness," *Journal of Applied Psychological Measurement*. 3:437 – 452.
- Suanthong, Surinthorn. (1998). *An investigation of factors affecting test equating in latent trait theory (Item difficulty . Deviation type)*. Dissertation abstract. Ph.D .University of Texas.
- Wright, B.D. and Stone, M.H. (1979). *Best Test Design*. Chicago : Mesa Press.
- Warm, Thomas A. (1978). *A Primer of Item Response Theory*.
- Yang, Wen – Ling. (1997) . *The effect of content homogeneity and equating method on the accuracy of commom-item test equating*. Dissertation abstract . Ph.D . Michigan State University.

ภาคผนวก

ภาคผนวก ก

- ตารางวิเคราะห์รายละเอียดเนื้อหาวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 เรื่องสมการ
- แบบทดสอบ ฉบับ X
- แบบทดสอบ ฉบับ Y

ตาราง 11 ตารางวิเคราะห์รายละเอียด เนื้อหาเรื่อง สมการและการแก้สมการ

ลำดับที่	เนื้อหา	พฤติกรรม	ทักษะการ คำนวณ	ความ เข้าใจ	การนำ ไปใช้	การ วิเคราะห์	รวม
1	สมการ		-	5	-	-	5
2	สมการที่เป็นจริง		10	-	-	-	10
3	สมการที่มีตัวไม่ทราบค่า		-	5	-	-	5
4	คำตอบของสมการ		15	-	-	-	15
5	สมบัติการเท่ากันเกี่ยวกับการบวก และลบ		-	7	-	-	7
6	การแก้สมการโดยใช้การเท่ากันเกี่ยว กับการบวกและลบ		9	-	-	-	9
7	สมบัติการเท่ากันเกี่ยวกับการคูณและ หาร		-	9	-	-	9
8	การแก้สมการโดยใช้สมบัติการเท่า กันเกี่ยวกับการคูณและหาร		8	-	-	-	8
9	การแก้โจทย์ปัญหาด้วยสมการ		-	6	15	5	26
10	การเขียนสมการจากข้อความที่ กำหนดให้		-	6	-	-	6
รวม			42	38	15	5	100

แบบทดสอบวิชาคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6 เรื่อง สมการ
(ฉบับ X)

คำชี้แจง

1. แบบทดสอบฉบับนี้มีคำถามทั้งหมด 30 ข้อ และให้เวลาในการทำ 30 นาที
2. คำถามเป็นแบบปรนัย ชนิดเลือกตอบ ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุด เพียงคำตอบเดียว ก,ข,ค,ง,และ จ ที่ให้ไว้ เมื่อนักเรียนเลือกคำตอบใด ให้ทำเครื่องหมายทับตรงตัวเลือกที่ต้องการ ดังตัวอย่างต่อไปนี้

1.) ก ข ค ง จ

☐ ☒ ☐ ☐ ☐

ถ้านักเรียนขีดคำตอบไปแล้ว ต้องการเปลี่ยนคำตอบใหม่ ให้ทำเครื่องหมายขีดทับข้อที่ไม่ต้องการให้เห็นชัดเจนทุกครั้ง แล้วค่อยขีดคำตอบใหม่ ดังนี้

1.) ก ข ค ง จ

☐ ☒ ☐ ☐ ☒

3. แบบทดสอบนี้มีทั้งหมด 4 หน้า
4. อย่าขีดเขียนหรือทำเครื่องหมายใดๆ ในแบบทดสอบนี้ เพราะข้อสอบนี้จะต้องนำไปใช้อีก

1. ข้อใดคือสมการที่เป็นจริง

ก. $30 \times 20 = 3 \times 100$

ข. $8 \times 20 = 2 \times 80$

ค. $(7 \times 10) + (7 \times 16) = 7 \times 25$

ง. $2 + 20 + 4,000 = 4,202$

ช. $\frac{1}{4} + \frac{3}{4} = \frac{7}{15} + \frac{8}{15}$

2. ข้อใดคือสมการที่เป็นจริง

ก. $17 + 4 = (5 \times 5) - 4$

ข. $100 + 50 + 9 = 119 + 30$

ค. $75 - 4 = (35 - 5) \times 1$

ง. $(25 \times 2) + 3 = 7 \times 5$

ช. $(4 \times 9) + 9 = (8 \times 2) \times 3$

3. ข้อใดคือสมการที่เป็นจริง

ก. $(55 \div 5) - 1 = (2 \times 6) - 2$

ข. $(0.5 \times 0.8) = 0.4$

ค. $35 \times 3 = (30 \times 4) - 25$

ง. $(4 \times 40) - 60 = (20 \times 5) - 10$

จ. $16 \times 3 = (3 \times 4) \times 5$

4. ข้อใดคือสมการที่เป็นจริง

ก. $400 + 40 + 5 = 1,000 - 545$

ข. $(3 \times 17) + (17 \times 2) = (3 \times 2) \times 17$

ค. $(27 \div 3) + 2 = (13 \times 2) - 3$

ง. $(3 \times 8) + (8 \times 2) = 8 \times 5$

จ. $(4 + 5) \times 2 = (5 \times 2) + 4$

5. ข้อใดคือสมการที่เป็นจริง

ก. $3 + (8 \times 2) = (5 \times 3) + 5$

ข. $(35 - 10) \div 5 = 100 \div (5 \times 4)$

ค. $25 - 12 = 23 - 11$

ง. $84 \div 2 = 40 + 3$

จ. $(2 + 7) \times 3 = (3 \times 3) \times 4$

6. ข้อใดเป็นคำตอบของสมการ $m \div 22 = 8$

ก. 167

ข. 176

ค. 30

ง. 14

จ. 96

7. ข้อใดเป็นคำตอบของสมการ $405 = j \times 45$

ก. 3

ข. 6

ค. 7

ง. 8

จ. 9

8. ข้อใดเป็นคำตอบของสมการ $56 - x = 27$

ก. 38

ข. 29

ค. 25

ง. 27

จ. 83

9. ข้อใดเป็นสมบัติการเท่ากับเกี่ยวกับการบวก

ก. ถ้า $k - 1 = 3$ แล้ว $k - 1 + 1 = 3 + 1$

ข. ถ้า $x + 1 = 2$ แล้ว $x + 1 - 1 = 2 - 1$

ค. ถ้า $d + 2 = 3$ แล้ว $d + 2 - 2 = 3 - 2$

ง. ถ้า $g - 2 = 4$ แล้ว $g - 2 - 2 = 4 - 2$

จ. ถ้า $j + 5 = 7$ แล้ว $j + 5 - 5 = 7 - 5$

10. ข้อใดคือคำตอบของสมการ $\frac{5}{6} + b = \frac{4}{3}$

ก. $\frac{2}{3}$

ข. $\frac{1}{2}$

ค. $\frac{4}{6}$

ง. $\frac{5}{6}$

จ. $\frac{1}{3}$

11. ข้อใดคือคำตอบของสมการ $x - \frac{1}{7} = \frac{6}{7}$

ก. $\frac{2}{7}$

ข. $\frac{3}{7}$

ค. $\frac{4}{7}$

ง. $\frac{5}{7}$

จ. 1

12. ข้อใดคือคำตอบของสมการ $X - 79 = 38$

ก. 27

ข. 39

ค. 41

ง. 117

จ. 171

13. ข้อใดเป็นสมบัติการเท่ากันเกี่ยวกับการคูณและการหาร

ก. ถ้า $5 \times 4 = (2+3) \times 4$ แล้ว $5 \times 4 \div 4 = (2+3) \times 4 \div 4$

ข. ถ้า $2 \times 6 = (3+3) \times 2$ แล้ว $2 \times 6 \div = (3+3) \times 2 \div 2$

ค. ถ้า $1 \times 6 = 2 \times 3$ แล้ว $1 \times 6 \div 6 = (2 \times 3) \times 6$

ง. ถ้า $2 \times 8 = 4 \times 4$ แล้ว $2 \times 8 \div = (4 \times 4) \div 2$

จ. ถ้า $4 \times 3 = 2 \times 6$ แล้ว $4 \times 3 \div 3 = 2 \times 6 \div 2$

14. ข้อใดเป็นสมบัติการเท่ากันเกี่ยวกับการคูณ

ก. ถ้า $2 \div 2 = 1$ แล้ว $(2 \div 2) \times 2 = 1 \times 2$

ข. ถ้า $9 \div 3 = 3$ แล้ว $9 \div 3 \times 3 = 3 \times \frac{1}{3}$

ค. ถ้า $14 = 7 \times 2$ แล้ว $14 \div 2 = (7 \times 2) \div 2$

ง. ถ้า $2 \times 10 = 20$ แล้ว $2 \times 10 \div 10 = 20 \div 10$

จ. ถ้า $4 \times 9 = 36$ แล้ว $4 \times 9 \div 9 = 36 \div 9$

15. จงแก้สมการ $x \div 5 = 75$

ก. 305

ข. 315

ค. 325

ง. 350

จ. 375

16. จงแก้สมการ $b \times 12 = 288$

ก. 6

ข. 9

ค. 12

ง. 16

จ. 24

17. ต่ายมีเงิน ก บาท ซื้อสมุดไป 35 บาท เหลือ 78 บาท ข้อใดถูกต้อง

ก. $k - 35 = 78$, $k = 113$

ข. $k + 35 = 78$, $k = 43$

ค. $78 - k = 35$, $k = 43$

ง. $78 + k = 35$, $k = 113$

จ. $k - 78 = 35$, $k = 113$

18. เลี้ยงเป็ดไว้ 150 ตัว เป็นเพศผู้ x ตัว เป็นเพศเมีย 39 ตัว มีเป็ดตัวผู้กี่ตัว ข้อใดถูกต้อง

ก. $150 - 39 = x$, $x = 111$

ข. $x \times 150 = 39$, $x = 111$

ค. $x - 39 = 150$, $x = 111$

ง. $150 + 39 = 150$, $x = 111$

จ. $39 - x = 150$, $x = 111$

19. เต๋ยมียเงิน 680 บาท ตูมียเงิน ๗ บาท เมื่อรวมกันแล้วมีเงิน 850 บาท อยากทราบว่าตูมียเงินเท่าใด

ก. 140 บาท

ข. 150 บาท

ค. 160 บาท

ง. 170 บาท

จ. 180 บาท

20. ถ้า 14 เท่าของเงินจำนวนหนึ่ง คิดเป็น 1,120 บาท จงหาจำนวนนั้น
- 30
 - 40
 - 60
 - 80
 - 90
21. ก้อยอ่านหนังสือ ค วันๆละ 9 หน้า ก้อยอ่านได้ทั้งสิ้น 153 หน้า อยากทราบว่า ก้อยอ่านหนังสือกี่วัน
- 13
 - 15
 - 17
 - 19
 - 21
22. จอยขายขนมไป จ กล่อง ราคากล่องละ 18 บาท ได้เงินทั้งสิ้น 162 บาท จอยขายขนมไปทั้งหมดกี่กล่อง
- 9
 - 18
 - 36
 - 72
 - 144
23. ปู่มีเงินจำนวนหนึ่งแบ่งให้หลาน 7 คนๆละเท่ากันได้คนละ 95 บาท จงหาว่าปู่มีเงินเท่าใด
- 625
 - 635
 - 645
 - 655
 - 665
24. มีข้าวสาร 150 กระสอบ ขายไป y กระสอบ เหลือข้าวสาร 41 กระสอบ เขียนเป็นสมการได้อย่างไร
- $150-41=y$
 - $150+41=y$
 - $y+150=41$
 - $150-y=41$
 - $41-y=150$
25. จำนวนหนึ่งหารด้วย 9 ได้ 8 เขียนเป็นสมการได้อย่างไร
- $z \div 9=8$
 - $z \div 8=9$
 - $x-9=8$
 - $x+9=8$
 - $y-8=9$
26. รถคันหนึ่งแล่นได้ระยะทาง 10 กม. ใช้น้ำมันไป k ลิตร ถ้ามีระยะทางทั้งหมด 140 กม. จะใช้น้ำมันกี่ลิตร และถ้าหากมีเงินอยู่ 500 บาท จะพอซื้อน้ำมันหรือไม่ ข้อนี้หาคำตอบไม่ได้เพราะเหตุใด
- จำนวนน้ำมันที่ใช้ไป
 - ราคาน้ำมัน
 - ระยะทางที่แล่นไป
 - จำนวนเงินทั้งหมด
 - จำนวนระยะทางทั้งหมด
27. แม่มีอายุเป็น 4 เท่าของลูก ปัจจุบันแม่มีอายุ 52 ปี จงหาว่าลูกมีอายุเท่าใด
- 10
 - 11
 - 12
 - 13
 - 14

28. ออฟเลี้ยงเปิด ก ตัว นับขารวมกันได้ 90

ขา ออฟเลี้ยงเปิดกี่ตัว

ก. 5

ข. 15

ค. 25

ง. 35

จ. 45

29. มีเก้าอี้ ข ตัว จัดเป็น 17 แถว แถวละ 5

ตัวพอดี ดังนั้นจะมีเก้าอี้ทั้งหมดกี่ตัว

ก. 55

ข. 68

ค. 75

ง. 85

จ. 95

30. มีขนม ง ชิ้น แบ่งใส่ถุง 4 ถุงๆละ 15 ชิ้น

เขียนเป็นสมการได้อย่างไร

ก. $4 \div g = 15$

ข. $g \div 4 = 15$

ค. $15 \div g = 4$

ง. $15 \div 4 = g$

จ. $15 \times g = 4$

แบบทดสอบวิชาคณิตศาสตร์
ชั้นประถมศึกษาปีที่ 6 เรื่อง สมการ
(ฉบับ Y)

คำชี้แจง

1. แบบทดสอบฉบับนี้มีคำถามทั้งหมด 30 ข้อ และให้เวลาในการทำ 30 นาที
2. คำถามเป็นแบบปรนัย ชนิดเลือกตอบ ให้นักเรียนเลือกคำตอบที่ถูกต้องที่สุด เพียงคำตอบเดียว ก,ข,ค,ง,และ จ ที่ให้ไว้ เมื่อนักเรียนเลือกคำตอบใด ให้ทำเครื่องหมายทับตรงตัวเลือกที่ต้องการ ดังตัวอย่างต่อไปนี้

1.) ก ข ค ง จ

☐ ☒ ☐ ☐ ☐

ถ้านักเรียนขีดคำตอบไปแล้ว ต้องการเปลี่ยนคำตอบใหม่ ให้ทำเครื่องหมายขีดทับข้อที่ไม่ต้องการให้เห็นชัดเจนทุกครั้ง แล้วค่อยขีดคำตอบใหม่ ดังนี้

1.) ก ข ค ง จ

☐ ☒ ☐ ☐ ☒

3. แบบทดสอบนี้มีทั้งหมด 4 หน้า
4. อย่าขีดเขียนหรือทำเครื่องหมายใด ๆ ในแบบทดสอบนี้ เพราะข้อสอบนี้จะต้องนำไปใช้อีก

1. ข้อใดคือสมการที่เป็นจริง

- ก. $(6 + 9) \times 2 = 12 \times 9$
- ข. $1,269 = (500 \times 2) + (40 \times 5) + (3 \times 24)$
- ค. $8 \times 7 = (5 \times 10) + 7$
- ง. $(2 \times 15) \div 3 = (50 \div 10) + 3$
- จ. $(3 \times 9) + 18 = (18 - 8) \times 5$

2. ข้อใดคือสมการที่เป็นจริง

- ก. $(16 \times 2) - 1 = 3 \times (5 \times 2)$
- ข. $28 - 9 = 19$
- ค. $19 \div 1 = 20$
- ง. $110 \div 11 = 11 - 1$
- จ. $61 - 9 = (13 \times 4) - 3$

3. ข้อใดคือสมการที่เป็นจริง

- ก. $625 \div 5 = 15 \times 5$
- ข. $73 - 16 = (25 \times 2) + 8$
- ค. $14 \times 3 = 50 - (4 \times 3)$
- ง. $45 - 9 = (7 - 2) \times 7$
- จ. $(23 \times 17) + (3 \times 23) = (17 + 3) \times 23$

4. ข้อใดคือสมการที่เป็นจริง

- ก. $64 \div 16 = 8 \div 4$
- ข. $(2 \times 9) \times 3 = (3 \times 6) \times 4$
- ค. $(150 \div 3) - (100 \div 4) = 75 \div 3$
- ง. $1,556 = 1,000 + 500 + 60 + 5$
- จ. $(76 - 6) \div 10 = 10 - 4$

5. ข้อใดคือสมการที่เป็นจริง

- ก. $100 \div 5 = 200$
- ข. $15 \times (12 - 10) = (12 \times 10) - 80$
- ค. $5 \times 60 = (15 \times 10) + (80 \times 2)$
- ง. $2,504 = 2,000 + 50 + 4$
- จ. $(3 \times 4) + 5 = (4 \times 5) - 3$

6. ข้อใดเป็นคำตอบของสมการ $891 \div x = 99$

- ก. 3
- ข. 792
- ค. 6
- ง. 18
- จ. 9

7. ข้อใดเป็นคำตอบของสมการ $\frac{1}{8} + A = \frac{1}{4}$

- ก. 1
- ข. $\frac{1}{8}$
- ค. $\frac{1}{4}$
- ง. $\frac{1}{16}$
- จ. $\frac{1}{32}$

8. ข้อใดเป็นคำตอบของสมการ $0.45 + X = 1041$

- ก. 0.63
- ข. 0.54
- ค. 0.64
- ง. 1086
- จ. 0.96

9. ข้อใดเป็นสมบัติการเท่ากันของการบวก

- ก. ถ้า $m - 2 = 4$ แล้ว $m - 2 - 2 = 4 + 2$
- ข. ถ้า $n + 3 = 7$ แล้ว $n + 3 - 3 = 7 - 3$
- ค. ถ้า $r - 4 = 6$ แล้ว $r - 4 + 4 = 6 + 4$
- ง. ถ้า $p + 5 = 8$ แล้ว $p + 5 - 5 = 8 - 5$
- จ. ถ้า $k + 1 = 2$ แล้ว $k + 1 - 1 = 2 + 1$

10. ข้อใดคือคำตอบของสมการ $a + 0.62 = 1.58$

- ก. 0.39
- ข. 0.69
- ค. 2.55
- ง. 2.20
- จ. 0.96

11. ข้อใดคือคำตอบของสมการ $3.96 + m = 9.36$

- ก. 3.36
- ข. 6.36
- ค. 5.40
- ง. 12.62
- จ. 13.32

12. ข้อใดคือคำตอบของสมการ $x + 0.48 = 4.8$

- ก. 0
- ข. 1.56
- ค. 2.30
- ง. 4.32
- จ. 5.28

13. ข้อใดเป็นสมบัติการเท่ากันเกี่ยวกับการหาร

- ก. ถ้า $4.5 \times 3.2 = 14.4$ แล้ว $4.5 \times 3.2 \div 3.2 = 14.4 \div 4.5$
- ข. ถ้า $5 \times 6 = 30$ แล้ว $5 \times 6 \div 6 = 30 \div 5$
- ค. ถ้า $3 \times 9 = 27$ แล้ว $3 \times 9 \div 9 = 27 \div 3$
- ง. ถ้า $9 \times 2 = 18$ แล้ว $9 \times 2 \div 2 = 18 \div 9$
- จ. ถ้า $1 \times 5 = 5$ แล้ว $1 \times 5 \div 5 = 5 \div 5$

14. ข้อใดเป็นสมบัติการเท่ากันเกี่ยวกับการหาร

- ก. ถ้า $8 \times 4 = 32$ แล้ว $8 \times 4 \div 8 = 32 \div 4$
- ข. ถ้า $15 = 3 \times 5$ แล้ว $15 \div 5 = 3 \times 5 \div 5$
- ค. ถ้า $2 \times 6 = 12$ แล้ว $2 \times 6 \div 2 = 12 \times 2$
- ง. ถ้า $4 \times 9 = 36$ แล้ว $4 \times 9 \div 9 = 36 \div 4$
- จ. ถ้า $3 \times 6 = 18$ แล้ว $3 \times 6 \div 6 = 18 \div 3$

15. จงแก้สมการ $y \div 98 = 294$

- ก. 2
- ข. 3
- ค. 4
- ง. 6
- จ. 8

16. จงแก้สมการ $s \div 85 = 340$

- ก. 2
- ข. 3
- ค. 4
- ง. 5
- จ. 6

17. มีทุเรียน ข ผล ฝรั่ง 39 ผล รวมมีผลไม้ 45 ผล ข้อใดถูกต้อง

- ก. $x - 39 = 45$, $x = 6$
- ข. $39 - x = 45$, $x = 45$
- ค. $x + 39 = 45$, $x = 6$
- ง. $39 + 45 = x$, $x = 84$
- จ. $45 - 39 = x$, $x = 39$

18. รูปสี่เหลี่ยมจัตุรัสมีความยาวด้านละ k ซม วัดความยาวรอบรูปได้ 68 ซม ข้อใดถูกต้อง

- ก. $k \times 68 = 4$, $k = 17$
- ข. $4 \times k = 68$, $k = 17$
- ค. $68 - k = 4$, $k = 17$
- ง. $68 - 4 = k$, $k = 17$
- จ. $4 \times 68 = k$, $k = 17$

19. แม่มีอายุเป็น 4 เท่าของลูก ปัจจุบันแม่มีอายุ 52 ปี จงหาว่าลูกมีอายุเท่าใด

- ก. 10
- ข. 11
- ค. 12
- ง. 13
- จ. 14

20. ออฟเลี้ยงเป็ด k ตัว นับขารวมกันได้ 90

ขา ออฟเลี้ยงเป็ดกี่ตัว

- ก. 5
- ข. 15
- ค. 25
- ง. 35
- จ. 45

21. เดินออกกำลังกายได้ระยะทางวันละ 7 กม. เดินอยู่ k วัน ได้ระยะทาง 126 กม. อยากทราบว่าเดินกี่วัน
- ก. 16
ข. 17
ค. 18
ง. 19
จ. 20
22. ก้อยอ่านหนังสือ ค วันๆละ 9 หน้า ก้อยอ่านได้ทั้งสิ้น 153 หน้า อยากทราบว่าก้อยอ่านหนังสือกี่วัน
- ก. 13
ข. 15
ค. 17
ง. 19
จ. 21
23. มีแก้วอิ๋ ซ ตัว จัดเป็น 17 แถว แถวละ 5 ตัวพอดี ดังนั้นจะมีแก้วอิ๋ทั้งหมดกี่ตัว
- ก. 55
ข. 68
ค. 75
ง. 85
จ. 95
24. มีเงินอยู่ x บาท ซื้อสมุด 61 บาท เหลือเงิน 17 บาท เขียนเป็นสมการได้อย่างไร
- ก. $x + 17 = 61$
ข. $61 - x = 17$
ค. $x + 61 = 17$
ง. $x - 61 = 17$
จ. $17 \times 61 = x$
25. มีขนม ง ชิ้น แบ่งใส่ถุง 4 ถุงๆละ 15 ชิ้น เขียนเป็นสมการได้อย่างไร
- ก. $4 \div 3 = 15$
ข. $3 \div 4 = 15$
ค. $15 \div 3 = 4$
ง. $15 \div 4 = 3$
จ. $15 \times 3 = 4$
26. รถคันหนึ่งแล่นได้ระยะทาง 10 กม. ใช้น้ำมันไป k ลิตร ถ้ามีระยะทางทั้งหมด 140 กม. จะใช้น้ำมันกี่ลิตร และถ้าหากมีเงินอยู่ 500 บาท จะพอซื้อน้ำมันหรือไม่ ข้อนี้หาคำตอบไม่ได้เพราะเหตุใด
- ก. จำนวนน้ำมันที่ใช้ไป
ข. ราคาน้ำมัน
จ. ระยะทางที่แล่นไป
ฉ. จำนวนเงินทั้งหมด
ช. จำนวนระยะทางทั้งหมด
27. จอยขายขนมไป จ กล่อง ราคากล่องละ 18 บาท ได้เงินทั้งสิ้น 162 บาท จอยขายขนมไปทั้งหมดกี่กล่อง
- ก. 9
ข. 18
ค. 36
ง. 72
จ. 144
28. ปู่มีเงินจำนวนหนึ่งแบ่งให้หลาน 7 คนๆละเท่ากันได้คนละ 95 บาท จงหาว่าปู่มีเงินเท่าใด
- ก. 625
ข. 635
ค. 645
ง. 655
จ. 665
29. ถ้า 14 เท่าของเงินจำนวนหนึ่ง คิดเป็น 1,120 บาท จงหาจำนวนนั้น
- ก. 30
ข. 40
ค. 60
ง. 80
จ. 90

30. เต๋ยมีเงิน 680 บาท ตูมีเงิน ๗ บาท เมื่อรวมกัน
แล้วมีเงิน 850 บาท อยากทราบว่ตูมีเงินเท่าใด

ก. 140 บาท

ข. 150 บาท

ค. 160 บาท

ง. 170 บาท

จ. 180 บาท

ภาคผนวก ข

ค่าไอเกินขององค์ประกอบที่มีค่าเกิน 1.0

ตาราง 12 ค่าไอเกนขององค์ประกอบที่มีค่าเกิน 1.0 ของแบบทดสอบวิชาคณิตศาสตร์
 ชั้นประถมศึกษาปีที่ 6 เรื่องสมการ

องค์ประกอบ	ค่าไอเกน	% of Varian	องค์ประกอบ	ค่าไอเกน	% of Varian
1	16.600	16.600	16	1.378	1.378
2	3.720	3.720	17	1.338	1.338
3	2.997	2.997	18	1.311	1.311
4	2.625	2.625	19	1.252	1.252
5	2.263	2.263	20	1.228	1.228
6	1.943	1.943	21	1.208	1.208
7	1.855	1.855	22	1.189	1.189
8	1.744	1.744	23	1.174	1.174
9	1.633	1.633	24	1.124	1.124
10	1.594	1.594	25	1.112	1.112
11	1.536	1.536	26	1.070	1.070
12	1.501	1.501	27	1.051	1.051
13	1.468	1.468	28	1.041	1.041
14	1.436	1.436	29	1.027	1.027
15	1.413	1.413	30	1.013	1.013

ภาคผนวก ค

ค่าพารามิเตอร์ของแบบทดสอบจากการทดลองเครื่องมือ

ตารางที่ 13 ผลการวิเคราะห์ค่าพารามิเตอร์จากการทดลองเครื่องมือ

เนื้อหา	พฤติกรรม	ข้อที่	a	b	c	CHISQUARE (PROP)	หมายเหตุ
1	ความเข้าใจ	1	0.597	-1.849	0.186	0.7845	ใช้ได้
		2	0.711	-1.235	0.168	0.1871	ใช้ได้
		3	0.632	-0.335	0.169	0.5818	ใช้ได้
		4	0.807	-0.874	0.174	0.7002	ใช้ได้
		5	0.784	-0.323	0.180	0.7437	ใช้ได้
2	ทักษะการ คำนวณ	6	0.933	0.262	0.166	0.5058	ใช้ได้
		7	1.148	0.269	0.295	0.2072	ใช้ได้
		8	1.195	0.443	0.196	0.8410	ใช้ได้
		9	1.286	0.268	0.206	0.3399	ใช้ได้
		10	1.232	0.714	0.219	0.4671	ใช้ได้
		11	1.411	0.565	0.234	0.4817	ใช้ได้
		12	1.264	0.720	0.262	0.2736	ใช้ได้
		13	1.274	0.220	0.155	0.7398	ใช้ได้
		14	1.053	0.306	0.183	0.5563	ใช้ได้
		15	1.233	0.371	0.160	0.5205	ใช้ได้
3	ความเข้าใจ	16	1.209	-0.902	0.156	0.3310	ใช้ได้
		17	0.043	-0.401	0.114	0.4013	ใช้ได้
		18	0.102	-0.248	0.134	0.6232	ใช้ได้
		19	0.979	-0.871	0.113	0.0556	ใช้ได้
		20	0.768	-0.935	0.124	0.0090	ใช้ได้
4	ทักษะการ คำนวณ	21	1.092	0.329	0.117	0.8651	ใช้ได้
		22	0.579	0.339	0.182	0.5835	ใช้ได้
		23	1.220	-0.259	0.074	0.0911	ใช้ได้
		24	1.076	-0.149	0.173	0.1258	ใช้ได้
		25	0.397	0.506	0.137	0.0142	ใช้ได้
		26	1.352	-0.004	0.144	0.4626	ใช้ได้
		27	0.846	0.379	0.105	0.1831	ใช้ได้
		28	0.912	0.757	0.134	0.2774	ใช้ได้
		29	1.374	0.299	0.284	0.0882	ใช้ได้
		30	1.093	0.357	0.235	0.4092	ใช้ได้
		31	1.089	1.906	0.199	0.4151	ใช้ได้
		32	1.243	1.367	0.376	0.6616	ใช้ได้
		33	0.749	-0.221	0.140	0.0442	ใช้ได้
		34	0.846	1.237	0.133	0.5298	ใช้ได้
		35	0.766	0.492	0.111	0.1991	ใช้ได้

ตาราง 13 (ต่อ)

เนื้อหา	พฤติกรรม	ข้อที่	a	b	c	CHISQUARE (PROP)	หมายเหตุ
5	ความเข้าใจ	36	0.495	0.510	0.126	0.0377	ใช้ได้
		37	0.572	1.764	0.209	0.0485	ใช้ได้
		38	0.814	0.200	0.152	0.6891	ใช้ได้
		39	0.152	0.797	0.196	0.0884	ใช้ได้
		40	0.712	1.146	0.193	0.7649	ใช้ได้
		41	0.575	1.167	0.154	0.2849	ใช้ได้
		42	1.092	0.769	0.205	0.9705	ใช้ได้
6	ทักษะการ คำนวณ	43	0.906	0.222	0.232	0.6569	ใช้ได้
		44	0.592	0.260	0.152	0.4122	ใช้ได้
		45	1.210	0.378	0.209	0.2285	ใช้ได้
		46	0.890	0.295	0.184	0.2100	ใช้ได้
		47	1.004	0.701	0.106	0.2631	ใช้ได้
		48	1.178	0.306	0.176	0.6991	ใช้ได้
		49	0.858	0.633	0.126	0.3419	ใช้ได้
		50	0.934	1.082	0.134	0.3145	ใช้ได้
		51	0.791	0.686	0.236	0.6152	ใช้ได้
7	ความเข้าใจ	52	0.849	0.778	0.239	0.9513	ใช้ได้
		53	0.815	1.099	0.142	0.9219	ใช้ได้
		54	1.220	2.327	0.173	0.3254	
		55	0.759	1.922	0.221	0.8120	ใช้ได้
		56	0.999	0.645	0.188	0.5041	ใช้ได้
		57	1.401	2.044	0.215	0.6811	
		58	1.421	0.351	0.227	0.9903	ใช้ได้
		59	0.734	1.845	0.104	0.5434	ใช้ได้
		60	0.873	0.761	0.113	0.4936	ใช้ได้
8	ทักษะการ คำนวณ	61	1.378	0.392	0.141	0.2390	ใช้ได้
		62	0.782	0.789	0.129	0.1619	ใช้ได้
		63	0.916	-0.392	0.165	0.5398	ใช้ได้
		64	1.744	0.484	0.359	0.3219	ใช้ได้
		65	1.151	-0.145	0.155	0.6680	ใช้ได้
		66	1.486	0.345	0.215	0.4740	ใช้ได้
		67	1.427	0.239	0.215	0.7417	ใช้ได้
		68	1.700	0.436	0.195	0.6281	ใช้ได้

ตาราง 13 (ต่อ)

เนื้อหา	พฤติกรรม	ข้อที่	a	b	c	CHISQUARE (PROP)	หมายเหตุ
9	ความเข้าใจ	69	1.164	0.227	0.140	0.7775	ใช้ได้
		70	0.793	0.256	0.144	0.8594	ใช้ได้
		71	1.326	1.108	0.150	0.9541	ใช้ได้
		72	1.177	0.443	0.245	0.5434	ใช้ได้
		73	1.060	0.724	0.245	0.9693	ใช้ได้
		74	1.315	0.496	0.198	0.4114	ใช้ได้
	การนำไปใช้	75	1.603	0.264	0.168	0.9090	ใช้ได้
		76	1.243	0.204	0.212	0.7244	ใช้ได้
		77	0.736	0.611	0.090	0.8387	ใช้ได้
		78	1.730	0.267	0.146	0.1722	ใช้ได้
		79	1.296	0.752	0.165	0.1385	ใช้ได้
		80	0.778	2.024	0.211	0.5221	
		81	0.670	1.493	0.197	0.9270	ใช้ได้
		82	0.614	0.282	0.147	0.4746	ใช้ได้
		83	1.332	0.342	0.189	0.3578	ใช้ได้
		84	1.457	0.421	0.250	0.1347	ใช้ได้
		85	1.768	0.763	0.153	0.8794	ใช้ได้
		86	1.361	0.334	0.203	0.6011	ใช้ได้
		87	1.187	0.381	0.239	0.5071	ใช้ได้
		88	1.339	1.343	0.272	0.4761	ใช้ได้
		89	1.175	0.369	0.186	0.4544	ใช้ได้
	การวิเคราะห์	90	1.648	0.589	0.207	0.9772	ใช้ได้
		91	0.628	0.562	0.261	0.6968	ใช้ได้
		92	0.781	0.597	0.194	0.4591	ใช้ได้
		93	0.743	0.102	0.119	0.1382	ใช้ได้
		94	0.569	0.767	0.190	0.3268	ใช้ได้
10	ความเข้าใจ	95	0.386	-	0.183	0.2194	
		96	0.447	-	0.172	1.000	
		97	0.950	-	1.053	0.0879	
		98	0.839	2.146	0.173	0.5183	
		99	0.559	0.736	0.175	0.3040	ใช้ได้
		100	0.473	2.427	0.278	0.4727	

ภาคผนวก ง

ค่าพารามิเตอร์ของแบบทดสอบทั้งสองฉบับ

ผู้วิจัยได้นำผลการตอบแบบทดสอบเทียบคะแนนมาวิเคราะห์เพื่อประมาณค่าพารามิเตอร์ของข้อสอบเป็นรายข้อ ซึ่งได้พารามิเตอร์ของข้อสอบในแบบทดสอบแต่ละฉบับ ดังตาราง 14

ตาราง 14 แสดงค่าพารามิเตอร์ของแบบทดสอบทั้งสองฉบับ

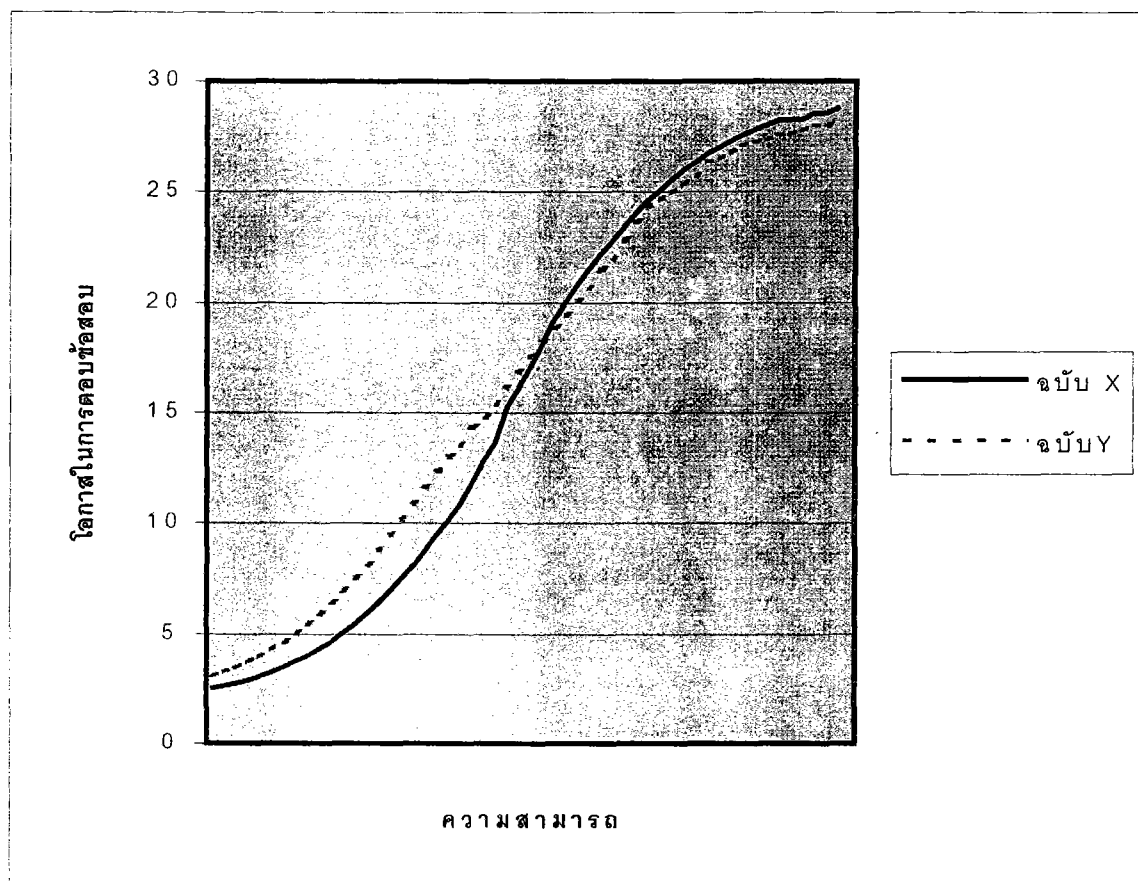
ข้อ	ฉบับ X			ฉบับ Y		
	a	b	c	a	b	c
1	0.597	-1.849	0.186	0.711	-1.235	0.168
2	0.632	-0.335	0.169	0.784	-0.323	0.180
3	0.993	0.262	0.166	1.274	0.220	0.155
4	1.148	0.269	0.295	1.286	0.268	0.206
5	1.232	0.714	0.219	1.264	0.720	0.262
6	1.053	0.306	0.183	1.233	0.371	0.160
7	1.209	-0.902	0.156	0.768	-0.935	0.124
8	1.092	0.329	0.117	0.579	0.339	0.182
9	0.749	-0.221	0.140	1.220	-0.259	0.074
10	1.243	1.367	0.376	0.846	1.237	0.133
11	0.846	0.379	0.105	1.093	0.357	0.235
12	0.712	1.146	0.193	0.575	1.167	0.154
13	1.092	0.769	0.205	0.152	0.797	0.196
14	0.814	0.200	0.152	0.906	0.222	0.232
15	0.592	0.260	0.152	0.890	0.295	0.184
16	1.004	0.701	0.106	0.791	0.686	0.236
17	1.178	0.306	0.176	1.210	0.378	0.209
18	0.759	1.922	0.221	0.734	1.845	0.104
19	0.849	0.778	0.239	0.873	0.761	0.113
20	1.378	0.392	0.141	1.486	0.345	0.215
21	1.744	0.484	0.359	1.700	0.436	0.195
22	1.164	0.227	0.140	0.793	0.256	0.144
23	1.177	0.443	0.245	1.315	0.496	0.198
24	1.296	0.752	0.165	1.768	0.763	0.153
25	1.648	0.589	0.207	0.781	0.597	0.194
26	1.187	0.381	0.239	1.332	0.342	0.189
27	1.603	0.264	0.168	1.730	0.267	0.146
28	1.457	0.421	0.250	0.736	0.611	0.090
29	0.628	0.562	0.261	0.569	0.767	0.190
30	0.569	0.767	0.190	0.559	0.736	0.175

ภาคผนวก จ

คุณภาพของแบบทดสอบทั้งสองฉบับ

ในการหาคุณภาพแบบทดสอบ ผู้วิจัยนำค่าพารามิเตอร์ประจำแต่ละข้อของแบบทดสอบทั้งสองฉบับ มาหาค่าโอกาสในการตอบถูก และค่าอินฟอร์เมชันของแบบทดสอบที่ระดับความสามารถต่างๆ ดังนี้

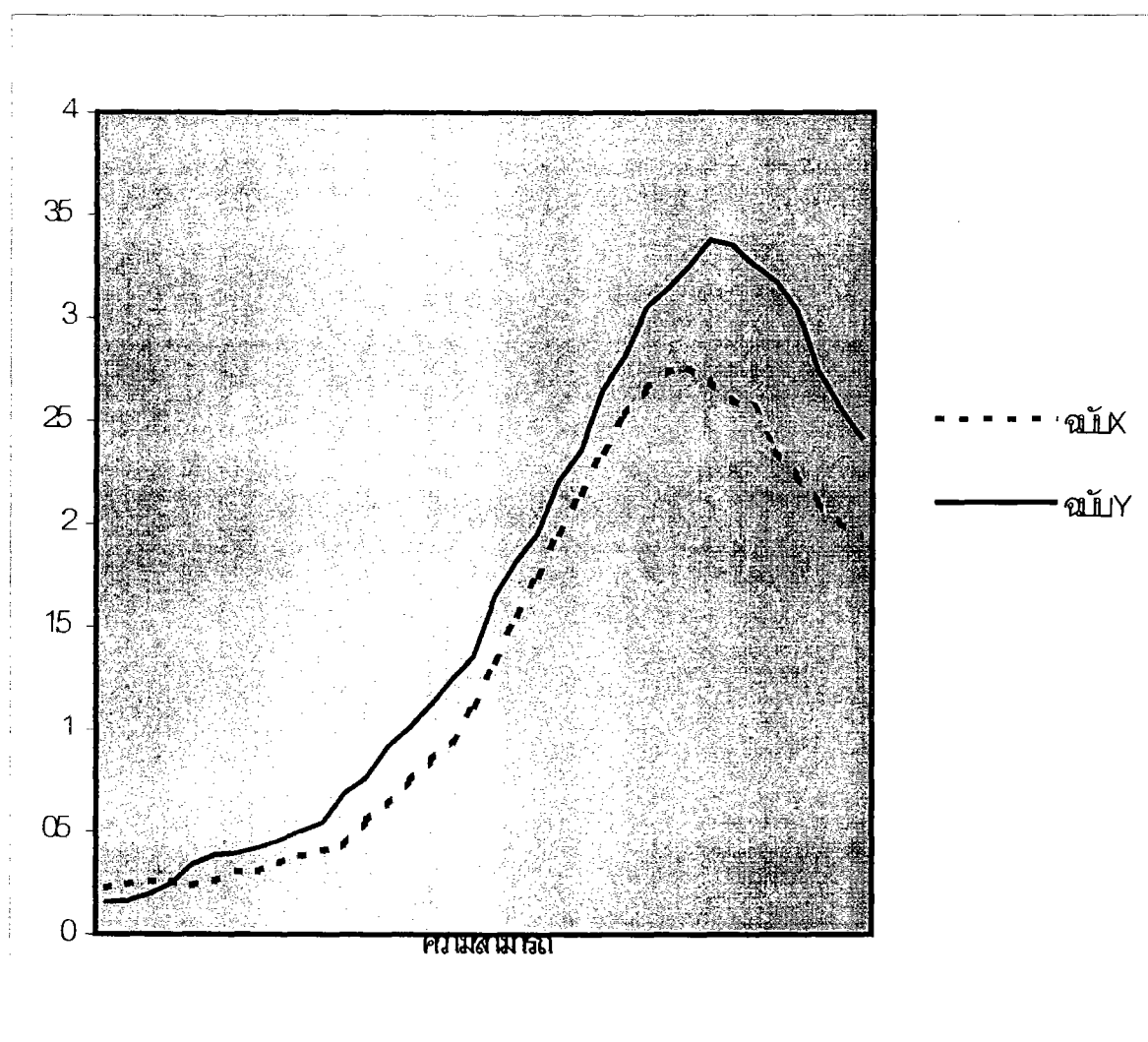
1.โอกาสในการตอบข้อสอบถูก ณ ระดับความสามารถต่างๆ แล้วนำมาเขียนโค้งลักษณะเฉพาะของแบบทดสอบ ซึ่งแสดงได้ในภาพประกอบ 7



ภาพประกอบ 7 แสดงโค้งลักษณะเฉพาะของแบบทดสอบทั้งสองฉบับ

จากภาพประกอบ 7 แสดงให้เห็นว่า ที่ระดับความสามารถ -3.0 ถึง 0.0 โอกาสในการตอบข้อสอบถูกของแบบทดสอบ ฉบับ Y สูงกว่าฉบับ X และที่ระดับความสามารถ 1.0 ถึง 3.0 แบบทดสอบทั้งสองฉบับมีโอกาสในการตอบข้อสอบถูกที่ใกล้เคียงกัน

2. ค่าสูงสุดของอินฟอร์เมชันของแบบทดสอบ ณ ระดับความสามารถต่างๆ แล้วนำมาเขียนโค้งอินฟอร์เมชันของแบบทดสอบที่ระดับความสามารถต่างๆ จะได้ดังภาพประกอบ 8



ภาพประกอบ 8 โค้งอินฟอร์เมชันของแบบทดสอบทั้งสองฉบับที่ระดับความสามารถต่างๆ

จากภาพประกอบ 8 แสดงให้เห็นว่า ที่ระดับความสามารถ -3.0 ถึง -2.0 แบบทดสอบทั้งสองฉบับให้ค่าอินฟอร์เมชันของแบบทดสอบที่ใกล้เคียงกัน แต่เมื่อพิจารณาที่ระดับความสามารถตั้งแต่ -2.0 ถึง 3.0 แล้วพบว่าฉบับ Y ให้ค่าอินฟอร์เมชันของแบบทดสอบสูงกว่าฉบับ X

ภาคผนวก จ

ผลการสร้างตารางเทียบคะแนนโดยวิธีเทียบคะแนนจริง

ตาราง 15 การคำนวณคะแนนที่ได้จากการเทียบคะแนนจริง

θ	x	y*	$Z_x - score$	$Z_y - score$	$CEEB_x$	$CEEB_y^*$
-3.0	2.1016	2.1865	-1.1661	-1.2132	383	378
-2.9	2.1500	2.2503	-1.1618	-1.2071	383	379
-2.8	2.2053	2.3237	-1.1568	-1.1999	384	380
-2.7	2.2686	2.4079	-1.1511	-1.1919	384	380
-2.6	2.3409	2.5045	-1.1446	-1.1784	385	382
-2.5	2.4235	2.6154	-1.1372	-1.1716	386	382
-2.4	2.5179	2.7423	-1.1288	-1.1593	387	384
-2.3	2.6257	2.8876	-1.1191	-1.1452	388	385
-2.2	2.7487	3.0536	-1.1081	-1.1291	389	387
-2.1	2.8892	3.2427	-1.0955	-1.1107	390	388
-2.0	3.0492	3.4578	-1.0812	-1.0898	391	391
-1.9	3.2316	3.7015	-1.0645	-1.0662	393	393
-1.8	3.4391	3.9766	-1.0462	-1.0395	395	396
-1.7	3.6748	4.2857	-1.0251	-1.0095	397	399
-1.6	3.9422	4.6311	-1.0011	-0.9759	399	402
-1.5	4.2447	5.0146	-0.9740	-0.9387	402	406
-1.4	4.5863	5.4374	-0.9434	-0.8879	405	411
-1.3	4.9706	5.9001	-0.9089	-0.8527	409	414
-1.2	5.4016	6.4026	-0.8703	-0.8039	412	419
-1.1	5.8827	6.9438	-0.8272	-0.7514	417	424
-1.0	6.4170	7.5222	-0.7793	-0.6952	422	430
-0.9	7.0067	8.1356	-0.7264	-0.6357	427	436
-0.8	7.6533	8.7811	-0.6685	-0.5730	433	442
-0.7	8.3563	9.4555	-0.6054	-0.5076	439	449
-0.6	9.1142	10.1550	-0.5375	-0.4396	446	456

ตาราง 15 (ต่อ)

θ	x	y^*	$Z_x - score$	$Z_y - score$	$CEEB_x$	$CEEB_y$
-0.5	9.9233	10.8753	-0.4649	-0.397	453	463
-0.4	10.7782	11.6117	-0.3883	-0.2982	461	470
-0.3	11.6719	12.3594	-0.3082	-0.2256	469	477
-0.2	12.5959	13.1134	-0.2254	-0.1524	477	484
-0.1	13.5408	13.8687	-0.4107	-0.0791	485	492
0.0	14.4966	14.6205	-0.0550	-0.0061	494	499
0.1	15.4535	15.3642	0.0307	0.0061	503	506
0.2	16.4020	16.0956	0.11576	0.1371	511	513
0.3	17.3332	16.8109	0.1992	0.2065	519	520
0.4	18.2395	17.5071	0.2805	0.2741	528	527
0.5	19.1145	18.1812	0.3589	0.3396	535	533
0.6	19.9530	18.8314	0.4341	0.4027	543	540
0.7	20.7511	19.4561	0.5056	0.4629	550	546
0.8	21.5063	20.0544	0.5733	0.5214	557	552
0.9	22.2170	20.6258	0.6370	0.5769	563	557
1.0	22.8826	21.1703	0.6967	0.6297	569	562
1.1	23.5034	21.6885	0.7523	0.6801	575	568
1.2	24.0802	22.1809	0.8040	0.7279	580	572
1.3	24.6143	22.6485	0.8519	0.7733	585	577
1.4	25.1075	23.0924	0.8961	0.8164	589	581
1.5	25.5615	23.5135	0.9368	0.8572	593	585
1.6	25.9786	23.9130	0.9742	0.8960	597	589
1.7	26.3608	24.2921	1.0084	0.9328	600	593
1.8	26.7104	24.6516	1.0398	0.9677	603	596
1.9	27.0296	24.9927	1.0684	1.0008	606	600
2.0	27.3204	25.3162	1.0945	0.0322	609	603

ตาราง 15 (ต่อ)

θ	x	y^*	$Z_x - score$	$Z_y - score$	$CEEB_x$	$CEEB_y$
2.1	27.5850	25.6228	1.1182	0.0620	611	606
2.2	27.8254	25.9134	1.3973	0.0920	613	609
2.3	28.0434	26.1886	1.1593	1.0069	615	611
2.4	28.2409	26.4489	1.1770	1.1442	617	614
2.5	28.4196	26.6950	1.1930	1.1661	619	616
2.6	28.5810	26.9274	1.2074	1.1887	620	618
2.7	28.7267	27.1465	1.2205	1.2099	622	620
2.8	28.8581	27.3529	1.2323	1.2299	623	622
2.9	28.9764	27.5470	1.2429	1.2488	624	624
3.0	29.0829	27.7293	1.2524	1.2665	625	626

ภาคผนวก ช

ผลการสร้างตารางเทียบคะแนนโดยวิธีเทียบคะแนนสังเกต

ตาราง 16 การคำนวณคะแนนที่ได้จากการเทียบคะแนนสังเกต

คะแนน	θ	X	Z_X -score	Y^*	Z_Y -score	CEEB _X	CEEB _Y *
0	-	0.0866	-1.6533	0.1064	-1.7372	-	-
1.	-	0.3456	-1.5810	0.4159	-1.7031	-	-
2	-	1.5645	-1.3256	1.6421	-1.6801	-	-
3	-	2.8713	-1.2954	3.1088	-1.5678	-	-
4	-3	3.1274	-1.2776	3.6431	-1.3471	370	365
5	-2.8	3.6453	-1.2421	3.9874	-1.3913	375	360
6	-2.1	5.3121	-1.0710	6.1532	-1.0586	392	394
7	-2.0	6.5942	-0.9394	7.9363	-0.8736	406	412
8	-1.9	7.4891	-0.8475	8.5310	-0.8080	415	419
9	-1.8	9.6436	-0.6263	10.4677	-0.5943	437	440
10	-1.7	10.6471	-0.5233	11.5931	-0.4702	447	453
11	-1.5	11.8707	-0.3977	12.7330	-0.3445	460	465
12	-1.2	12.6973	-0.3128	13.9584	-0.2093	468	479
13	-0.7	13.8644	-0.1930	14.5335	-0.1459	480	485
14	-0.5	14.3540	-0.1427	15.5617	-0.0325	485	496
15	0.2	15.7461	0.0002	16.7978	0.1039	500	510
16	0.3	16.1459	0.0413	17.4538	0.1762	504	517
17	0.4	17.3491	0.1679	17.4659	0.1776	516	517
18	0.5	18.8211	0.3159	18.6762	0.3111	531	531
19	0.6	19.1479	0.3494	18.1611	0.2543	534	525
20	0.8	20.6544	0.5041	19.2643	0.3759	550	537
21	1.4	21.7382	0.6154	19.388721.	0.3897	561	538
22	1.9	22.6120	0.7051	6740	0.6417	570	534
23	2.0	23.2124	0.7668	23.1073	0.7998	576	579
24	2.3	24.5781	0.9097	23.6675	0.8616	590	586
25	2.4	25.8973	1.0424	24.3381	0.9360	604	593
26	2.5	26.88	1.1433	25.4433	1.0575	614	605
27	2.7	27.675	1.2250	27.7664	1.3137	622	631
28	2.8	28.2375	1.2827	28.4003	1.3866	628	638
29	2.9	29.6774	1.4308	29.7466	1.5322	643	656
30	3.0	29.8366	1.4469	29.9639	1.5561	644	655

ภาคผนวก ซ

ผลการสร้างตารางเทียบคะแนนโดยวิธีอิกวิเปอร์เซนไทล์

ตาราง 17 คะแนนจากแบบทดสอบฉบับ X และแบบทดสอบ Y

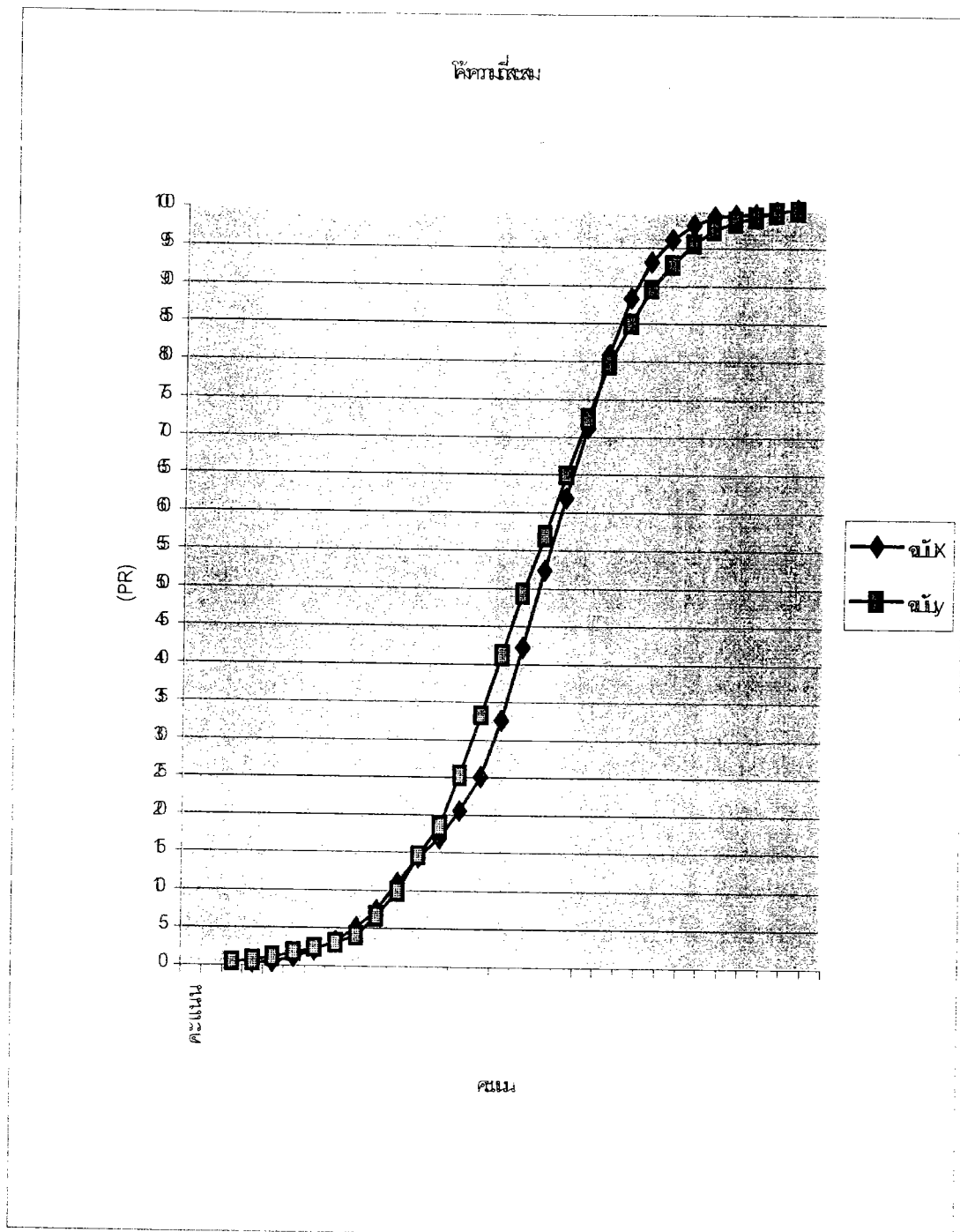
คะแนน	ความถี่ f จากแบบทดสอบ	
	ฉบับ X	ฉบับ Y
30	1	2
29	2	3
28	1	3
27	1	6
26	6	10
25	10	15
24	16	17
23	25	24
22	40	28
21	51	38
20	50	41
19	50	42
18	54	40
17	51	43
16	40	43
15	24	42
14	20	36
13	13	21
12	18	25
11	19	17
10	12	14
9	10	5
8	6	4
7	5	3
6	4	3
5	1	3
4	1	1
3	-	2
2	-	-
1	-	-
0	-	-
	531	531

ตาราง 18 ตำแหน่งเปอร์เซ็นต์ไทล์ของคะแนนจากแบบทดสอบ ฉบับ X และฉบับ Y

คะแนน	ตำแหน่งเปอร์เซ็นต์ไทล์ (PR) ของแบบทดสอบฉบับ	
	X	Y
30	100.00	100.00
29	99.81	99.62
28	99.44	99.06
27	99.24	98.49
26	99.05	97.36
25	97.93	95.48
24	96.05	92.66
23	93.003	89.46
22	88.32	84.93
21	80.79	79.66
20	71.19	72.50
19	61.77	64.78
18	52.35	56.87
17	42.18	49.34
16	32.58	41.24
15	25.05	33.15
14	20.53	25.24
13	16.76	18.46
12	14.31	14.50
11	10.92	9.79
10	7.34	6.59
9	5.08	3.95
8	3.20	3.01
7	2.07	2.26
6	1.13	1.69
5	0.38	1.13
4	0.19	0.56
3	-	0.38

ตาราง 18 -1 ตารางเทียบคะแนนแบบทดสอบฉบับ X และ Y

เปอร์เซ็นต์(PR)	คะแนนจากแบบทดสอบฉบับ	
	X	Y
100	30	30
99.8	29	29.3
99.4	28	28
97	25	27.2
96	24	26.4
95	23.8	26
93	23	25.2
88	22	23.2
80.8	21	22.1
71	20	20.8
61	19	19.2
52	18	18
42	17	16.2
32	16	15
25	15	14
20	14	1.2
16.5	13	12.6
14.5	12	12.2
10.5	11	11.2
9.5	10.8	10.8
7	10	10.3
5	9	9.6
3	8	8
2	7	7
1	6	6
0.5	5	5.1
0.38	4	5



ภาพประกอบ 9 โค้งความถี่สะสม แบบทดสอบฉบับ X และฉบับ Y

ตาราง 19 เทียบคะแนนแบบทดสอบฉบับ X และฉบับ Y ด้วยวิธีอีควิเปอร์เซนไทล์

เปอร์เซนไทล์	คะแนนจากแบบทดสอบ					
	X	$Z_x - score$	$CEEB_x$	Y^*	$Z_y - score$	$CEEB_y$
0.38	4	-1.4933	351	5	-1.4501	355
0.5	5	-1.372	363	5.1	-1.4378	357
1	6	-1.2509	375	6	-1.3260	368
2	7	-1.1296	388	7	-1.2017	380
3	8	-1.0084	400	8	-1.0775	393
5	9	-0.8872	412	9.6	-0.8788	413
7	10	-0.766	424	10.3	-0.7918	421
9.5	10.8	-0.669	434	10.8	-0.7269	428
10.5	11	-0.6448	436	11.2	-0.6800	432
14.5	12	-0.5236	448	12.2	-0.5557	445
16.5	13	-0.4020	460	12.6	-0.5060	450
20	14	-0.2812	472	13	-0.4563	455
25	15	-0.1600	484	14	0.1103	489
32	16	0.0387	503	15	0.043	496
42	17	0.0824	508	16.2	0.0035	500
52	18	0.2036	520	18	0.1647	516
61	19	0.3248	532	19.2	0.3138	531
71	20	0.4460	544	20.8	0.5125	551
80.8	21	0.5673	556	22.1	0.6740	567
88	22	0.6884	568	23.2	0.8107	581
93	23	0.8096	580	25.2	0.7762	577
95	23.8	0.9066	590	26	1.1585	615
96	24	0.9309	593	26.4	1.2081	620
97	25	1.0521	605	27.2	1.3076	630
99.4	28	1.4157	641	28	1.4069	640
99.8	29	1.5369	653	29.3	1.5684	656
100	30	1.6582	665	30	1.6554	665

ภาคผนวก ณ

ผลการสร้างตารางเทียบคะแนนทั้งสามวิธี

ตาราง 20 การเปรียบเทียบคะแนนที่ได้จากการเทียบคะแนนทั้งสามวิธี

คะแนน	θ	Z – score คะแนน จริง	Z-score คะแนน สังเกต	Z-score คะแนน Eq	คะแนน จริง	คะแนน สังเกต	คะแนน Eq	CEEB คะแนน จริง	CEEB คะแนน สังเกต	CEEB คะแนน Eq
0	-	-	-1.737	-	-	0.106	-	-	-	-
1	-	-	-1.703	-	-	0.415	-	-	-	-
2	-	-	-1.680	-	-	1.642	-	-	-	-
3	-	-	-1.567	-	-	3.108	-	-	-	-
4	-3.0	-1.213	-1.391	-1.450	2.186	3.643	5	378	370	355
5	-2.8	-1.199	-1.347	-1.437	2.323	3.987	5.1	384	375	357
6	-2.1	-1.110	-1.391	-1.326	3.242	6.153	6	390	392	368
7	-2.0	-1.089	-1.058	-1.261	3.457	7.936	7	391	406	380
8	-1.9	-1.066	-0.873	-1.077	3.701	8.531	8	393	415	393
9	-1.8	-1.004	-0.880	-0.878	3.976	10.46	9.6	395	437	413
10	-1.7	-0.101	-0.594	-0.791	4.285	11.59	10.3	397	447	421
11	-1.5	-0.938	-0.470	-0.726	5.014	12.73	10.8	402	460	428
12	-1.2	-0.804	-0.344	-0.68	6.402	13.95	11.2	412	468	432
13	-0.7	-0.507	-0.209	-0.557	9.455	14.53	12.2	439	480	445
14	-0.5	-0.368	-0.145	-0.506	10.87	15.56	12.6	453	485	450
15	0.2	0.137	-0.032	-0.580	16.09	16.79	13	511	500	455
16	0.3	0.236	0.123	0.110	16.81	17.45	14	519	504	489
17	0.4	0.274	0.176	0.043	17.50	17.49	15	528	516	496
18	0.5	0.339	0.177	0.003	18.18	18.67	16.2	535	531	500
19	0.6	0.462	0.311	0.164	18.83	18.16	18	543	534	516
20	0.8	0.521	0.254	0.3138	20.05	19.26	19.2	557	550	531
21	1.4	0.816	0.375	0.512	23.09	19.38	20.3	589	561	551
22	1.9	1.00	0.389	0.674	24.99	21.67	22.1	606	570	567
23	2.0	1.032	0.641	0.710	25.31	23.10	23.2	609	576	581
24	2.3	1.116	0.799	0.776	26.18	23.66	25.2	615	590	577
25	2.4	1.142	0.861	1.158	26.44	24.33	26	617	604	615
26	2.5	1.166	0.936	1.120	26.69	25.44	26.4	619	614	620
27	2.7	1.201	1.059	1.307	27.14	27.76	27.2	622	622	630
28	2.8	1.229	1.313	1.406	27.35	28.50	28	623	628	640
29	2.9	1.2488	1.386	1.568	27.54	29.74	29.3	624	643	656
30	3.0	1.266	1.532	1.65	27.72	29.96	30	625	644	665

ภาคผนวก ญ

คำดัชนีความสอดคล้อง (IOC)

ตาราง 21 ค่าดัชนีความสอดคล้อง (IOC) ของคำถามกับเนื้อหาและพฤติกรรมของแบบทดสอบ
ที่ใช้หาคุณภาพข้อสอบเป็นรายข้อ

ข้อ	IOC	ข้อ	IOC	ข้อ	IOC
1	1.0	35	1.0	69	1.0
2	1.0	36	1.0	70	0.8
3	0.6	37	1.0	71	1.0
4	0.8	38	0.8	72	0.8
5	0.8	39	1.0	73	0.8
6	1.0	40	1.0	74	0.6
7	1.0	41	0.8	75	1.0
8	1.0	42	1.0	76	0.8
9	0.8	43	1.0	77	1.0
10	1.0	44	0.8	78	0.6
11	1.0	45	1.0	79	1.0
12	0.8	46	1.0	80	1.0
13	1.0	47	0.8	81	0.8
14	0.6	48	0.8	82	0.8
15	1.0	49	0.6	83	1.0
16	0.6	50	1.0	84	1.0
17	1.0	51	1.0	85	1.0
18	0.6	52	1.0	86	0.6
19	0.8	53	1.0	87	0.8
20	0.6	54	1.0	88	1.0
21	0.8	55	0.8	89	1.0
22	0.8	56	0.8	90	0.6
23	1.0	57	0.6	91	1.0
24	0.6	58	0.6	92	0.8
25	0.6	59	0.6	93	1.0
26	1.0	60	1.0	94	1.0
27	1.0	61	1.0	95	1.0
28	0.8	62	1.0	96	1.0
29	0.8	63	1.0	97	1.0
30	0.8	64	1.0	98	1.0
31	0.6	65	0.6	99	0.8
32	1.0	66	1.0	100	1.0
33	1.0	67	1.0		
34	1.0	68	1.0		

ภาคผนวก ก

รายชื่อผู้เชี่ยวชาญ

รายชื่อผู้เชี่ยวชาญ

รศ.นิภา ศรีไพโรจน์	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
อ.ชวลิต รวยอาจิณ	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
อ.ระวีวรรณ พันธุ์พานิช	ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
อ.ทรงสิทธิ์ เจริญดี	อาจารย์ 3 ระดับ 8 โรงเรียนบ้านบึงกาฬ สปอ.บึงกาฬ สปจ.หนองคาย
ผอ.ประยูร นามวงศ์	ผู้อำนวยการโรงเรียนบ้านสุขสำราญ สปอ.บึงกาฬ สปจ.หนองคาย

ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ	นางสาว ยุพิน หลายเจริญ
วันเดือนปีเกิด	12 ธันวาคม 2514
สถานที่เกิด	83 หมู่ 4 บ้านนาโนน ต.บึงกาฬ อ.บึงกาฬ จ.หนองคาย 43140
สถานที่อยู่ปัจจุบัน	83 หมู่ 4 บ้านนาโนน ต.บึงกาฬ อ.บึงกาฬ จ.หนองคาย 43140
ตำแหน่งหน้าที่การงาน	อาจารย์ 1 ระดับ 4
สถานที่ทำงาน	โรงเรียนบ้านท่าโพธิ์ ต.บึงกาฬ อ.บึงกาฬ จ.หนองคาย 43140
ประวัติการศึกษา	
พ.ศ. 2533	มัธยมศึกษาปีที่ 6 โรงเรียนบึงกาฬ อ.บึงกาฬ จ. หนองคาย
พ.ศ. 2537	ค.บ.(เกียรตินิยมอันดับ 2) วิชาเอกการประถมศึกษา ในโครงการคุรุทายาท สถาบันราชภัฏเลย
พ.ศ. 2545	กศ.ม. (วิชาเอกการวัดผลการศึกษา) จาก มหาวิทยาลัย ศรีนครินทรวิโรฒ