

372.4024

11 2544

การเปรียบเทียบผลการเทียบคะแนนโดยวิธีเชิงเส้นตรง เมื่อใช้วิธีการแปลงคะแนนที่แตกต่างกัน
ในแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

ปริญญานิพนธ์
ของ
ทิพาภรณ์ ทุงฤทธิ์

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
มีนาคม 2544

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

7 2544

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปริญญานิพนธ์ฉบับนี้แล้ว
เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิต
วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

คณะกรรมการควบคุม

.....ประธาน
(รองศาสตราจารย์อังคณา สายยศ)
.....กรรมการ
(อาจารย์ ดร.สุพร เข้มเฮง)

คณะกรรมการสอบ

.....ประธาน
(รองศาสตราจารย์อังคณา สายยศ)
.....กรรมการ
(อาจารย์ ดร.สุพร เข้มเฮง)
.....กรรมการที่แต่งตั้งเพิ่มเติม
(อาจารย์ชวลิต รวยอาจิน)
.....กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ชูศรี วงศ์รัตนะ)

บัณฑิตวิทยาลัยอนุมัติให้รับปริญญานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัย
ศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย
(ศาสตราจารย์ ดร.เสริมศักดิ์ วิศาลาภรณ์)
วันที่ 9 เดือน มิถุนายน พ.ศ. 2544

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จลงได้ด้วยดี เพราะได้รับความกรุณาอย่างดียิ่งในการให้ความรู้ ให้คำปรึกษา แนะนำ ตลอดจนตรวจแก้ไขข้อบกพร่องต่างๆ จากรองศาสตราจารย์อังคณา สายยศ ประธานกรรมการควบคุมปริญญานิพนธ์ และดร.สุพร เข้มแข็ง กรรมการควบคุมปริญญานิพนธ์ ผู้วิจัยรู้สึกซาบซึ้งในความกรุณาและเอาใจใส่จากท่านอาจารย์ทั้งสองเป็นอย่างยิ่ง ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้

ในการตรวจสอบคุณภาพของเครื่องมือที่ใช้ในการวิจัย ผู้วิจัยขอขอบพระคุณ รศ.ดร.ส.วาสนา ประवालพุกภักดิ์ ผศ.ระวีวรรณ พันธุ์พานิช อาจารย์ชวลิต รวยอาจิณ ศึกษานิเทศกิจจิตร ทองสุข และอาจารย์สมพร นวลอุทัย ที่กรุณาให้คำแนะนำและเป็นผู้เชี่ยวชาญตรวจสอบคุณภาพเครื่องมือ

นอกจากนี้ผู้วิจัยขอขอบคุณสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี โรงเรียนในสังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี ตลอดจนนักเรียนชั้นประถมศึกษาปีที่ 6 ทุกคนที่ให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี และขอขอบคุณคุณคุณอนันต์ สงเคราะห์ที่ให้ความช่วยเหลือและช่วยอำนวยความสะดวกในการเก็บข้อมูล ตลอดจนช่วยเป็นกำลังใจ ให้คำปรึกษา แนะนำ ตลอดการทำงานวิจัยครั้งนี้จนสำเร็จลุล่วงไปได้ด้วยดี

ผู้วิจัยขอกราบน้อมระลึกถึงพระคุณอันสูงยิ่งของบิดามารดา ครูอาจารย์ทุกท่าน ซึ่งได้เมตตาให้การสนับสนุน ให้กำลังใจ และประสิทธิ์ประสาทความรู้ให้แก่ผู้วิจัยอย่างยิ่ง

ทิพาภรณ์ หงุฤทธิ์

สารบัญ

บทที่	หน้า
1 บทนำ.....	1
ภูมิหลัง.....	1
ความมุ่งหมายของการวิจัย.....	3
ความสำคัญของการวิจัย.....	3
ขอบเขตของการวิจัย.....	3
นิยามศัพท์เฉพาะ.....	4
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	6
ความหมายของการเทียบคะแนน.....	6
เงื่อนไขของการเทียบคะแนน.....	7
ประเภทของการเทียบคะแนน.....	8
แบบแผนการเก็บรวบรวมข้อมูล.....	8
รูปแบบการเทียบคะแนน.....	12
การเทียบคะแนนเชิงเส้นตรง.....	12
การเทียบคะแนนรูปแบบอิกวิเปอร์เซ็นต์ไทล์.....	15
การเทียบคะแนนรูปแบบอิงทฤษฎีการตอบข้อสอบ.....	16
วิธีการแปลงคะแนนตามวิธีการเทียบคะแนนเชิงเส้นตรง.....	19
ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน.....	22
ความเพียงพอของการเทียบคะแนน.....	23
งานวิจัยที่เกี่ยวข้อง.....	25
สมมติฐานในการวิจัย.....	29
3 วิธีการดำเนินการศึกษาค้นคว้า.....	30
ประชากร.....	30
กลุ่มตัวอย่าง.....	30
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	31
วิธีการดำเนินการสร้างแบบทดสอบ.....	32
วิธีการดำเนินการเก็บรวบรวมข้อมูล.....	34
ขั้นตอนการวิเคราะห์ข้อมูล.....	35
สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	35
4 ผลการวิเคราะห์ข้อมูล.....	41
สัญลักษณ์ที่ใช้ในการวิเคราะห์ข้อมูล.....	41
การนำเสนอผลการวิเคราะห์ข้อมูล.....	42

บทที่	หน้า
ผลการวิเคราะห์ข้อมูล.....	42
5 สรุปผล อภิปราย และข้อเสนอแนะ.....	47
ความมุ่งหมายของการวิจัย.....	47
กลุ่มตัวอย่าง.....	47
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	47
วิธีดำเนินการเก็บรวบรวมข้อมูล.....	47
การวิเคราะห์ข้อมูล.....	48
สรุปผลการวิเคราะห์ข้อมูล.....	48
อภิปรายผล.....	49
ข้อเสนอแนะ.....	50
บรรณานุกรม.....	51
ภาคผนวก.....	55
ก รายชื่อผู้เชี่ยวชาญที่ตรวจสอบคุณภาพเครื่องมือ.....	56
ข ตารางวิเคราะห์หลักสูตรวิชาคณิตศาสตร์.....	58
ค คุณภาพของแบบทดสอบจากการทดลองเครื่องมือ.....	62
ง แบบทดสอบที่ใช้ในการเทียบคะแนน 3 ฉบับ.....	76
จ คะแนนที่ได้จากการเทียบคะแนน 3 วิธี.....	85
ฉ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน 3 วิธี.....	87
ช ความเพียงพอของการเทียบคะแนน 3 วิธี.....	89
ประวัติย่อผู้วิจัย.....	91

บัญชีตาราง

ตาราง	หน้า
1 แบบแผนที่มีผู้สอบกลุ่มเดียว.....	9
2 แบบแผนที่มีผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล.....	9
3 แบบแผนที่มีผู้สอบกลุ่มเท่าเทียมกัน.....	9
4 แบบแผนที่มีผู้สอบกลุ่มสองกลุ่มโดยมีแบบทดสอบร่วม.....	10
5 แบบแผนที่มีกลุ่มผู้สอบที่ไม่เท่าเทียมกันโดยมีแบบทดสอบร่วม.....	10
6 จำนวนนักเรียนในกลุ่มตัวอย่างที่ใช้ในการเทียบคะแนน.....	31
7 รายละเอียดในการจัดลำดับการสอบเพื่อนำไปเทียบคะแนน.....	34
8 ค่าสถิติพื้นฐานของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์.....	42
9 คะแนนที่ได้จากการเทียบจากแบบทดสอบฉบับ X ไปสู่แบบทดสอบฉบับ Y ในรูปแบบการเทียบคะแนนทั้ง 3 วิธี.....	42
10 การเปรียบเทียบความแตกต่างของคะแนนที่เทียบได้จากการเทียบคะแนน 3 วิธี.....	43
11 การเปรียบเทียบความแตกต่างรายคู่ของการเทียบคะแนน 3 วิธี.....	43
12 ช่วงคะแนนและความคลาดเคลื่อนมาตรฐานของคะแนนที่เทียบได้ในการเทียบคะแนน 3 วิธี.....	44
13 การเปรียบเทียบความแตกต่างของค่าความคลาดเคลื่อนมาตรฐาน ที่ได้จากรูปแบบการเทียบคะแนน 3 วิธี.....	44
14 ความเพียงพอของการเทียบคะแนน 3 วิธี.....	45
15 การเปรียบเทียบความเพียงพอของการเทียบคะแนน 3 วิธี.....	45
16 คะแนนที่ได้จากการเทียบจากแบบทดสอบฉบับ X ไปสู่แบบทดสอบฉบับ Y ในรูปแบบการเทียบคะแนนทั้ง 3 วิธี.....	86
17 ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในแต่ละระดับคะแนน ในรูปแบบการเทียบทั้ง 3 วิธี.....	88
18 ความเพียงพอของการเทียบคะแนนทั้ง 3 วิธี ที่ระดับคะแนนเดียวกัน.....	90

บัญชีภาพประกอบ

ภาพประกอบ

หน้า

1	ขั้นตอนการสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์.....	32
---	--	----

งานวิจัยนี้ได้รับทุนสนับสนุนจาก
ดร.ชวาล แพรัตน์กุล
ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

บทที่ 1

บทนำ

ภูมิหลัง

ปัจจุบันการทดสอบเป็นกระบวนการที่สำคัญอย่างหนึ่งในระบบการศึกษา จะเห็นได้ว่า เริ่มตั้งแต่ก่อนเข้ารับการศึกษา ในขณะที่ศึกษา ก่อนจบการศึกษา จนกระทั่งสมัครเข้าทำงาน จะต้องมีการทดสอบทั้งสิ้น และในการทดสอบแต่ละครั้งเครื่องมือที่ใช้ในการทดสอบที่ใช้กันมากที่สุดคือ แบบทดสอบ ซึ่งแบบทดสอบที่ดีนั้นสงบ ลักษณะ (2525 : 47) ได้กล่าวไว้ว่า จะต้องสามารถวัดได้ตรงกับความสามารถที่แท้จริงของแต่ละบุคคล และเกิดความยุติธรรมแก่ทุกคนด้วย จากที่มีการทดสอบบ่อยครั้ง และเป็นสิ่งที่แทบจะหลีกเลี่ยงไม่ได้ของแต่ละบุคคล จนเป็นเหตุกระตุ้นให้ผู้สอบเกิดความกระตือรือร้นและหาหนทางเพิ่มพูนคะแนนด้วยวิธีการต่างๆ รวมถึงการฝึกซ้อมความชำนาญในการทำข้อสอบจากข้อสอบเก่าๆ ที่ใช้มาแล้ว จึงทำให้ผู้ที่รับผิดชอบเกี่ยวกับการสอบจะต้องจัดทำแบบทดสอบขึ้นใหม่ และแบบทดสอบที่สร้างขึ้นใหม่นี้ จะต้องสามารถวัดคุณลักษณะ (Trait) เดียวกัน เพื่อที่จะให้เกิดความเหมาะสมกับสภาพการทดสอบในแต่ละสถานการณ์ จึงทำให้เกิดความคิดในการสร้างแบบทดสอบคู่ขนาน ซึ่งในการสร้างแบบทดสอบคู่ขนานนั้น มักจะประสบปัญหาคือไม่สามารถจะสร้างแบบทดสอบให้คู่ขนานกันได้อย่างสมบูรณ์ เพราะในทางปฏิบัตินั้นมักจะทำได้เพียงให้แบบทดสอบทั้งสองฉบับเกิดความใกล้เคียงด้านเนื้อหา ค่าสถิติของแบบทดสอบรายข้อและแบบทดสอบทั้งฉบับเท่ากัน (Gulliksen, 1950 : 174) จึงมีผู้คิดค้นวิธีการที่จะทำให้คะแนนจากแบบทดสอบที่วัดคุณลักษณะเดียวกันแต่ต่างฉบับกัน สามารถนำมาเปรียบเทียบกันได้อย่างมีความหมายขึ้น ทั้งนี้เพื่อประโยชน์ในการดำเนินการสอบให้เกิดความยุติธรรม และเกิดความเหมาะสมกับสถานการณ์ในการสอบ (นิคม ดังคะพิภพ. ม.ป.ป. : 8-1) วิธีการดังกล่าวคือ การเทียบคะแนน (Equating) ซึ่งสามารถจำแนกได้เป็นสองลักษณะใหญ่ๆ (สงบ ลักษณะ. 2525 : 25 ; อ้างอิงมาจาก วิรัช วรรณรัตน์. 2525 : 69-70) คือ

1. การเทียบคะแนนในแนวนอนหรือแนวนอน (Horizontal Equating) เป็นการเทียบคะแนนระหว่างแบบทดสอบสองชุดที่มีความยากพอๆ กัน กล่าวคือ ต้องการเทียบคะแนนจากแบบทดสอบฉบับหนึ่งว่าจะเท่ากับที่หน่วยของแบบทดสอบอีกฉบับหนึ่ง ดังนั้นในการทดสอบจะต้องสอบบุคคลกลุ่มเดียวกัน

2. การเทียบคะแนนในแนวตั้งหรือแนวตั้ง (Vertical Equating) เป็นการเทียบคะแนนระหว่างแบบทดสอบสองชุดที่มีความยากต่างกัน เมื่อนำไปสอบกับบุคคลที่มีความสามารถต่างกัน กล่าวคือ ต้องการเทียบคะแนนจากแบบทดสอบสองชุด เมื่อความยากของข้อทดสอบต่างกันและเมื่อสอบกับบุคคลที่มีความสามารถต่างกันซึ่งการเทียบคะแนนในระดับต่างๆ ของแบบทดสอบเป็นสิ่งที่พึงปรารถนา เพื่อว่าผู้สอบจะได้รับคะแนนที่เหมือนกัน ไม่ว่าเขาจะสอบฉบับที่ง่ายหรือฉบับที่ยาก

ข้อตกลงเบื้องต้นของการเทียบคะแนนคือ แบบทดสอบสองฉบับใดๆ จะเทียบคะแนนกันได้ก็ต่อเมื่อแบบทดสอบสองฉบับนั้นคู่ขนานกันในด้านต่อไปนี้

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

ในการเทียบคะแนน เป็นการกำหนดเงื่อนไขและสร้างกฎการแปลงคะแนนในแต่ละครั้งที่ทำการเทียบคะแนน ซึ่งมีขั้นตอนในการคำนวณที่ยาก และต้องใช้กลุ่มประชากรค่อนข้างมากในการเทียบคะแนนแต่ละครั้ง ทำให้เป็นเรื่องยุ่งยากในการนำไปใช้ในวิชาชีพครูในสภาพการณ์จริง ผู้วิจัยจึงต้องการศึกษารูปแบบของการเทียบคะแนนที่ไม่ยากนักในการคำนวณ และใช้กลุ่มตัวอย่างไม่มาก สามารถทำการเทียบคะแนนในระดับประถมศึกษาในสภาพจริงได้ โดยมุ่งศึกษาผลการเทียบคะแนนเชิงเส้นตรงด้วยเหตุผล 3 ประการ ดังนี้

1. การเทียบคะแนนเชิงเส้นตรง เป็นวิธีการแปลงคะแนนที่ง่าย และสะดวกที่สุดที่จะใช้ (Kolen & Whitney. 1982 : 291)

2. การเทียบคะแนนเชิงเส้นตรง ให้ผลดีกว่ารูปแบบอื่นในสถานการณ์สอบวัดผลสัมฤทธิ์ (Petersen, Cook & Stocking. 1983 : 155) เนื่องจากในการทดสอบเพื่อวัดผลสัมฤทธิ์ทางการเรียนตามหลักสูตรในชั้นเรียนมีขอบเขตเนื้อหาที่จำกัดในหลักสูตร แบบสอบชุดใหม่ที่สร้างขึ้นจะประกอบด้วยข้อสอบที่มีลักษณะต่างๆ ใกล้เคียงกับแบบทดสอบที่สร้างขึ้นในชุดก่อนๆ เมื่อแบบทดสอบวัดเนื้อหาเดียวกัน วิธีการเทียบคะแนนเชิงเส้นตรงจึงให้ผลดีกว่าวิธีอื่น

3. การเทียบคะแนนเชิงเส้นตรงยังคงใช้ได้ดีกับกลุ่มตัวอย่างขนาดเล็ก โดยเฉพาะการเทียบคะแนนในแนวนอน (Kolen & Whitney. 1982 : 292) ครู-อาจารย์ยังสามารถนำไปปฏิบัติได้จริงในสภาพการเรียนการสอนปัจจุบันในชั้นเรียนซึ่งมีจำนวนกลุ่มตัวอย่างไม่มากนัก

ดังนั้นในการวิจัยครั้งนี้ ผู้วิจัยจึงมีความสนใจที่จะศึกษาในเชิงเปรียบเทียบเกี่ยวกับผลของการเทียบคะแนนเชิงเส้นตรงเมื่อมีวิธีการแปลงคะแนนแตกต่างกันตามที่แองกอฟฟ์ได้เสนอไว้ 3 วิธีคือ (Angoff. 1971 : 569-586)

1. วิธีการแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ (Random Group-One Test Administered to Each Group) รูปแบบนี้ผู้สอบมาจากการสุ่มประชากรกลุ่มเดียวกัน แบ่งเป็น 2 กลุ่ม คือ กลุ่ม ที่ 1 สอบแบบทดสอบฉบับ X และกลุ่มที่ 2 สอบแบบทดสอบฉบับ Y

2. วิธีการแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบทั้งสองฉบับที่เรียงอันดับการสอบต่างกัน (Random Groups-Both Forms Administered to Each Group, Counterbalanced) ลักษณะเช่นเดียวกับรูปแบบที่ 1 เลือกกลุ่มตัวอย่างขนาดใหญ่แล้วสุ่มแบ่งออกเป็น 2 กลุ่ม โดยกลุ่มหนึ่งเป็นกลุ่มที่ 1 สอบแบบทดสอบฉบับ X แล้วตามด้วยแบบทดสอบฉบับ Y อีกกลุ่มหนึ่งเป็นกลุ่มที่ 2 สอบแบบทดสอบฉบับ Y แล้วตามด้วยแบบทดสอบฉบับ X เพื่อป้องกันความคลาดเคลื่อนในการสอบ

3. วิธีการแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ โดยเพิ่มแบบทดสอบร่วมที่ต้องสอบทั้งสองกลุ่ม (Random Groups-One Test Administered to Each Group, Common Equating Test Administered to Both Groups) รูปแบบนี้เพิ่มแบบทดสอบร่วมขึ้นมาอีกหนึ่งฉบับ เป็นแบบทดสอบร่วม เพื่อใช้ปรับความแตกต่างที่อาจจะพบระหว่างกลุ่มที่ 1 และกลุ่มที่ 2 ในระหว่างการดำเนินการสอบ กลุ่มที่ 1 จะได้รับแบบทดสอบฉบับ X และแบบทดสอบร่วม และกลุ่มที่ 2 จะได้รับแบบทดสอบฉบับ Y และแบบทดสอบร่วม

จาก 3 วิธีดังกล่าวนี้ ผู้วิจัยต้องการจะศึกษาว่า ผลการเทียบคะแนนทั้ง 3 วิธีนี้มีความแตกต่างกันอย่างไร และวิธีการเทียบคะแนนวิธีใดจะมีความเหมาะสมในการเทียบมากที่สุดที่จะสามารถนำไปปฏิบัติได้จริงในสภาพการเรียนการสอนในปัจจุบัน

ความมุ่งหมายของการวิจัย

1. เพื่อเปรียบเทียบคะแนนที่เทียบได้ของการเทียบคะแนนในแนวนอน โดยวิธีเชิงเส้นตรง เมื่อมีการแปลงคะแนนที่แตกต่างกัน
2. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในแนวนอน โดยวิธีเชิงเส้นตรง เมื่อมีการแปลงคะแนนที่แตกต่างกัน
3. เพื่อเปรียบเทียบความเพียงพอของการเทียบคะแนนในแนวนอน โดยวิธีเชิงเส้นตรง เมื่อมีการแปลงคะแนนที่แตกต่างกัน

ความสำคัญของการวิจัย

ผลการศึกษาครั้งนี้ทำให้ทราบว่า การแปลงคะแนนที่แตกต่างกัน 3 วิธีนี้จะให้ผลการเทียบคะแนนเชิงเส้นตรงที่แตกต่างกันหรือไม่ นอกจากนั้นยังพิจารณาว่าวิธีการเทียบคะแนนแบบใดมีคุณภาพในการเทียบคะแนนมากที่สุด โดยพิจารณาจากค่าความคลาดเคลื่อนมาตรฐาน และดัชนีความเพียงพอในการเทียบคะแนน ผลจากการเทียบคะแนนครั้งนี้จะเป็นประโยชน์อย่างยิ่งในการนำหลักการเทียบคะแนนไปใช้ในการวัดและประเมินผลการเรียนการสอนในชั้นเรียนจริงๆ เนื่องจากเป็นวิธีที่สะดวกและง่ายต่อการปฏิบัติ และนำไปใช้ในวิชาชีพครูต่อไป

ขอบเขตของการวิจัย

1. ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี ปีการศึกษา 2542 จำนวน 5,290 คน
2. กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี จำนวน 808 คน ได้มาจากการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีขนาดโรงเรียนเป็นชั้น (Strata) และมีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit)
3. ตัวแปรที่ใช้ในการศึกษาค้นคว้า
 - 3.1 ตัวแปรอิสระ คือ การเทียบคะแนนโดยวิธีเชิงเส้นตรงที่ใช้วิธีการแปลงคะแนนต่างกัน 3 วิธี ดังนี้
 - 3.1.1 วิธีการแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ (Random Group-One Test Administered to Each Group)
 - 3.1.2 วิธีการแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบทั้งสองฉบับที่เรียงอันดับการสอบต่างกัน (Random Groups-Both Forms Administered to Each Group, Counterbalanced)
 - 3.1.3 วิธีการแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ และมีแบบทดสอบร่วมที่ต้องสอบทั้งสองกลุ่ม (Random Groups-One Test Administered to Each Group, Common Equating Test Administered to Both Groups)
 - 3.2 ตัวแปรตาม คือ
 - 3.2.1 คะแนนที่เทียบได้
 - 3.2.2 ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน
 - 3.2.3 ความเพียงพอของการเทียบคะแนน

นิยามศัพท์เฉพาะ

1. การเทียบคะแนน (Equating) หมายถึง กระบวนการที่ใช้วิธีการทางสถิติเพื่อแปลงคะแนนที่ได้จากแบบทดสอบต่างฉบับให้มาอยู่ในคะแนนมาตรฐานเดียวกันหรือหน่วยวัดเดียวกัน

2. การเทียบคะแนนในแนวนอน (Horizontal Equating) หมายถึง การเทียบคะแนนระหว่างแบบทดสอบสองฉบับที่มีความยากพอๆ กัน และทดสอบโดยใช้บุคคลในระดับชั้นเดียวกัน

3. การเทียบคะแนนเชิงเส้นตรง (Linear Equating) หมายถึง กระบวนการที่ใช้วิธีการทางสถิติเพื่อปรับคะแนนจากแบบทดสอบต่างฉบับกัน ที่มีโครงสร้างเนื้อหาเดียวกัน ให้สามารถเปรียบเทียบกันได้ โดยยึดหลักการแปลงคะแนนให้เป็นคะแนนมาตรฐาน และกำหนดให้คะแนนมาตรฐานจากแบบทดสอบสองฉบับมีค่าเท่ากัน

4. วิธีการแปลงคะแนน หมายถึง การนำคะแนนที่ได้จากการสอบแบบทดสอบสองฉบับมาเปรียบเทียบกัน โดยยึดคะแนนที่ได้จากการสอบของนักเรียนที่มีวิธีการสอบแตกต่างกัน ดังนี้

4.1 วิธีการแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ (Random group-one test administered to each group)

4.2 วิธีการแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบทั้งสองฉบับที่เรียงอันดับการสอบต่างกัน (Random Group-Both Forms Administered to Each Group, Counterbalanced)

4.3 วิธีการแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ และมีแบบทดสอบร่วมที่ต้องสอบทั้งสองกลุ่ม (Random Groups-One Test Administered to Each Group, Common Equating Test Administered to Both Groups)

5. แบบทดสอบเทียบคะแนน (Equating Test) หมายถึง แบบทดสอบที่ใช้ในการสอบกลุ่มตัวอย่างกลุ่มที่ 1 และกลุ่มที่ 2 ซึ่งในการวิจัยครั้งนี้เป็นแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ เรื่อง สมการและการแก้สมการ ชั้นประถมศึกษาปีที่ 6 จำนวน 2 ฉบับ คือ ฉบับ X และ ฉบับ Y ฉบับละ 20 ข้อ

6. แบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ (Mathematics Achievement Test) หมายถึง แบบทดสอบที่วัดความรู้ของนักเรียนที่ได้เรียนไปแล้วในวิชาคณิตศาสตร์ เรื่อง สมการและการแก้สมการ ซึ่งมีข้อคำถามเป็นแบบเลือกตอบ 5 ตัวเลือก

7. แบบทดสอบร่วม (Common Equating Test) หมายถึง แบบทดสอบที่ใช้สอบกับกลุ่มตัวอย่างทั้งสองกลุ่ม และนำคะแนนที่ได้จากแบบทดสอบนี้ไปหาความเพียงพอของการเทียบคะแนน ซึ่งในการวิจัยครั้งนี้เป็นแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ เรื่อง สมการและการแก้สมการ ที่มีข้อคำถามไม่ซ้ำกับแบบทดสอบฉบับ X และแบบทดสอบฉบับ Y จำนวน 10 ข้อ

8. คะแนนที่เทียบได้ หมายถึง คะแนนที่ได้จากการเทียบคะแนนในวิธีการแปลงคะแนนตามข้อที่ 4.1, 4.2 และ 4.3

9. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน หมายถึง ความคลาดเคลื่อนที่เกิดจากการนำคะแนนที่แปลงแล้วจากแบบทดสอบฉบับ X มีคะแนนเบี่ยงเบนไปจากกับคะแนนเดิมของแบบทดสอบฉบับเดียวกัน ในการแปลงคะแนนในแต่ละวิธี

10. กลุ่มเทียบคะแนน หมายถึง กลุ่มตัวอย่างที่ได้ทำการทดสอบแตกต่างกันตามวิธีการแปลงคะแนนแต่ละวิธี และนำผลการสอบไปใช้ในการแปลงคะแนนสำหรับการสร้างตารางเทียบคะแนน

11. กลุ่มทานผล หมายถึง กลุ่มตัวอย่างที่ได้รับการสอบแบบทดสอบทั้งสองฉบับ และแบบทดสอบร่วม เพื่อนำผลการสอบไปใช้ในการตรวจสอบหาความเพียงพอของการแปลงคะแนนในแต่ละวิธี

12. ความเพียงพอของการเทียบคะแนน หมายถึง ค่าที่แสดงถึงคุณภาพของการเทียบคะแนนโดยยึดหลักการประเมินของปีเตอร์เซนและคนอื่นๆ (Petersen and others. 1982 : 93-94) ซึ่งมีเกณฑ์ในการพิจารณา ดังนี้

ระดับที่น่าพอใจอย่างมาก	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	$C > (.20SD_x)^2$

13. ผู้เชี่ยวชาญ หมายถึง บุคคลที่มีความรู้ ความสามารถ และประสบการณ์ทางการวัดผลการศึกษา และทำการสอนวิชาคณิตศาสตร์มาเป็นเวลาไม่น้อยกว่า 5 ปี จำนวน 5 ท่าน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยขอเสนอรายละเอียดของเอกสาร และงานวิจัยที่เกี่ยวข้องดังนี้

1. ความหมายของการเทียบคะแนน
2. เงื่อนไขของการเทียบคะแนน
3. ประเภทของการเทียบคะแนน
4. แบบแผนของการเก็บรวบรวมข้อมูล
5. รูปแบบการเทียบคะแนน
6. วิธีการแปลงคะแนนตามวิธีการเทียบคะแนนเชิงเส้นตรง
7. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน
8. ความเพียงพอของการเทียบคะแนน
9. งานวิจัยที่เกี่ยวข้อง

1. ความหมายของการเทียบคะแนน

นักการศึกษาหลายท่านได้ให้ความหมายของการเทียบคะแนนไว้ดังนี้

กัลลิกเซน (Gulliksen, 1950 : 298-304) ให้ความหมายของการเทียบคะแนนว่า เป็นการนำคะแนนจากแบบทดสอบสองฉบับในวิชาเดียวกัน มาทดสอบกับกลุ่มตัวอย่างที่คล้ายคลึงกัน แล้วนำมาเปรียบเทียบกัน โดยมีข้อแม้ว่าต้องแปลงคะแนนให้เป็นมาตรฐานเดียวกัน มีการกระจายคะแนนเป็นโค้งปกติ หรือทำเป็นคะแนนเปอร์เซ็นต์ไทล์ แล้วจึงนำมาเปรียบเทียบกันได้

ฟลานาแกน (Flanagan, 1951 : 745-747) ให้ความหมายของการเทียบคะแนนว่า เป็นกระบวนการเปรียบเทียบคะแนนจากแบบทดสอบสองฉบับในวิชาเดียวกัน และทดสอบกับกลุ่มประชากรเดียวกัน คะแนนที่ได้มาจะมีการกระจายของคะแนนจริงที่เหมือนกัน

แองกอฟ (Angoff, 1971 : 562) ให้ความหมายของการเทียบคะแนนว่า เป็นวิธีการเปลี่ยนระบบหน่วยคะแนนของแบบทดสอบฉบับหนึ่งไปสู่ระบบหน่วยคะแนนของแบบทดสอบอีกฉบับหนึ่ง เพื่อให้ได้คะแนนที่แปลงจากแบบทดสอบทั้งสองฉบับมีความเท่าเทียมกัน

ปีเตอร์เซนและคนอื่นๆ (Petersen and others, 1989 : 242) กล่าวว่า วิธีการเทียบคะแนนเป็นวิธีการเชิงประจักษ์ (Empirical Procedure) ในการสร้างความสัมพันธ์ระหว่างคะแนนดิบ (Raw Score) จากแบบทดสอบสองฉบับที่สามารถใช้คะแนนจากแบบทดสอบฉบับหนึ่งไปสู่อีกฉบับหนึ่งได้

สงบ ลักษณะ (2525 : 22) ได้กล่าวถึงการเทียบคะแนนไว้ว่า การเทียบคะแนนเป็นการนำคะแนนจากแบบทดสอบ ต่างฉบับกันที่วัดในสิ่งเดียวกัน แต่ไม่จำเป็นต้องเป็นข้อสอบคู่ขนานกัน จะถือว่าเทียบเท่ากันได้ ถ้านักเรียนคนใดก็ตามทำข้อสอบต่างฉบับแล้ว ได้คะแนนจริงหรือความสามารถที่แท้จริงเท่ากัน

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

1. กระบวนการที่ทำให้แบบทดสอบสองฉบับใดๆ มีความทัดเทียมกันหรือเท่ากันในเชิงของโครงสร้าง
2. การใช้วิธีการทางสถิติเพื่อปรับ (Adjust) คะแนนที่ได้จากแบบทดสอบแต่ละฉบับให้อยู่ในมาตรฐานเดียวกันและเทียบกันได้

วิรัช วรรณรัตน์ (2530 : 69) ได้ให้ความหมายของการเทียบคะแนนไว้ว่า เป็นการนำเอาคะแนนที่สังเกตได้จากแบบทดสอบต่างฉบับ ซึ่งแต่ละฉบับต่างวัดในสิ่งเดียวกันมาเทียบกันหรือกล่าวได้ว่าการเทียบคะแนน เป็นวิธีการเปลี่ยนระบบคะแนนของแบบทดสอบฉบับหนึ่ง ไปสู่ระบบคะแนนของแบบทดสอบอีกฉบับหนึ่ง

จากความหมายที่กล่าวมานั้น พอสรุปได้ว่า การเทียบคะแนนเป็นการใช้กระบวนการทางสถิติ ที่ทำให้คะแนนจากแบบทดสอบต่างฉบับกันที่วัดในเนื้อหาเดียวกัน มาเปรียบเทียบกับกันได้ในระบบคะแนนมาตรฐานเดียวกัน

2. เงื่อนไขของการเทียบคะแนน

แฮมเบิลตัน และสวามินาธานกล่าวว่า การเทียบคะแนนของแบบทดสอบสองฉบับจะต้องมีข้อกำหนด (Requirement) ดังนี้ (Hambleton & Swaminathan. 1985 : 199 - 200)

1. ความทัดเทียมกัน (Equity) ลอร์ด (Lord. 1980 : 195) กล่าวว่า “ถ้าการเทียบคะแนนของแบบทดสอบฉบับ X และ Y มีความยุติธรรมต่อผู้สอบแล้ว จะต้องไม่มีความแตกต่างที่ใดๆ ระดับความสามารถ (θ) ไม่ว่าจะสอบด้วยแบบทดสอบฉบับ X หรือ Y ก็ตาม” และมีข้อกำหนดดังนี้

- 1.1 แบบทดสอบที่วัดคุณลักษณะ (Trait) หรือความสามารถ (Ability) ต่างกัน จะไม่สามารถเทียบคะแนนกันได้

- 1.2 คะแนนแบบทดสอบที่มีค่าความเชื่อมั่นไม่เท่ากันแล้ว ไม่สามารถเทียบเคียงกันได้

- 1.3 คะแนนของแบบทดสอบที่มีระดับความยากต่างกัน ในการเทียบคะแนนตามแนวคิด ไม่สามารถจะเทียบคะแนนกันได้ เนื่องจากคะแนนจริงของแบบทดสอบจะมีความสัมพันธ์กันแบบไม่เป็นเส้นตรง และทำให้แบบทดสอบจะมีค่าความเชื่อมั่นไม่เท่ากัน 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 $x(y)$ เป็นฟังก์ชันแบบหนึ่งต่อหนึ่งของ y

- 1.5 คะแนนที่อาจผิดพลาดจากแบบทดสอบฉบับ X และ Y ไม่สามารถจะเทียบกันได้ นอกจากว่าแบบทดสอบฉบับ X และ Y คู่ขนานกัน

- 1.6 แบบทดสอบที่มีความเชื่อมั่นที่ดีสามารถเทียบกันได้

2. ความไม่แปรเปลี่ยนไปตามกลุ่ม (Invariance Across Group) คือ คะแนนของกลุ่มต่างๆ ที่ผ่านการเทียบแล้ว จะมีค่าไม่แตกต่างกันไม่ว่าจะเทียบจากกลุ่มตัวอย่างใด

3. ความสมมาตร (Symmetry) คือ ผลของการเทียบคะแนนจากแบบทดสอบฉบับ X ไปสู่ฉบับ Y หรือจากฉบับ Y ไปสู่ฉบับ X จะต้องเหมือนกัน

4. ความเป็นมิติเดียวกันของแบบทดสอบ (Unidimensionality of the Test) คือ แบบทดสอบทั้งสองฉบับจะต้องมีวัดคุณลักษณะเดียวกัน

ส่วนเลวิน (Holland & Rubin. 1982 : 10 ; citing Levin. 1955) ได้เสนอเงื่อนไขเกี่ยวกับการเทียบคะแนนไว้ว่า แบบทดสอบสองฉบับใดๆ ที่จะเทียบคะแนนกันได้ก็ต่อเมื่อแบบทดสอบทั้งสองฉบับนั้นคู่ขนานกัน (Parallel) ในด้านโครงสร้าง (Structure) เวลา (Timing) รูปแบบ (Format) ชนิดของข้อสอบ (Item Type) และเนื้อหาวิชา (Subject Matter)

ลอร์ด (Lord. 1980) ได้กำหนดเงื่อนไขของการปรับเทียบคะแนนว่า แบบทดสอบสองฉบับจะสามารถนำมาเปรียบเทียบกันได้ก็ต่อเมื่อแบบทดสอบทั้งสองมีคุณสมบัติ 4 ประการคือ

1. แบบทดสอบทั้งสองฉบับจะต้องวัดความสามารถเดียวกัน (Same Ability) คือ แบบทดสอบทั้งสองวัดคุณลักษณะเดียวกัน คุณลักษณะนี้อาจเป็นคุณลักษณะแฝง หรือความสามารถ หรือทักษะ อย่างใดอย่างหนึ่งก็ได้
2. มีความเสมอภาค (Equity) คือ เมื่อทุกกลุ่มมีความสามารถเดียวกัน การแจกแจงคะแนนของแบบทดสอบฉบับ Y หลังจากมีการแปลงคะแนนแล้วจะมีการแจกแจงเหมือนกับการแจกแจงของคะแนนจากแบบทดสอบฉบับ X
3. ประชากรไม่แปรเปลี่ยน (Population Invariance) คือ การแปลงคะแนนต้องเป็นไปในลักษณะเดียวกันไม่ว่าคะแนนจะมาจากกลุ่มตัวอย่างใดก็ตาม
4. มีความสมมาตร (Symmetry) หมายถึง การแปลงคะแนนสามารถแปรเปลี่ยนกลับได้ เช่น การแปลงคะแนนกลับจากฉบับ X ไปยังฉบับ Y มีผลเช่นเดียวกันกับฉบับ Y ไปสู่ฉบับ X

3. ประเภทของการเทียบคะแนน

สงบ ลักษณะ (2525 : 23) ได้แบ่งประเภทของการเทียบคะแนนออกเป็น 2 กรณี คือ

3.1 การเทียบคะแนนตามแนวนอน (Horizontal Equating) เป็นการเทียบคะแนนระหว่างแบบทดสอบสองฉบับที่มีความยากพอๆ กัน กล่าวคือ ต้องการเทียบคะแนนจากแบบทดสอบฉบับหนึ่งว่าจะเท่ากับกี่หน่วยคะแนนของแบบทดสอบอีกฉบับหนึ่ง ดังนั้นในการทดสอบจะต้องสอบกับบุคคลกลุ่มเดียวกัน

3.2 การเทียบคะแนนตามแนวตั้ง (Vertical Equating) เป็นการเทียบคะแนนระหว่างแบบทดสอบสองชุดที่มีความยากต่างกัน เมื่อนำไปทดสอบกับบุคคลที่มีความสามารถต่างกัน กล่าวคือ ต้องการเทียบคะแนนจากแบบทดสอบสองชุดเมื่อความยากของข้อสอบต่างกัน และเมื่อสอบกับบุคคลที่มีความสามารถต่างกัน การเทียบคะแนนในลักษณะนี้จะกระทำโดยมีแบบสอบร่วมอีกฉบับหนึ่ง

4. แบบแผนการเก็บรวบรวมข้อมูล

ข้อมูลจากการสอบโดยใช้แบบทดสอบหลายฉบับและเก็บรวบรวมข้อมูล โดยใช้รูปแบบต่างกันนั้น ข้อมูลที่ได้จากบางรูปแบบไม่สามารถนำมาเทียบคะแนนได้ เช่น ข้อมูลที่ได้จากการสอบโดยใช้แบบทดสอบสองฉบับ กับผู้สอบที่มีความสามารถแตกต่างกันสองกลุ่ม กับผู้สอบที่มีความแตกต่างกันสองกลุ่มจะนำคะแนนที่ได้จากแบบทดสอบต่างฉบับที่จะนำมาเปรียบเทียบกันได้ ต้องได้มาจากผู้สอบกลุ่มเดียวกัน หรือผู้สอบกลุ่มเท่าเทียมกัน หรือได้จากผู้สอบที่สอบข้อสอบร่วมกันรูปแบบใดรูปแบบหนึ่ง แบบแผนการเก็บรวบรวมข้อมูล ซึ่งสามารถนำคะแนนจากแบบทดสอบต่างฉบับกันมาเปรียบเทียบกันได้ ปีเตอร์เซ็น และคณะสามารถจำแนกได้ดังนี้ (Petersen & others. 1989 : 244-246)

แบบแผนที่ 1 กลุ่มผู้สอบกลุ่มเดียว (Single Group Design)

เป็นแบบแผนที่ง่ายที่สุดในการเก็บรวบรวมข้อมูลโดยใช้กลุ่มตัวอย่างผู้สอบกลุ่มเดียว ทำได้โดยให้ผู้สอบแต่ละคนทำแบบทดสอบทั้งสองฉบับ ผู้สอบจะทำแบบทดสอบฉบับหนึ่งแล้วตามด้วยแบบทดสอบอีกฉบับหนึ่ง การให้ทำแบบทดสอบทั้งสองฉบับ ทำให้ไม่มีปัญหาในเรื่องความยากของแบบทดสอบต่างฉบับ แต่การสอบครั้งหลัง ผู้สอบอาจได้รับผลกระทบจากการสอบครั้งแรก เนื่องจากการเรียนรู้ ความเมื่อยล้า และการฝึกฝน

ตาราง 1 แบบแผนที่ผู้สอบกลุ่มเดียว

กลุ่มตัวอย่าง	แบบทดสอบฉบับที่ 1	แบบทดสอบฉบับที่ 2
P1	*	*

แบบแผนที่ 2 ผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล (Counterbalanced Random Group Design)

ในทางปฏิบัติ ลำดับของการทดสอบก่อนหลัง ย่อมส่งผลต่อการสอบ ดังนั้นแบบแผนที่จึงปรับมาจากแบบแผนแรกโดยจัดให้ผู้สอบเป็นกลุ่มสุ่ม 2 กลุ่ม แต่ละกลุ่มย่อยได้รับการสอบด้วยแบบทดสอบสองฉบับ ในลักษณะการจัดลำดับก่อนหลัง สลับให้เกิดความสมดุล การจัดแบบนี้เพื่อให้การสอบโดยใช้แบบทดสอบสองฉบับได้รับอิทธิพลจากองค์ประกอบในเรื่องการเรียนรู้ ความเมื่อยล้า และการฝึกฝนในลักษณะที่สมดุลกัน

ตาราง 2 แบบแผนที่ผู้สอบกลุ่มเดียวที่ได้รับการจัดให้สมดุล

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

แบบแผนที่ 3 ผู้สอบกลุ่มเท่าเทียมกัน (Equivalent Group Design)

ในทางปฏิบัติ การให้ผู้สอบทำแบบทดสอบทั้งสองฉบับ อาจจะทำได้ยาก การใช้ผู้สอบกลุ่มเท่าเทียมกันเป็นรูปแบบที่สามารถนำมาใช้ได้ แบบแผนที่จำเป็นต้องจัดกลุ่มผู้สอบให้มีความคล้ายคลึงกันมากเท่าที่จะทำได้ ผู้สอบในแต่ละกลุ่มจะทำแบบทดสอบฉบับเดียว ดังนั้นจึงหลีกเลี่ยงปัญหาเรื่องการเรียนรู้ ความเมื่อยล้า และการฝึกฝนจากการสอบครั้งแรกได้ แต่จะไม่มีข้อมูลที่จะนำมาใช้ปรับความแตกต่างของกลุ่มตัวอย่างทั้งสอง วิธีที่จะลดความแตกต่างระหว่างกลุ่มทำได้โดยใช้กลุ่มตัวอย่างขนาดใหญ่

ตาราง 3 แบบแผนที่ผู้สอบกลุ่มเท่าเทียมกัน

กลุ่มตัวอย่าง	แบบทดสอบฉบับที่ 1	แบบทดสอบฉบับที่ 2
P1	*	
P2		*

แบบแผนที่ 4 ผู้สอบกลุ่มสุ่มสองกลุ่มโดยมีแบบทดสอบร่วม (Anchor Test Random Group Design)

การใช้แบบแผนนี้อาจมีข้อจำกัดในเรื่องความแตกต่างระหว่างกลุ่มตัวอย่าง การใช้แบบทดสอบร่วมด้วยทำให้สามารถปรับความแตกต่างระหว่างกลุ่มได้ตามแบบแผนนี้ กลุ่มสุ่มแต่ละกลุ่มจะได้รับการสอบแบบทดสอบกลุ่มละ 1 ฉบับ และทั้งสองกลุ่มจะได้รับการสอบด้วยแบบทดสอบร่วมด้วย ซึ่งอาจจะจัดไว้ในแบบทดสอบเรียกว่าแบบทดสอบร่วมภายใน (Internal Anchor Test) หรือจัดแยกไว้จากแบบทดสอบที่ต้องการเทียบคะแนนก็ได้ เรียกว่า แบบทดสอบร่วมภายนอก (External Anchor Test)

ตาราง 4 แบบแผนที่ผู้สอบกลุ่มสุ่มสองกลุ่มโดยมีแบบทดสอบร่วม

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

แบบแผนที่ 5 กลุ่มผู้สอบที่ไม่เท่าเทียมกันโดยมีแบบทดสอบร่วม (Anchor Test Non-Equivalent Group Design)

ในการสอบบางครั้งจำเป็นต้องใช้แบบทดสอบฉบับใหม่กับผู้สอบกลุ่มใหม่ในโปรแกรมการสอบเดียวกัน การสอบโดยใช้แบบทดสอบต่างฉบับกันและเวลาสอบต่างกันนี้ ผู้สอบไม่ได้มาจากกลุ่ม ประชากรเดียวกัน จึงจำเป็นต้องใช้แบบทดสอบร่วมในการปรับความแตกต่างระหว่างกลุ่ม ตามแบบแผนนี้ผู้สอบในแต่ละกลุ่มจะได้รับแบบทดสอบ 1 ฉบับและแบบทดสอบร่วม ซึ่งต้องมีเนื้อหาและระดับความยากคล้ายกับแบบทดสอบที่ต้องการเทียบคะแนนให้มากที่สุดเท่าที่จะเป็นไปได้ คะแนนจากแบบทดสอบร่วมจะใช้เพื่อลดความลำเอียงในการเทียบคะแนนจากกลุ่มที่ต่างกัน

ตาราง 5 แบบแผนที่ผู้สอบกลุ่มที่ไม่เท่าเทียมกันโดยมีแบบทดสอบร่วม

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

ส่วนลอร์ด (Hambleton, 1985 : 198 ; citing Lord, 1982 : 145) ได้แบ่งรูปแบบในการเก็บรวบรวมข้อมูลเพื่อเทียบคะแนนนั้นออกเป็น 3 รูปแบบ คือ

1. รูปแบบที่ใช้กลุ่มผู้สอบกลุ่มเดียว หรือใช้ผู้สอบร่วมโดยกลุ่มเดียวทำแบบทดสอบทั้งสองชุด การประมาณค่าความยาก จึงไม่เกี่ยวข้องกับความสามารถของของผู้สอบ แต่ลำดับการสอบชุดใดก่อนหลังอาจทำให้เกิดการเรียนรู้ การฝึกฝน และความเมื่อยล้า ซึ่งจะมีผลกระทบต่อคะแนนผลการสอบได้ ดังนั้นเพื่อขจัดความลำเอียงเหล่านี้จึงดัดแปลงรูปแบบนี้โดยผู้สอบเป็นสองกลุ่ม แต่ละกลุ่มทำแบบทดสอบทั้งสองชุดในลักษณะของการจัดลำดับก่อนหลังต่างกันโดยวิธีสุ่ม

2. รูปแบบที่ใช้กลุ่มที่เท่าเทียมกัน รูปแบบนี้กลุ่มผู้สอบที่เท่าเทียมกันสองกลุ่ม ซึ่งอาจได้มาโดยการสุ่ม หรือเป็นกลุ่มที่ได้จากการจับคู่กัน (Matching) ทางด้านความรู้และความสามารถ ทั้งนี้เพื่อให้คะแนนผลการสอบจากแบบทดสอบแต่ละชุดที่จะนำมาเทียบคะแนนนั้นไม่มีผลของความแตกต่างในความสามารถระหว่างกลุ่มผู้สอบ อย่างไรก็ตามกลุ่มผู้สอบทั้งสอง อาจมีความแตกต่างในการแจกแจงความสามารถ ถึงแม้จะมีเพียงเล็กน้อยก็อาจทำให้เกิดความลำเอียงในการเทียบคะแนนได้

3. รูปแบบที่ใช้แบบสอบร่วม รูปแบบนี้ผู้สอบสองกลุ่มที่แตกต่างกัน (อาจเป็นกลุ่มที่ได้มาจากการสุ่มหรือไม่ได้มาจากการสุ่ม) แต่ละกลุ่มทำแบบทดสอบเพียง 1 ชุด ซึ่งแต่ละชุดจะประกอบด้วยข้อสอบร่วมจำนวนหนึ่ง โดยอาจสอบรวมหรือสอบแยกต่างหากจากแบบทดสอบทั้งสองชุดก็ได้ แบบสอบร่วมจะต้องคู่ขนานหรือเป็นตัวแทนแบบทดสอบทั้งสองชุดนั้นให้มากที่สุด แบบสอบร่วมจะกำหนดสเกลร่วมของคะแนนจากแบบทดสอบแต่ละชุดที่สอบทั้งสองกลุ่ม เพื่อนำไปปรับค่าความยากของข้อสอบ หรือค่าความสามารถของผู้สอบจากแบบทดสอบต่างชุดให้อยู่ในสเกลร่วมกัน

ส่วนมาร์โค (Macro, 1981) ได้รวบรวมแบบแผนการเก็บรวบรวมข้อมูลไว้ทั้งหมด 5 วิธี ดังนี้

1. กลุ่มนักเรียนกลุ่มเดียวตอบแบบทดสอบ 2 ชุด คือ ชุดใหม่กับชุดเก่า การทดสอบจะกระทำในวันเดียวหรือ 2 วันติดต่อกัน แต่หลักการที่สำคัญคือ ช่วงระยะห่างของการทำแบบทดสอบ 2 ชุด จะต้องเป็นเวลาสั้นๆ เพื่อมิให้เกิดประสบการณ์แทรกซ้อนที่มีผลกระทบต่อคะแนน ลำดับการทดสอบชุดใดก่อนไม่ใช่วิทยาเพราะถือว่าองค์ประกอบต่างๆ อัน ได้แก่ การเรียนรู้ การฝึกฝน ความล้า มีผลต่อคะแนนน้อยมาก

2. กลุ่มนักเรียน 2 กลุ่ม แต่ละกลุ่มได้รับการทดสอบทั้ง 2 ชุด ในลักษณะการจัดลำดับการสอบก่อน-หลัง สลับให้เกิดความสมดุลกัน แบบแผนที่ 2 นี้ ดัดแปลงมาจากแบบแผนที่ 1 เป็นแบบแผนที่จะรับประกันอิทธิพลขององค์ประกอบการเรียนรู้ การฝึกฝน และความล้า เป็นการขจัดความลำเอียงที่จะเกิดขึ้นกับผลการสอบชุดใดชุดหนึ่ง

3. กลุ่มนักเรียน 2 กลุ่มที่ได้มาโดยการสุ่มแต่ละกลุ่มแล้วสอบแบบทดสอบฉบับใดฉบับหนึ่งเพียงฉบับเดียว สิ่งสำคัญของแบบแผนที่ 3 คือต้องมีกลุ่มสุ่ม 2 กลุ่ม ที่มีความคล้ายคลึงกันด้านความรู้ ความสามารถ หรือทักษะที่ต้องการวัด เพื่อให้คะแนนสอบที่ได้มา ไม่ใช่เป็นผลของความแตกต่างของความสามารถของกลุ่ม

4. กลุ่มนักเรียน 2 กลุ่มที่ได้มาจากการสุ่ม โดยแต่ละกลุ่มจะสอบแบบทดสอบฉบับใดฉบับหนึ่งเพียงฉบับเดียว และทั้งสองกลุ่มจะต้องสอบแบบทดสอบร่วม (Anchor Test) แบบทดสอบร่วมนี้อาจแยกสอบต่างหากจากแบบทดสอบที่ใช้เทียบคะแนน เรียกว่า แบบทดสอบร่วมภายนอก (External Anchor Test) หรืออาจนำมารวมกับแบบทดสอบเทียบคะแนนแต่ละฉบับ เรียกว่า แบบสอบร่วมภายใน (Internal Anchor Test) ส่วนที่เป็นแบบทดสอบร่วมควรประกอบด้วยคำถามที่คล้ายคลึงกับคำถามในแบบทดสอบที่ต้องการเทียบคะแนนทั้งสองชุด ประโยชน์ของแบบทดสอบร่วมที่มีต่อการเทียบคะแนนของแบบทดสอบทั้งสองชุดจะมีมากหรือน้อย ขึ้นอยู่กับความสัมพันธ์ของแบบทดสอบร่วมที่มีต่อแบบทดสอบทั้งสองชุดนั้น เหตุที่แบบแผนที่ 4 ได้เปรียบ เพราะคะแนนชุดเก่าและใหม่สามารถปรับ เพื่อสะท้อนให้เห็นความแตกต่างที่ปรากฏอยู่ในกลุ่มทั้งสองตามผลของแบบทดสอบร่วมได้

5. กลุ่มนักเรียนที่ไม่ได้มาจากการสุ่ม 2 กลุ่ม แต่ละกลุ่มทำแบบทดสอบเพียงชุดเดียว และทำแบบทดสอบร่วมเหมือนกันอีก เช่นเดียวกับแบบแผนที่ 4 กลุ่มผู้สอบที่ไม่ได้มาจากการสุ่ม มักจะเกิดขึ้นกับการสอนตามธรรมชาติ ปัจจัยสำคัญของแบบแผนนี้อยู่ที่แบบทดสอบร่วม ซึ่งจะต้องมีความคล้ายคลึงกันมากที่สุดกับแบบทดสอบ 2 ชุดที่ต้องการเทียบคะแนน โดยปกติแล้วไม่มีวิธีการทางสถิติใดที่จะทำการปรับคะแนนในกรณีกลุ่มตัวอย่างไม่ได้มาจากการสุ่ม แต่จะทำได้ก็ต่อเมื่อมีแบบทดสอบร่วมที่มีความเป็นคู่ขนานกับแบบทดสอบชุดใหม่และเก่า หน้าที่ของแบบทดสอบร่วม คือ กำหนดจุดอ้างอิงให้เกิดขึ้นกับทั้งสองชุดในทำนองเดียวกับการ

กำหนดจุดเดือด หรือจุดเยื้องเชิงของน้ำเพื่อใช้กำหนดมาตราบนเทอร์โมมิเตอร์บนฟาร์เรนไฮต์ให้เท่ากับ องศาบนเทอร์โมมิเตอร์แบบเซลเซียสได้ แบบนี้ใช้กับการเทียบคะแนนในแบบทดสอบ SAT มาแล้ว

แองกอฟ (วิรัช วรรณรัตน์. 2530 : 70 ; อ้างอิงมาจาก Angoff. 1976 : 569) จัดแบบแผนการเก็บรวบรวมข้อมูลไว้ 6 ลักษณะคือ

แบบที่ 1 กลุ่มเดียวทำแบบทดสอบชุดเดียวที่คล้ายกัน กล่าวคือให้กลุ่มหนึ่งสอบแบบทดสอบชุด ก อีกกลุ่มหนึ่งสอบแบบทดสอบชุด ข โดยแบบทดสอบชุด ก และ ข มีลักษณะคู่ขนานกัน แล้วเทียบคะแนนแบบสอบทั้งสองชุดด้วยสมการเชิงเส้นตรง

แบบที่ 2 สองกลุ่มทำแบบทดสอบสองชุด โดยให้กลุ่มหนึ่งสอบแบบทดสอบชุด ก แล้วตามด้วยแบบทดสอบชุด ข อีกกลุ่มหนึ่งสอบด้วยแบบทดสอบชุด ข แล้วตามด้วยแบบทดสอบชุด ก แล้วนำคะแนนมาเทียบกันโดยใช้สมการเส้นตรง

แบบที่ 3 สองกลุ่มทำแบบทดสอบสองชุด โดยมีแบบทดสอบชุดหนึ่งเป็นแบบทดสอบร่วม กล่าวคือให้กลุ่มหนึ่งสอบแบบทดสอบชุด ก และ ค อีกกลุ่มหนึ่งสอบแบบทดสอบชุด ข และ ค แล้วนำคะแนนจากแบบทดสอบชุด ก และชุด ข มาเทียบกัน โดยใช้แบบทดสอบชุด ค เป็นตัวปรับร่วม

แบบที่ 4 สองกลุ่มทำแบบทดสอบสองชุด โดยชุดหนึ่งเป็นแบบทดสอบร่วม ซึ่งเหมือนแบบที่ 3 จะแตกต่างกันที่แบบที่ 4 ไม่สลับกลุ่มที่ใช้ในการให้แบบสอบ

แบบที่ 5 วิธีการอื่นๆ ที่เกี่ยวข้องกับคะแนน โดยแบ่งเป็น

- แบบทดสอบ ก และ ข เทียบคะแนนไปยังแบบทดสอบ ค ที่เป็นแบบทดสอบร่วม
- พยากรณ์แบบทดสอบ ก และ ข โดยแบบทดสอบร่วม

แบบที่ 6 วิธีการเทียบคะแนนภายใต้คะแนนรายข้อ เช่น วิธีการของ Thurstone และวิธีการเทียบคะแนนแบบ Swineford-Fan

5. รูปแบบการเทียบคะแนน

รูปแบบการเทียบคะแนน มีผู้เสนอไว้หลายรูปแบบพอสรุปได้ดังนี้

5.1 การเทียบคะแนนเชิงเส้นตรง

รูปแบบการเทียบคะแนนเชิงเส้นตรง เป็นรูปแบบที่ได้รับความนิยมมาเป็นเวลานาน (ภาวิณี ศรีสุข วัฒนานันท์. 2529 : 35) โดยใช้สมการเส้นตรงเป็นรูปแบบเป็นรูปแบบอยู่ในนิยามที่ว่า คะแนนจากแบบทดสอบ 2 ชุด โดยที่แบบทดสอบทั้ง 2 ชุด ต่างเป็นแบบทดสอบที่วัดในสิ่งเดียวกัน (Parallel Measures) จะพิจารณาตัดสินให้เทียบกันได้ก็ต่อเมื่อคะแนนมาตรฐานทั้งสองชุดมีค่าเท่ากัน คือ

$$Z_y = \frac{Y - M_y}{S_y}$$

$$Z_x = \frac{X - M_x}{S_x}$$

$$Z_y = Z_x$$

$$\frac{Y - M_y}{S_y} = \frac{X - M_x}{S_x}$$

เมื่อ	X	คือ	คะแนนจากแบบทดสอบฉบับ X
	Y	คือ	คะแนนจากแบบทดสอบฉบับ Y
	M_x	คือ	ค่าเฉลี่ยของแบบทดสอบฉบับ X
	M_y	คือ	ค่าเฉลี่ยของแบบทดสอบฉบับ Y
	S_x	คือ	ค่าเบี่ยงเบนมาตรฐานของแบบทดสอบฉบับ X
	S_y	คือ	ค่าเบี่ยงเบนมาตรฐานของแบบทดสอบฉบับ Y
	Z_x	คือ	คะแนนมาตรฐานของแบบทดสอบฉบับ X
	Z_y	คือ	คะแนนมาตรฐานของแบบทดสอบฉบับ Y

ในกรณีรูปแบบการแจกแจงของคะแนนแบบทดสอบฉบับ X เหมือนกับแบบทดสอบฉบับ Y ใช้การเทียบรูปแบบเชิงเส้นตรงได้สะดวกง่าย และมีความเป็นปรนัย เพราะจะเป็นไปตามข้อตกลงเบื้องต้น

การเทียบคะแนนแบบเชิงเส้นตรงโดยใช้แบบสอบร่วม

บรูเนอและฮอลแลนด์ (Braun & Holland. 1982 : 23) เป็นผู้คิดค้นรูปแบบเชิงเส้นตรงโดยใช้แบบทดสอบร่วมขึ้น วิธีนี้เริ่มด้วยการมีข้อตกลงเบื้องต้นของการเทียบคะแนนมีฟังก์ชันเป็นเส้นตรงดังนี้

$$e_x(Y) = \mu_x + \frac{\sigma_x}{\sigma_y}(Y - \mu_y)$$

เมื่อ μ_x , σ_x , μ_y และ σ_y เป็นค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนจากแบบทดสอบฉบับ X และฉบับ Y ตามลำดับ และในการหาค่า μ_x , σ_x , μ_y และ σ_y จะหาได้ก็ต่อเมื่อได้กำหนดเงื่อนไขเป็นข้อตกลงเบื้องต้น 2 ประการ คือค่าเฉลี่ยและความแปรปรวนอย่างมีเงื่อนไขของ x เมื่อกำหนดแบบทดสอบร่วม (V) เดียวกับ y และเป็นข้อตกลงตามรูปแบบ (Model Assumption) เพราะเป็นการยอมรับการทำค่าคาดหวังของค่าเฉลี่ยและค่าความแปรปรวน

ผลจากข้อตกลงเบื้องต้น สรุปหลักการเทียบคะแนนได้ดังนี้ ถ้าค่าคาดหวังเฉลี่ย และค่าความแปรปรวนในแบบทดสอบร่วมเท่ากันแล้ว จะได้ค่า μ_x , σ_x , μ_y และ σ_y เพื่อจะนำไปแทนค่าในสมการข้างบน จะได้ฟังก์ชันของการเทียบคะแนนเชิงเส้นตรงของ x ไปสู่ y (Braun & Holland. 1982 : 23 - 24)

ขั้นตอนของการเทียบคะแนนตามรูปแบบเชิงเส้นตรง

รวบรวมข้อมูลจากกลุ่มตัวอย่าง ที่ให้กลุ่มตัวอย่างที่ 1 ทำแบบทดสอบฉบับ X และกลุ่มตัวอย่างที่ 2 ทำแบบทดสอบฉบับ Y นำมาประกอบค่าประชากรต่างๆ เพื่อนำมาแทนค่าในสมการ

$$X^* = e_x(Y) = A_y + B$$

กระทำด้วย วิธีการประมาณด้วยความเป็นไปได้สูงสุด (Maximum Likelihood Method) (Angoff. 1984 : 105) ตามขั้นตอนดังนี้

1) หาค่าประมาณต่างๆ ตามสูตรดังนี้

$$\hat{\mu}_x = M_{x1} + b_{xv1}(\hat{\mu}_v - M_{v1})$$

$$\hat{\sigma}_x^2 = S_{x1}^2 + b_{xv1}^2(\hat{\sigma}_v^2 - S_{v1}^2)$$

$$\hat{\mu}_y = M_{y2} + b_{yv2}(\hat{\mu}_v^2 - M_{v2})$$

$$\hat{\sigma}_y^2 = S_{y2}^2 + b_{yv2}^2(\hat{\sigma}_v^2 - S_{v2}^2)$$

เมื่อ	$\hat{\mu}_x, \hat{\sigma}_x^2$	คือ	ค่าประมาณคะแนนเฉลี่ย และความแปรปรวนของกลุ่มตัวอย่างทั้งหมดที่ทำแบบทดสอบฉบับ X
	$\hat{\mu}_y, \hat{\sigma}_y^2$	คือ	ค่าประมาณคะแนนเฉลี่ย และความแปรปรวนของกลุ่มตัวอย่างทั้งหมดที่ทำแบบทดสอบฉบับ Y
	M_{x1}, S_{x1}^2	คือ	ค่าเฉลี่ย และความแปรปรวนของคะแนนแบบทดสอบฉบับ X ในกลุ่มที่ 1
	M_{y2}, S_{y2}^2	คือ	ค่าเฉลี่ย และความแปรปรวนของคะแนนแบบทดสอบฉบับ Y ในกลุ่มที่ 2
	M_{v1}, S_{v1}^2	คือ	ค่าเฉลี่ย และความแปรปรวนของคะแนนแบบทดสอบรวมในกลุ่มที่ 1
	M_{v2}, S_{v2}^2	คือ	ค่าเฉลี่ย และความแปรปรวนของคะแนนแบบทดสอบรวมในกลุ่มที่ 2
	b_{xv1}	คือ	สัมประสิทธิ์ถดถอยของคะแนน จากแบบทดสอบฉบับ X จากแบบทดสอบรวมในกลุ่มที่ 1
	b_{yv2}	คือ	สัมประสิทธิ์ถดถอยของคะแนน จากแบบทดสอบฉบับ Y จากแบบทดสอบรวมกลุ่มที่ 2

2) หาค่าประมาณความชัน (A) และค่าคงที่ (B) จากสูตร

$$A = \frac{b_{yv2}[S_{x1}^2 + b_{xv1}^2(\hat{\sigma}_v^2 - S_{v1}^2)]}{b_{xv1}[S_{y2}^2 + b_{yv2}^2(\hat{\sigma}_v^2 - S_{v2}^2)]}$$

$$B = \hat{\mu}_x - A\hat{\mu}_y$$

3) แทนค่าประมาณประชากรลงในสมการเส้นตรง

$$X^* = A_y + B$$

4) ใช้สมการสร้างตารางการเทียบคะแนนจากแบบทดสอบฉบับ X ไปสู่คะแนนของแบบทดสอบฉบับ Y

5.2 วิธีการเทียบคะแนนรูปแบบอิกวิเปอร์เซ็นไทล์

แองกอฟ (Angoff, 1984 : 101) กล่าวว่า รูปแบบอิกวิเปอร์เซ็นไทล์เริ่มจากการแจกแจงของคะแนนจากแบบทดสอบฉบับที่ 1 และฉบับที่ 2 มีลักษณะคล้ายกัน หรือจะแตกต่างกันบ้างก็เพียงเล็กน้อย การเทียบคะแนนทำได้โดยใช้คะแนน ณ ตำแหน่งเปอร์เซ็นไทล์เดียวกันของคะแนน 2 ชุดนั้น ผลของการเทียบคะแนนทำด้วยกราฟ ขั้นตอนการแปลงคะแนนคือ เลือกกลุ่มตัวอย่างที่มีความสามารถกระจาย คือ มีทั้งเก่ง ปานกลาง และอ่อน แล้วแบ่งกลุ่มตัวอย่างออกเป็นสองกลุ่ม ให้กลุ่มหนึ่งทำแบบทดสอบฉบับที่ 1 และกลุ่มที่ 2 ทำแบบทดสอบฉบับที่ 2 ผลการทดสอบนำมาแจกแจงเป็นคะแนนชุดที่ 1 และชุดที่ 2 คำนวณหาจุดกลางเปอร์เซ็นไทล์ของการแจกแจง อ่านและทำเครื่องหมายสำหรับค่าคะแนนชุดที่ 1 และชุดที่ 2 ของคะแนนการแจกแจงที่เทียบกันบนกระดาษกราฟ ทำจุดประมาณ 30 จุด แล้วลากเส้นเชื่อมเป็นเส้นกราฟ ทำการปรับให้เรียบ เส้นกราฟนี้จะใช้อ่านค่าแบบทดสอบฉบับที่ 1 ที่สมนัยกับฉบับที่ 2 จากนั้นทำตารางแปลงคะแนนสำเร็จ โดยปกติการเทียบคะแนนรูปแบบอิกวิเปอร์เซ็นไทล์ให้ภาพการแปลงคะแนนที่สะท้อนถึงระดับความยากง่ายของแบบทดสอบสองฉบับ ถ้าแบบทดสอบสองฉบับมีค่าความยากใกล้เคียงกัน เส้นกราฟจะมีลักษณะใกล้เคียงเส้นตรง แต่ถ้ามีความยากต่างกัน เส้นกราฟจะเป็นเส้นโค้ง คะแนนสมมูลที่เกิดขึ้นจะถูกยึดหยุ่นคะแนนดิบให้คงรักษาไว้ให้เหมือนชุดก่อนตามต้องการ

การเทียบคะแนนตามรูปแบบอิกวิเปอร์เซ็นไทล์โดยใช้แบบทดสอบร่วม

บรูเนและฮอลแลนด์ (Braun & Holland, 1982 : 19-22) ได้แสดงถึงความสัมพันธ์ของแบบทดสอบร่วมในการเทียบคะแนนตามรูปแบบอิกวิเปอร์เซ็นไทล์ คือ กลุ่มตัวอย่างสองกลุ่มที่สุ่มมาอย่างอิสระจากประชากรทั้งหมด (T) โดลกกลุ่มตัวอย่างที่ 1 คือ Q เขียนเป็นสมการได้

$$T = fp + (1 - f) Q$$

โดยที่ p และ Q ขึ้นอยู่กับน้ำหนักของ f และ (1 - f) ตามลำดับ การคำนวณการแจกแจงของตัวอย่างร่วมทำได้โดยการคำนวณการแจกแจงตามเงื่อนไขของกลุ่มตัวอย่างที่ 1 และกลุ่มที่ 2 ก่อนแล้วหาค่าเฉลี่ยที่ถ่วงน้ำหนักด้วย f และ 1 - f ซึ่งได้จากสมการ คือ

$$f = N_1 / (N_1 + N_2)$$

เมื่อ N_1 และ N_2 คือ ขนาดของกลุ่มตัวอย่างที่ 1 และกลุ่มที่ 2 ตามลำดับ ค่าของ f คือ ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1 และกลุ่มตัวอย่างที่ 2

ในรูปแบบการเทียบคะแนนแบบอิกวิเปอร์เซ็นไทล์มีข้อตกลงเพื่อหาค่าประมาณการแจกแจงของแบบทดสอบฉบับที่ 1 และแบบทดสอบฉบับที่ 2 ในประชากรทั้งหมดซึ่งได้ฟังก์ชันของการเทียบคะแนนเป็น

$$e_x(Y) = F_T(G_T(Y))$$

ขั้นตอนการเทียบคะแนนตามรูปแบบอิกวิเปอร์เซ็นไทล์ มีขั้นตอนดังนี้

- 1) ทำตารางประมาณการแจกแจงความถี่ของกลุ่มตัวอย่างทั้งสองกลุ่ม
- 2) นำค่าความถี่ประมาณของกลุ่มตัวอย่างทั้งสองกลุ่มหาค่าตำแหน่งเปอร์เซ็นไทล์ของแต่ละขั้นคะแนน

3) เขียนกราฟแสดงความสัมพันธ์ของคะแนนชุดที่ 1 และชุดที่ 2 เพื่อแสดงคะแนนที่สมมูล โดยการลากเส้นกราฟให้ผ่านจุดต่างๆ ให้มากที่สุดแล้วเกลากลั่นโค้งให้เรียบ

4) ทำตารางคะแนนสมมูลของคะแนนชุดที่ 1 และชุดที่ 2 โดยอ่านจากกราฟ

5.3 วิธีการเทียบคะแนนตามรูปแบบอิงทฤษฎีการตอบข้อสอบ

ลอร์ด (Lord. 1977 : 195) ได้กล่าวถึงการเทียบคะแนนรูปแบบอิงทฤษฎีการตอบข้อสอบว่า การเทียบคะแนนของแบบทดสอบ 2 ชุดใดๆ ที่ใช้ได้กับทุกคนจะต้องเป็นไปตามความแตกต่างของบุคคลที่ระดับความสามารถ (θ) เดียวกัน และได้เสนอข้อกำหนดที่สำคัญของการเทียบคะแนนไว้ 3 อย่าง คือ

1) ความทัดเทียมกัน (Equity) หมายถึง ทุกๆ ความสามารถ θ การแจกแจงความถี่แบบมีเงื่อนไขของคะแนนแปลง $X(Y)$ หรือ Y ที่ θ ที่กำหนดให้ต้องเหมือนกับการแจกแจงความถี่แบบมีเงื่อนไขของคะแนน X

2) ความไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง (Invariance Across Groups) หมายถึง คะแนนแปลง $X(Y)$ จะคงที่ โดยไม่ขึ้นกับตัวแปรด้านอื่นๆ ของประชากรที่นำมาเทียบคะแนน

3) ความสมมาตร (Symmetry) หมายถึง การเทียบคะแนนจะต้องเหมือนกันไม่ว่าการเทียบนั้นจะเป็นการเทียบจากแบบทดสอบฉบับ X ไปยังแบบทดสอบฉบับ Y หรือแบบทดสอบฉบับ Y ไปยังแบบทดสอบฉบับ X ก็ตาม

ซึ่งข้อกำหนดนี้ จะมีในการเทียบคะแนนที่สอดคล้องกับข้อมูลที่เหมาะสมกับทฤษฎีการตอบข้อสอบด้วย และคะแนนจากการสอบนั้นไม่มีค่าความคลาดเคลื่อนรวมอยู่นั้นคือ มีเฉพาะคะแนนจริง

ในการคำนวณจะต้องหาค่าประมาณของประชากรในแต่ละชุดของแบบทดสอบแล้วนำมาเทียบเป็นคะแนนที่สมมูลในค่าความสามารถ (θ) เดียวกัน และในทางปฏิบัติเราไม่สามารถหาคะแนนจริงได้ ซึ่งลอร์ด (Lord. 1982 : 164-174) ได้เสนอวิธีการเทียบโดยใช้คะแนนจากการสังเกต (Observed Score) ได้

5.3.1 การเทียบคะแนนรูปแบบอิงทฤษฎีการตอบข้อสอบโดยใช้แบบทดสอบร่วม

การออกแบบเพื่อศึกษาการเทียบคะแนนโดยใช้ทฤษฎีการตอบข้อสอบ ได้ใช้แบบทดสอบร่วมช่วยพิจารณาและปรับความแตกต่างของความสามารถ 2 กลุ่มโดยแต่ละกลุ่มจะสอบเพียงฉบับเดียว และแบบทดสอบร่วมเท่านั้น วิธีการเทียบคะแนนรูปแบบนี้ มีการเทียบคะแนน 2 วิธีกล่าวคือ

1. วิธีการเทียบคะแนนโดยใช้คะแนนจริง (True-Score Equating)

โดยใช้ความสัมพันธ์เชิงคณิตศาสตร์ระหว่างความสามารถและจำนวนคะแนนตามข้อที่ทำถูก ซึ่งจะได้คะแนนจริงจากแบบทดสอบ 2 ฉบับ ดังนี้

$$\xi = \xi(\theta) = \sum_{i=1}^n P_i(\theta)$$

$$\eta = \eta(\theta) = \sum_{j=1}^m P_j(\theta)$$

เมื่อ	ξ	คือ	คะแนนจริงจากแบบทดสอบฉบับ X
	η	คือ	คะแนนจริงจากแบบทดสอบฉบับ Y
	n	คือ	จำนวนข้อของแบบทดสอบฉบับ X
	m	คือ	จำนวนข้อของแบบทดสอบฉบับ Y

ถ้าแทนค่าความสามารถ (θ) ลงในสมการทั้งสอง ก็จะได้คะแนนจริงที่สมมูลกันบนมาตราเดียวกัน และสามารถแทนค่าความสามารถ (θ) ลงในสมการที่ θ เท่ากัน จะหาคะแนนสมมูลของแบบทดสอบฉบับ X และฉบับ Y ได้โดยคำนวณจากค่า $P_i(\theta)$ และ $P_j(\theta)$ โดยประมาณค่า $P_i(\theta)$ และ $P_j(\theta)$ ที่ได้จากค่าพารามิเตอร์ (Lord, 1977 : 180) โดยการใช้โปรแกรมโลจิส (LOGIST) วิเคราะห์ข้อมูลการตอบโดยมีข้อตกลงเบื้องต้นว่า ความสามารถ (θ) ใดๆ ของผู้สอบไม่มีความแตกต่างจากแบบทดสอบฉบับหนึ่งไปยังอีกแบบทดสอบหนึ่ง และถือว่าส่วนที่ไม่ได้สอบคือ ส่วนที่ทำไม่ถึง (Lord, 1980 : 20)

2. วิธีการเทียบมาตราโดยใช้คะแนนสังเกต (Observed-Score Equating)

จากการที่กะประมาณคะแนนจริงในทางปฏิบัติมีปัญหา ซึ่งลอร์ด (Lord, 1980 : 197) ได้กล่าวว่า การกะประมาณคะแนนของผู้สอบจากคำตอบ $\hat{\xi} = \sum_{i=1}^n P_i(\theta)$ ที่แท้จริง คือ คะแนนที่มีค่าความคลาดเคลื่อนรวมอยู่ด้วย ไม่มีคุณสมบัติเป็นคะแนนจริง ลอร์ดจึงเสนอการเทียบคะแนนด้วยคะแนนสังเกต เป็นวิธีการเทียบคะแนน โดยค่าประมาณที่ทำได้ในทางปฏิบัติที่สถานการณ์การสอบไม่สามารถจัดให้ผู้สอบทำแบบทดสอบได้ทั้ง 2 ชุด โดยการกะประมาณค่าความสามารถ (θ) ของผู้สอบ ใช้การแจกแจงของ r ในกลุ่มที่ประมาณ $r(\theta)$ คือ กลุ่มค่าความยาก ค่าอำนาจจำแนก และค่าสัมประสิทธิ์การเดา การแจกแจงของคะแนนดิบที่ได้จากสมการดังนี้

$$\hat{\phi}(x, y) = \frac{1}{N} \sum_{N=1}^N \hat{\phi}_x(X/\hat{\phi}_a) \hat{\phi}_y(Y/\hat{\phi}_a)$$

$$\hat{\phi}(x, y) = \int_{-\infty}^{\infty} \hat{\phi}_x(X/\theta) \hat{\phi}_y(y/\theta) \hat{y}(\theta) d\theta$$

จากสมการเป็นการแจกแจงร่วมกันของตัวแปร 3 ตัว คือ θ , X และ Y จากที่ θ เป็นตัวระบุคะแนนจริง ξ และ η การแจกแจงนี้ จึงเป็นการแจกแจงร่วมกันของตัวแปร 4 ตัว คือ ξ , η , X , และ Y ซึ่งกะประมาณได้โดยการระบุความสัมพันธ์เชิงอิควิเปอร์เซ็นไทล์ระหว่าง X และ Y

จากทฤษฎีการเทียบคะแนนดิบของแบบทดสอบฉบับเดียว ด้วยวิธีการเทียบคะแนนจริง วิธีการเทียบคะแนนสังเกตดังกล่าวมานั้น ซึ่งจะมีทั้งข้อดีและข้อเสีย วิธีการเทียบคะแนนจริง ξ และ η ให้ความหมายของคะแนนสมมูลเฉพาะคะแนนที่อยู่เหนือค่าเฉลี่ยการเดา ส่วนวิธีที่ใช้การกะประมาณจากคะแนนสังเกตก็มีจุดอ่อนที่เป็นเพียงการเทียบคะแนนอย่างประมาณเท่านั้น ซึ่งลอร์ดได้กล่าวว่าทั้งสองวิธีมีความสอดคล้องกันมาก แต่การสรุปเปรียบเทียบอ้างอิงทั่วไปจะต้องทำอย่างพิถีพิถัน (Lord, 1980 : 203) ซึ่งในปี ค.ศ.1984 ลอร์ดและวินเกสกี (ภาวิณี ศรีสุขวัฒนานนท์, 2529 : 45 ; อ้างอิงมาจาก Lord & Wingersky, 1984) ได้ศึกษาเปรียบเทียบผลการเทียบคะแนนตามรูปแบบอิงทฤษฎีการตอบข้อสอบด้วยวิธีคะแนนจริงและคะแนนสังเกต ผลปรากฏว่าทั้งสองวิธีไม่มีข้อแตกต่างกันเลย และได้วิจารณ์ไว้ว่า วิธีคะแนนสังเกตมีความซับซ้อนในการคำนวณมากกว่าวิธีคะแนนจริง

ขั้นตอนการเทียบคะแนนตามรูปแบบอิงทฤษฎีการตอบข้อสอบ

1) ขั้นตอนการประมาณคะแนนจริง โดยกะประมาณค่าพารามิเตอร์ของข้อสอบ และค่าความสามารถของข้อสอบรายข้อ ใช้กะประมาณค่าความเป็นไปได้สูงสุด และหาค่าประมาณคะแนนจริงของผู้สอบ

2) ขั้นตอนการเทียบคะแนน นำค่าคะแนนจริงของผู้สอบมาจับคู่ ที่ค่าความสามารถเดียวกัน และปรับขยายคะแนนจริงที่ต่ำกว่าคะแนนการเดาให้มีความหมายขึ้น แล้วจึงนำคะแนนจริงของแบบทดสอบแต่ละชุดที่คะแนนความสามารถเดียวกันในแต่ละคู่มาสร้างเป็นตารางเทียบคะแนน

5.3.2 โมเดลของทฤษฎีการตอบข้อสอบ

เนื่องจากข้อมูลที่ได้มาจากการสอบมีหลายลักษณะ ได้แก่ ข้อมูลแบบทวิ (Dichotomous) ข้อมูลแบบพหุ (Multi-Chotomous) และข้อมูลแบบต่อเนื่อง (Continuous) ดังนั้นจึงมีผู้พัฒนาโมเดลเพื่อให้สอดคล้องกับลักษณะข้อมูลดังกล่าวขึ้นมามากมาย แต่สำหรับข้อมูลที่เป็นแบบทวิ โมเดลที่นิยมใช้ได้แก่ โมเดลลอจิสติก (Logistic Model) ชนิด หนึ่ง สอง และสามพารามิเตอร์ ซึ่งแต่ละโมเดลมีรายละเอียดดังนี้ (Hambleton & Swaminathan. 1985 : 25)

1) รูปแบบที่ใช้พารามิเตอร์ 2 ตัว (Two-Parameters Logistic Model)

โมเดลนี้เสนอโดยเบิร์นบอม (Birnbbaum) เมื่อปี 1957 เป็นโมเดลโค้งคุณลักษณะข้อสอบ และเป็นฟังก์ชันของการแจกแจงที่มี 2 พารามิเตอร์ สามารถเขียนได้ดังนี้

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

เมื่อ $P_i(\theta)$	คือ	ความน่าจะเป็นในการตอบข้อสอบที่ i ถูก เมื่อผู้สอบมีความสามารถ θ
a_i	คือ	ค่าอำนาจจำแนกของข้อสอบข้อที่ i
b_i	คือ	ค่าความยากของข้อสอบข้อที่ i
D	คือ	ค่าปรับสเกล เท่ากับ 1.702
e	คือ	ค่าคงที่ เท่ากับ 2.7183

2) รูปแบบที่ใช้พารามิเตอร์ 3 ตัว (Three-Parameters Logistic Model)

รูปแบบ 3 พารามิเตอร์ สามารถเขียนได้จากรูปแบบ 2 พารามิเตอร์ โดยเพิ่มพารามิเตอร์ตัวที่ 3 คือ ค่าการเดา ซึ่งมีสูตรในการคำนวณ ดังนี้

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

เมื่อ $P_i(\theta)$	คือ	ความน่าจะเป็นในการตอบข้อสอบที่ i ถูก เมื่อผู้สอบมีความสามารถ θ
a_i	คือ	ค่าอำนาจจำแนกของข้อสอบข้อที่ i
b_i	คือ	ค่าความยากของข้อสอบข้อที่ i
D	คือ	ค่าปรับสเกล เท่ากับ 1.702
e	คือ	ค่าคงที่ เท่ากับ 2.7183
C_i	คือ	ค่าการเดาของข้อสอบข้อที่ i ทั้งนี้รูปแบบนี้ไม่สามารถประมาณค่าพารามิเตอร์ที่อยู่ต่ำกว่าการเดาได้

3) รูปแบบที่ใช้พารามิเตอร์ 1 ตัว (One-Parameter Logistic Model)

ซึ่งราสช์ (Rasch) ได้พัฒนาขึ้นมาโดยใช้รูปแบบนี้เป็นกรณีพิเศษของรูปแบบ 2 พารามิเตอร์ที่เบิร์นบอม (Birnbaum) ได้พัฒนาขึ้น ข้อตกลงเบื้องต้นที่เพิ่มขึ้นคือ ค่าอำนาจจำแนกของข้อสอบ (a) เป็นค่าคงที่ของข้อสอบแต่ละข้อและค่าความยากของข้อสอบแปรเปลี่ยนไป ซึ่งสูตรในการคำนวณมีดังนี้

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

เมื่อ	$P_i(\theta)$	คือ	ความน่าจะเป็นในการตอบข้อสอบที่ i ถูก เมื่อผู้สอบมีความสามารถ θ
	a_i	คือ	ค่าอำนาจจำแนกของข้อสอบข้อที่ i = 1
	b_i	คือ	ค่าความยากของข้อสอบข้อที่ i
	D	คือ	ค่าปรับสเกล เท่ากับ 1.702
	e	คือ	ค่าคงที่ เท่ากับ 2.7183

6. วิธีการแปลงคะแนนตามวิธีการเทียบคะแนนเชิงเส้นตรง

แองกอฟฟ์ (Angoff, 1971 : 569-586) ได้รวบรวมและเสนอรูปแบบการเทียบคะแนน ซึ่งได้บรรยายการออกแบบเพื่อรวบรวมข้อมูลและจัดกระทำในทางสถิติเพื่อระบุการแปลงคะแนนไว้ดังนี้

6.1 รูปแบบที่ 1 (Random Group-One Test Administered to Each Group) รูปแบบนี้ผู้สอบมาจากการสุ่มจากประชากรกลุ่มเดียวกัน แบ่งเป็น 2 กลุ่ม คือ กลุ่มที่ 1 สอบแบบทดสอบฉบับ X กลุ่มที่ 2 สอบแบบทดสอบฉบับ Y

ขั้นแรกคำนวณค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของคะแนนแบบสอบแบบทดสอบฉบับ X โดยกลุ่มที่ 1 และแบบสอบแบบทดสอบฉบับ Y โดยกลุ่มที่ 2 จัดให้คะแนนมาตรฐานของทั้งสองกลุ่มเท่ากัน

$$Z_y = Z_x$$

$$Z_y = \frac{Y - M_y}{S_y}$$

$$Z_x = \frac{X - M_x}{S_x}$$

$$\frac{Y - M_y}{S_y} = \frac{X - M_x}{S_x}$$

เมื่อ	Y	คือ	คะแนนสอบที่ได้จากแบบทดสอบฉบับ Y
	M_y	คือ	ค่าเฉลี่ยของคะแนนแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2
	S_y	คือ	ค่าเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2

X	คือ	คะแนนสอบที่ได้แบบทดสอบฉบับ X
M_x	คือ	ค่าเฉลี่ยของคะแนนแบบทดสอบฉบับ X จากการสอบ โดยกลุ่มที่ 1
S_x	คือ	ค่าเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1

เมื่อจัดเทอมในสมการใหม่ จะได้สมการเชิงเส้นตรง คือ

$$Y = \frac{S_y}{S_x} X + M_y - \frac{S_y}{S_x} M_x$$

ซึ่งจะอยู่ในรูป	Y	=	AX + B
เมื่อ	A	=	S_y / S_x
	B	=	$M_y - AM_x$

6.2 รูปแบบที่ 2 (Random Groups-Both Administered to Each Group, Counterbalanced)

ลักษณะเช่นเดียวกับรูปแบบที่ 1 เลือกกลุ่มตัวอย่างขนาดใหญ่แล้วแบ่งสุ่มออกเป็น 2 กลุ่ม โดยกลุ่มที่ 1 สอบแบบทดสอบฉบับ X แล้วตามด้วยแบบทดสอบฉบับ Y กลุ่มที่ 2 สอบแบบทดสอบฉบับ Y แล้วตามด้วยแบบทดสอบฉบับ X ในการที่จะป้องกันความคลาดเคลื่อนที่จะเกิดขึ้นในการสอบ โดยเฉพาะอย่างยิ่งเมื่อสอบแบบทดสอบฉบับ X และแบบทดสอบฉบับ Y พร้อมกัน ควรจะแยกแบบทดสอบทั้งสอง ฉบับติดกันโดยครึ่งหนึ่งให้แบบทดสอบฉบับ X อยู่ด้านบนและอีกครึ่งหนึ่งให้แบบทดสอบฉบับ Y อยู่ด้านบน และในการสอบควรจะแจกสลับเรียงตามผู้สอบ คือ XY, YX, XY, YX...

สมการที่ใช้ในการเทียบคะแนนรูปแบบที่ 2 คือ

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

เมื่อ	M_x	=	$\frac{1}{2} (M_{x1} + M_{x2} - K_x)$
	M_y	=	$\frac{1}{2} (M_{y1} + M_{y2} - K_y)$
	S_x	=	$\frac{1}{2} (S_{x1}^2 + S_{x2}^2)$
	S_y	=	$\frac{1}{2} (S_{y1}^2 + S_{y2}^2)$

เมื่อสมการนี้จัดลงในสมการเส้นตรงจะได้สมการเชิงเส้นในรูปของ $Y = AX + B$ โดยที่

$$A = \sqrt{\frac{S_{y1}^2 + S_{y2}^2}{S_{x1}^2 + S_{x2}^2}}$$

$$B = \frac{1}{2} (M_{y1} + M_{y2}) - \frac{A}{2} (M_{x1} + M_{x2})$$

เมื่อจัดให้อยู่ในรูปสมการเชิงเส้นตรง คือ

$$Y = \sqrt{\frac{S_{y1}^2}{S_{x1}^2} + \frac{S_{y2}^2}{S_{x2}^2}} X + \frac{1}{2} (M_{y1} + M_{y2}) - \frac{A}{2} (M_{x1} + M_{x2})$$

6.3 รูปแบบที่ 3 (Random Groups-One Test Administered to Each Group,

Common Equating Test Administered to Both Group)

คะแนนสอบที่ได้ส่วนใหญ่ขึ้นกับลักษณะ 2 ประการ คือ ความสามารถรายบุคคลและลักษณะของแบบทดสอบ ในการที่จะเปรียบเทียบผลการสอบของแต่ละบุคคลซึ่งสอบแบบทดสอบที่แตกต่างกันจึงจำเป็นต้องปรับคะแนนเสียก่อน กล่าวคือ การเทียบคะแนนนั้นจะเป็นเพียงการเทียบความแตกต่างของคะแนนอันเป็นผลมาจากความแตกต่างของแต่ละบุคคล หรือของกลุ่ม ถ้ากลุ่มที่ 1 และกลุ่มที่ 2 ไม่ได้สุ่มมาจากประชากรเดียวกัน ความแตกต่างระหว่างกลุ่มอาจเป็นองค์ประกอบที่มีนัยสำคัญมากในการแปรเปลี่ยนค่าของ A และ B ในสมการ และจะเป็นผลทำให้เกิดความลำเอียง (Bias) ในวิธีการเทียบคะแนน แม้ว่ากลุ่มจะถูกเลือกมาโดยวิธีการสุ่ม ความแตกต่างเล็กน้อยระหว่างกลุ่มอาจจะเกิดขึ้นได้ ถ้าไม่ป้องกันก็จะเกิดความลำเอียงในสมการการแปลงคะแนน จะเป็นผลต่อการเปรียบเทียบภายหลังในการที่จะควบคุมให้บังเกิดผล การเทียบมาตรฐานจำเป็นต้องใช้คะแนนฟอร์ม U ซึ่งเป็นแบบทดสอบที่เพิ่มขึ้นหรือแบบทดสอบที่ร่วมกันระหว่างแบบทดสอบฉบับ X และแบบทดสอบฉบับ Y แบบสอบรวมนี้ใช้สำหรับปรับความแตกต่างที่อาจจะพบระหว่างกลุ่มที่ 1 และกลุ่มที่ 2 ในการบริหารแบบทดสอบ กลุ่มที่ 1 จะได้รับแบบทดสอบฉบับ X และแบบทดสอบฉบับ U และกลุ่มที่ 2 จะได้รับแบบทดสอบฉบับ Y และแบบทดสอบฉบับ U

ลอร์ด (Lord, 1955) ได้พัฒนาสมการสำหรับรูปแบบนี้โดยใช้การประมาณค่าความเป็นไปได้สูงสุด (Maximum Likelihood Estimated) ประมาณค่าเฉลี่ยและความแปรปรวนของแบบทดสอบฉบับ X และแบบทดสอบฉบับ Y ดังนี้

$$\begin{aligned}\mu_x &= M_{x1} + b_{xu1} (\mu_u - M_{u1}) \\ \mu_y &= M_{y2} + b_{yu2} (\mu_u - M_{u2}) \\ \sigma_x^2 &= S_{x1}^2 + b_{xu1}^2 (\sigma_u^2 - S_{u1}^2) \\ \sigma_y^2 &= S_{y2}^2 + b_{yu2}^2 (\sigma_u^2 - S_{u2}^2)\end{aligned}$$

เมื่อ

$$\begin{aligned}\mu_u &= M_{ut} \\ \sigma_u^2 &= S_{ut}^2 \\ t &= n_1 + n_2\end{aligned}$$

ซึ่งการประมาณค่าเหล่านี้ใช้กับสมการ $Y = AX + B$

$$\begin{aligned}\text{โดยที่ } A &= \sigma_y / \sigma_x \\ B &= \mu_y - A\mu_x\end{aligned}$$

7. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน

ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนจะพิจารณาจากองค์ประกอบดังต่อไปนี้

7.1 ความลำเอียงของการเทียบคะแนน

ความลำเอียงเป็นองค์ประกอบหนึ่งของความคลาดเคลื่อนซึ่งจะต้องพิจารณากันอย่างรอบคอบ เมื่อต้องการนำผลของการเทียบคะแนนไปใช้ บรูน์และฮอลแลนด์ (Braun & Holland. 1982) ได้เสนอบทวิเคราะห์ไว้ว่า ความลำเอียงทางสถิติ มีความหมาย 2 ประการ คือ

1) ความแตกต่างของค่าเฉลี่ยของค่าตัวประมาณตลอดจนการทำการสุ่มซ้ำ ๆ จากประชากรเดียวกันกับค่าของประชากรที่ถูกประมาณ

2) ความแตกต่างของฟังก์ชันการเทียบคะแนนโดยประมาณ (Estimated Equating Function) ของค่าที่แท้จริงของฟังก์ชันของการเทียบคะแนน (Equating Function)

ความลำเอียงของการเทียบคะแนนนี้มีสาเหตุที่ทำให้เกิดความลำเอียงหลายประการ และที่สำคัญมีอยู่ 2 ประการ คือ

1. ความแปรผันของประชากร (Population Variability) เนื่องจากการเทียบคะแนนไม่ได้จัดกระทำกับประชากรกลุ่มเดียวโดยตลอด ฟังก์ชันของการเทียบคะแนนที่สร้างขึ้นเพื่อใช้กับประชากรหนึ่ง อาจนำไปใช้กับประชากรหนึ่งไม่ได้ เช่น $X(Y)$ เป็นการเทียบคะแนนจากแบบทดสอบฉบับ Y ไปยังแบบทดสอบฉบับ X ในประชากร P และ $Y(Z)$ เป็นการเทียบคะแนนจากแบบทดสอบฉบับ Z ไปยังแบบทดสอบฉบับ Y ในประชากร Q แล้ว จะสรุปว่า ฟังก์ชันเชิงซ้อน $X(Y(Z))$ เป็นการเทียบคะแนนจากแบบทดสอบฉบับ Z ไปยังแบบทดสอบฉบับ X ในประชากร P หรือ Q ย่อมไม่ได้ ในแบบแผนการเทียบคะแนนที่ใช้แบบสอบร่วมได้แก้ไขปัญหานี้ โดยการสร้างประชากรสังเคราะห์ (Synthesis Population) ซึ่งเป็นผลรวมของประชากร P และ Q ตามสัดส่วน แต่ยังไม่รับประกันว่า จะไม่มีความลำเอียงในทุก ๆ ประชากรที่แปรเปลี่ยนไป

2. ความคลาดเคลื่อนของรูปแบบ (Model Error) ความคลาดเคลื่อนนี้ เกิดขึ้นจากการจัดกระทำข้อสมมติฐานที่ว่าด้วยรูปแบบการแจกแจงให้ง่ายขึ้นอย่างผิดพลาด แบ่งออกเป็น 2 พวก คือ

2.1 ความคลาดเคลื่อนของรูปแบบที่ทดสอบได้ (Testable Model Error) เช่น ความเป็นเส้นตรงของฟังก์ชันการถดถอย สามารถทดสอบได้จากข้อมูลจำนวนมาก ซึ่งจะเห็นผลได้ว่า มีความเพียงพอหรือไม่

2.2 ความคลาดเคลื่อนของรูปแบบที่ทดสอบไม่ได้ (Nontestable Model Error) เป็นกรณีที่ไม่สามารถหาข้อมูลมาทดสอบเพื่อพิสูจน์ว่า สมมติฐานที่ว่าไว้เพียงพอหรือไม่

7.2 ความแปรผันเชิงสุ่ม และความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน

วิธีการเทียบคะแนนทุกวิธี ไม่ว่าจะอยู่ในรูปแบบการเทียบคะแนนใด เมื่อกลุ่มตัวอย่างผู้สอบเป็นกลุ่มที่สุ่มมาจากประชากรเดียว หรือหลายประชากร ย่อมมีความผันแปรเชิงสุ่มเกิดขึ้น จึงนิยมใช้เทคนิคการประมาณค่าความคลาดเคลื่อนเชิงสุ่ม (Estimated Sampling Error) กับการเทียบคะแนนต่างๆ ความคลาดเคลื่อนเชิงสุ่มมีสมมติฐานเบื้องต้นว่า กลุ่มตัวอย่างมาจากการสุ่ม และใช้ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน (Standard Error of Equating: SEE) เป็นการวัดความแปรผันประเภทนี้

ความคลาดเคลื่อนมาตรฐานโดยอิงรูปแบบทฤษฎีตอบสนองข้อสอบ ค่าประมาณคะแนนจำนวนข้อที่ตอบถูก (Number-Right Score) ξ ของผู้สอบที่ทำแบบทดสอบฉบับ X และ η ของผู้ที่ทำแบบทดสอบฉบับ Y มีค่าเท่ากับฟังก์ชันคุณลักษณะที่ประเมินที่ระดับความสามารถ θ เดียวกัน ให้ความหมายของการเท่ากันในเชิงคะแนนสมมูล ในทางปฏิบัติจึงนำความสัมพันธ์เชิงฟังก์ชัน ξ และ η มาใช้เทียบคะแนน X และ Y จากคะแนนที่สอบได้จริงของคนทั้งสองกลุ่ม การประมาณเพื่อให้เทียบ η ไปยัง ξ ตามฟังก์ชัน จะต้องใช้ค่าประมาณของประชากรข้อสอบ ค่าประมาณเหล่านี้ คือ ที่มาของความคลาดเคลื่อนเชิงสุ่ม (Sampling Error) ในการเทียบมาตรฐาน (Lord, 1981)

ลอร์ด (Angoff, 1984: 96 ; citing by Lord, 1950) ได้สร้างสูตรสำหรับคำนวณความคลาดเคลื่อนมาตรฐานของรูปแบบเชิงเส้นตรง โดยใช้สูตรการประมาณความคลาดเคลื่อนดังนี้

$$SEE = 2\hat{\sigma}_x^2(1-\hat{r}^2) \frac{(1+\hat{r}^2)Z_y^2 + 2}{N_t}$$

เมื่อ	SEE	คือ	ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน
	$\hat{\sigma}_x^2$	คือ	ค่าประมาณความแปรปรวนของคะแนนแบบทดสอบฉบับ X ที่แปลงแล้ว
	$\hat{r}$	คือ	ค่าประมาณของความสัมพันธ์ระหว่างแบบทดสอบฉบับ X และฉบับ Y
	Z_y	คือ	คะแนนมาตรฐานของแบบทดสอบฉบับ Y หาได้จาก $\frac{Y - M_y}{S_y}$

8. ความเพียงพอของการเทียบคะแนน

วิธีการประเมินความเพียงพอมีผู้เสนอแนวคิดและวิธีการปฏิบัติดังนี้

8.1 ดัชนีตรวจสอบความเพียงพอของเจเกอร์ (Jaeger)

เจเกอร์ (Jaeger, 1981 : 26) ได้เสนอดัชนี 5 ตัว เพื่อตรวจสอบความเพียงพอของการใช้รูปแบบเชิงเส้นตรงว่า เทคนิควิธีที่ได้นำมาใช้ในการเทียบคะแนนเพียงพอกับการปรับความแตกต่างระหว่างการแจกแจงของคะแนนจากแบบทดสอบหรือไม่ หรือจำเป็นต้องมีรูปแบบอื่นที่มีความสลับซับซ้อนมากขึ้น ดัชนี 5 ตัว มีดังนี้

1) ดัชนีความคล้ายคลึงของการแจกแจงคะแนนสะสมของแบบทดสอบฉบับเก่าและฉบับใหม่ โดยปรับความแตกต่างระหว่างค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน การทดสอบความเหมือนของการแจกแจง โดยใช้แบบทดสอบด้วย The Kolmogorov-Smirnov Two Sample Test (Jaeger. 1981 : 27 ; citing by Sminov. 1948)

2) รูปแบบของการแจกแจงคะแนนดิบกับคะแนนแปลง (Shape of the Raw Score Transformation) ดัชนีตัวนี้บอกความเพียงพอของการเทียบคะแนนด้วยรูปแบบเชิงเส้น ซึ่งสามารถอธิบายความแตกต่างในการแจกแจงคะแนนดิบทั้งสองชุดในการเทียบคะแนนด้วยรูปแบบเชิงเส้น ถ้าสามารถอธิบายความแตกต่างในการแจกแจงคะแนนดิบทั้งสองชุดได้อย่างเพียงพอ ก็สามารถยอมรับได้ว่าการแปลงคะแนนดิบจากแบบทดสอบฉบับใหม่ไปยังแบบทดสอบฉบับเก่าเป็นเส้นตรงอย่างแน่นอน

3) ความคงเส้นคงวาของผลลัพธ์ของการเทียบคะแนนตามรูปแบบเชิงเส้นตรงกับการเทียบที่ตำแหน่งเปอร์เซ็นต์ไทล์ (Consistency of Linear and Equipercentile Equating Results) การวิเคราะห์นี้อาศัยข้อตกลงที่เป็นสมมติฐานเบื้องต้นว่า ถ้ารูปแบบเชิงเส้นตรงมีความเพียงพอแล้ว ฟังก์ชันของรูปแบบการเทียบคะแนนที่ตำแหน่งเปอร์เซ็นต์ไทล์จะแปรผันไปโดยสุ่มรอบๆ ฟังก์ชันของรูปแบบเชิงเส้นตรงที่สมนัยกัน

4) ความคล้ายคลึงของการแจกแจงความยากของข้อสอบ (Similarity of Item Difficult Distributions) โดยอาศัยหลักการที่ว่า การใช้รูปแบบเชิงเส้นตรงมีความเพียงพออย่างแท้จริงกับแบบทดสอบที่มีคุณสมบัติเป็นคู่ขนานกัน ถ้ามีความเบี่ยงเบนจากความเป็นคู่ขนานมากเท่าใด แสดงว่า ต้องการรูปแบบการเทียบคะแนนที่ซับซ้อนขึ้น เพราะการแจกแจงของแบบทดสอบที่ไม่ใช่คู่ขนานจะมีความแตกต่างเกิดขึ้นในระดับโมเมนต์ที่สูงขึ้น

5) ความคล้ายคลึงของค่าอำนาจจำแนกของข้อสอบ (Similarity of Item Discrimination Distributions) มีเหตุผลทำนองเดียวกับดัชนีตัวที่ 4

8.2 ดัชนีความแตกต่าง (Discrepancy Indices)

ปีเตอร์เซ็น มาร์โค และสจวร์ต (Petersen, Macro & Stewart. 1982) ได้เสนอวิธีประเมินความเพียงพอของรูปแบบการเทียบคะแนนโดยหาความแตกต่างระหว่างคะแนนแปลง (An Estimated Criterion Score : t') ซึ่งเป็นผลมาจากการเทียบคะแนน กับคะแนนเกณฑ์ (Criterion Score : t) ที่สมนัยกัน ถ้าความแตกต่างมีค่าน้อย หมายความว่า ความคลาดเคลื่อนที่เกิดขึ้นจากการใช้รูปแบบการเทียบคะแนนนั้นน้อยด้วย รูปแบบการเทียบคะแนนดังกล่าวย่อมมีความเหมาะสมกับสถานการณ์ที่ใช้

ดัชนีความแตกต่างที่นำมาเปรียบเทียบระหว่างรูปแบบและสถานการณ์ที่ต่างๆ กัน คือ กำลังสองของค่าเฉลี่ยของความแตกต่างที่ถ่วงน้ำหนักด้วยความแปรปรวนของคะแนนเกณฑ์ (The Weighed Mean-Square Difference) เป็นค่ามาตรฐาน ค่าที่คำนวณออกมานี้ เรียกว่า ความแปรปรวนรวม (The Total Error) ซึ่งมีสูตรดังนี้

$$C = \Sigma(X_i - X_j)^2 / nS_x^2$$

เมื่อ	C	คือ	ค่าดัชนีความแตกต่างระหว่างรูปแบบการเทียบคะแนน
	X_i	คือ	คะแนนจากแบบทดสอบฉบับ X ของผู้เข้าสอบคนที่ i
			ของกลุ่มทวนผล

X_i	คือ	คะแนนจากแบบทดสอบฉบับ X ที่ได้จากการนำคะแนนจากแบบทดสอบฉบับ Y ไปเทียบจากตารางการเทียบคะแนนที่วิเคราะห์ได้คนที่ i ของกลุ่มทวนผล
n	คือ	จำนวนคนของกลุ่มสอบทวนผล
S_x^2	คือ	ค่าความแปรปรวนของการแจกแจงคะแนนจากแบบทดสอบฉบับ X ของกลุ่มทวนผล

8.3 ดัชนีเปรียบเทียบเปอร์เซ็นต์ไทล์ (The Percentile Comparison Index) เป็นมาตรฐานการวัดความไม่สอดคล้องระหว่างการแจกแจงของคะแนนในแบบทดสอบฉบับ X กับแบบทดสอบฉบับ Y ที่ได้แปลงไปอยู่ในมาตราคะแนนของ X แล้ว ตามวิธีที่ได้พัฒนาขึ้น ดัชนีเปรียบเทียบเปอร์เซ็นต์ไทล์ คือ ค่าเฉลี่ยกำลังสองของความแตกต่าง (The Mean-Square Difference) ที่ได้จากการแจกแจงคะแนนต่างๆ ของเกณฑ์ X กับคะแนนแปลง X ที่แปลงมาจาก Y ด้วยวิธีเทียบคะแนนที่ระบุไว้ ณ ตำแหน่งเปอร์เซ็นต์ไทล์เดียวกัน ดัชนีนี้เสนอโดย โคลเลน (Kolen. 1982) ได้แนะนำให้ใช้ข้อมูลจากกลุ่มตัวอย่างสอบทวนผล ซึ่งสุ่มมาจากประชากรเดียวกันกับกลุ่มตัวอย่างที่ได้พัฒนาตารางคะแนนแปลง สูตรการคำนวณ คือ

$$C = \sum (X_i - X)^2 / nk$$

เมื่อ n คือ จำนวนคะแนนดิบของกลุ่มสอบทวนผล
 k คือ จำนวนข้อสอบในแบบสอบรวมที่ใช้

ค่า C ที่ได้ ถ้ามีค่าน้อยจะให้ความหมายว่า รูปแบบการเทียบคะแนนที่นำมาสร้างตารางแปลงนั้น มีความเหมาะสมและเพียงพอที่จะให้ผลการแปลงคะแนนอย่างคงเส้นคงวา

9. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการเทียบคะแนนมีทั้งงานวิจัยต่างประเทศ และงานวิจัยในประเทศ สำหรับงานวิจัยต่างประเทศ มีผู้ทำการวิจัยหลายคนดังนี้

โคลเลน (Kolen. 1981 : 1-11) ได้ศึกษารูปแบบของการเทียบคะแนนระหว่างรูปแบบดั้งเดิมสองวิธี คือ รูปแบบเชิงเส้นตรง และรูปแบบอิกวอิเพอร์เซ็นต์ไทล์ กับรูปแบบการเทียบคะแนนโดยใช้ไค์ลักษณะข้อสอบ ที่ศึกษาทั้ง 1, 2 และ 3 พารามิเตอร์ แต่ละวิธีเทียบด้วยค่าประมาณจากคะแนนจริง (Estimated True-Score Equating) ใช้ข้อมูลจากโครงการ IOWA Test Educational Development (ITED) ปี ค.ศ. 1978 ศึกษา 2 แบบ คือ เทียบจากฟอร์มที่มีความยากเท่าเทียมกัน และเทียบฟอร์มที่มีความยากต่างกัน กลุ่มตัวอย่างคือนักเรียนระดับ 1 คือ เกรด 9 และ 10 ระดับ 2 คือ เกรด 11 และ 12 จำนวนทั้งหมด 10,728 คน ผลการวิจัยพบว่า การเทียบคะแนนทั้งหมดแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05 และพบว่า วิธีการเทียบโดยใช้ไค์ลักษณะของข้อสอบแบบสามพารามิเตอร์ ให้ผลอย่างแม่นยำมากที่สุด ส่วนวิธีที่ให้ผลไม่แน่นอนที่สุดคือ วิธีรูปแบบเชิงเส้นตรง

ปีเตอร์เซน, มาร์โค และสจวร์ต (Petersen, Marco & Stewart. 1982 : 71-185) ได้ศึกษาความเพียงพอของการเทียบคะแนนเชิงเส้นตรงหลายๆ วิธี และวิธีอิกวิเปอร์เซ็นไทล์ในสถานการณ์ต่างๆ พบว่า รูปแบบการเทียบคะแนนเชิงเส้นตรงโดยใช้วิธีของพอททอฟ (Potthoff) ให้ผลที่ไม่น่าพอใจในเกือบทุกสถานการณ์ที่ศึกษา ในกรณีที่กลุ่มตัวอย่างขนาดเท่ากัน รูปแบบอิกวิเปอร์เซ็นไทล์ให้ผลที่มีความคลาดเคลื่อนรวมมากกว่ารูปแบบเชิงเส้นตรง และเมื่อแบบสอบถามสองชุดมีความสัมพันธ์เชิงเส้นโค้ง อันเนื่องมาจากมีความยากต่างกัน และจากกลุ่มตัวอย่างที่ไม่คล้ายคลึงกัน รูปแบบอิกวิเปอร์เซ็นไทล์ให้ผลการเทียบคะแนนดีกว่ารูปแบบเชิงเส้นตรง

วูดดรัฟ (Woodruff. 1986) ได้ศึกษาวิธีการเทียบคะแนนแบบเส้นตรง โดยใช้แบบแผนของ CINEP และแบบจำลองคะแนนจริงสัมพันธ์ โดยใช้กลุ่มตัวอย่าง 4,000 คน ทดสอบกับแบบทดสอบ 5 ฉบับ ที่เป็นฟอร์มเก่า มีความยาวแตกต่างกันคือ ตั้งแต่จำนวน 25 ถึง 115 ข้อ และใช้แบบทดสอบที่เป็นฟอร์มเก่า ร่วมกับแบบทดสอบเชื่อมโยง มาทดสอบกับกลุ่มตัวอย่าง จำนวน 4,000 คน นำคะแนนที่ได้มาศึกษาวิธีการเทียบคะแนนสองวิธี ซึ่งผลการศึกษาพบว่า วิธีการเทียบคะแนนทั้งสองวิธีให้ความเพียงพอในการปฏิบัติ แม้ว่าแบบทดสอบทั้ง 5 ฉบับจะวัดในลักษณะที่ต่างกัน

คูก และปีเตอร์เซน (Cook & Petersen. 1987 : 225-244) ได้ศึกษาปัญหาเกี่ยวกับการทำตามข้อตกลงและวิธีการเทียบคะแนนแบบ IRT โดยดูผลจาก ความคลาดเคลื่อนของการสุ่ม ลักษณะของกลุ่มตัวอย่าง และลักษณะของข้อสอบรวม โดยใช้วิธีการเทียบคะแนนแบบอิกวิเปอร์เซ็นไทล์ ทดสอบกับกลุ่มตัวอย่าง 6 กลุ่ม คือ สองกลุ่มมาจากการสุ่ม กลุ่มหนึ่งเป็นผู้ชายทั้งหมด อีกกลุ่มหนึ่งเป็นผู้หญิงทั้งหมด อีกหนึ่งกลุ่มเป็นผู้มีอายุตั้งแต่ 21-23 ปี และกลุ่มสุดท้ายเป็นผู้มีอายุตั้งแต่ 29 ปีขึ้นไป กลุ่มตัวอย่างนี้มาคะแนนมาจากการสอบ GMAT ที่เทียบคะแนนโดยวิธี IRT การทดสอบใช้แบบทดสอบ 3 ฟอร์ม สอบต่างเวลากันโดยมีข้อสอบรวม 58 ข้อ ซึ่งมีความสัมพันธ์กันสูง ผลของการศึกษาพบว่า ไม่ใช่วิธีที่เกิดจากแบบทดสอบที่มีโครงสร้างเหมือนกัน และชี้ให้เห็นว่าการสอบในเวลาที่แตกต่างกันมีผลต่อการเทียบคะแนน

โคเป (Cope. 1987) ได้เปรียบเทียบการเทียบคะแนนแบบเส้นตรง 5 วิธี ระหว่างวิธีของแองกอฟ ตามแบบแผนการเก็บรวบรวมข้อมูลที่ 5 ซึ่งมี 3 วิธีย่อยกับวิธีของทักเกอร์ (Tucker) และลีไวน์ (Levine) ในกรณีที่แบบทดสอบมีความเชื่อมั่นเท่ากันและใช้แบบทดสอบรวม ผลการศึกษาพบว่า ในสถานการณ์ที่ศึกษาภายใต้แบบแผนที่ 5 ตามวิธีของแองกอฟให้คะแนนดิบที่สมมูลใกล้เคียงกับวิธีการของทักเกอร์และลีไวน์ โดยพบว่าคะแนนดิบที่เท่าเทียมกันจากแบบแผนที่ 5 มีค่าอยู่ระหว่างคะแนนจากวิธีการของทักเกอร์และลีไวน์ ซึ่งสรุปได้ว่า ผลการวิจัยนี้สนับสนุนวิธีการเทียบคะแนนแบบเส้นตรง ตามแบบแผนที่ 5 โดยมีข้อตกลงที่มีข้อจำกัดน้อยกว่าวิธีของทักเกอร์และลีไวน์ ดังนั้น ในทางปฏิบัติควรใช้วิธีการตามแบบแผนที่ 5 ซึ่งมีวิธีการย่อย 3 วิธี โดยเลือกใช้วิธีการใดวิธีการหนึ่งก็ได้

พาแชล, ฮอจตัน และครอมเลย์ (Parshall, Houghton & Kromley. 1995 : 37-54) ศึกษาความคลาดเคลื่อนในการเทียบคะแนน และความลำเอียงทางสถิติในกลุ่มตัวอย่างขนาดเล็ก โดยใช้วิธีเทียบคะแนนเชิงเส้นตรง โดยใช้แบบแผนข้อสอบรวมที่มีประชากรแตกต่างกัน ศึกษาในกลุ่มตัวอย่างขนาด 15, 25, 50, 100 และ 500 คน จากประชากรเทียม 2 กลุ่ม กลุ่มละ 1,000 คน ใช้วิธีการสุ่มแบบใส่คืนที่ ขนาดกลุ่มตัวอย่างขนาดละ 1,000 คน เครื่องมือที่ใช้เป็นแบบทดสอบ 2 ฉบับที่มีลักษณะคู่ขนานกัน และมีข้อสอบรวมภายในมี 5 เนื้อหาวิชา จากผลการศึกษาพบว่า มีความคลาดเคลื่อนของการสุ่มน้อยมาก เมื่อคะแนนใกล้เคียง และจะเพิ่มขึ้นเรื่อยๆ เช่นเดียวกับการกระจายของคะแนนจากค่าเฉลี่ย ส่วนความคลาดเคลื่อนมาตรฐานของคะแนนจะเพิ่มขึ้น เมื่อขนาดของกลุ่มตัวอย่างลดลง และความลำเอียงทางสถิติมีผลกับขนาดของกลุ่มตัวอย่างขนาดน้อยๆ เมื่อคะแนนเบี่ยงเบนออกจากค่าเฉลี่ย 1.5SD

สรุปจากงานวิจัยต่างประเทศพบว่า ในกรณีที่กลุ่มตัวอย่างมีขนาดเท่ากัน การเทียบคะแนนรูปแบบ อิกวิเปอร์เซ็นไทล์ให้ผลที่มีความคลาดเคลื่อนรวมมากกว่ารูปแบบการเทียบคะแนนเชิงเส้นตรง แต่ในกรณีที่ กลุ่มตัวอย่างที่ไม่คล้ายคลึงกัน รูปแบบอิกวิเปอร์เซ็นไทล์ให้ผลการเทียบคะแนนดีกว่ารูปแบบเชิงเส้นตรง ส่วน ความคลาดเคลื่อนมาตรฐานในการเทียบคะแนน และความลำเอียงทางสถิติ เมื่อศึกษากับกลุ่มตัวอย่างขนาดเล็ก โดยใช้วิธีการเทียบคะแนนเชิงเส้นตรง พบว่ามีค่าความคลาดเคลื่อนของการสุ่มน้อยมาก เมื่อคะแนนใกล้เคียง ค่าเฉลี่ย ส่วนความคลาดเคลื่อนมาตรฐานจะเพิ่มขึ้นเมื่อขนาดของกลุ่มตัวอย่างลดลง

ส่วนงานวิจัยในประเทศ มีผู้ทำการวิจัยหลายคนดังนี้

ชูชีพ พงษ์สมบุรณ์ (2528 : 112-118) ได้เปรียบเทียบการเทียบมาตรฐานระหว่างรูปแบบผู้ใช้แบบสอบ รวมกับรูปแบบที่ใช้แบบสอบหลักกับกลุ่มตัวอย่างจากโครงการตรวจสอบคุณภาพการศึกษาชั้นมัธยมศึกษาปีที่ 6 ที่สอบระหว่างวันที่ 13-14 ธันวาคม 2526 จำนวน 3,721 คน โดยใช้วิธีการเทียบมาตรา 3 วิธี คือ การเทียบ เชิงเส้นตรง การเทียบแบบอิกวิเปอร์เซ็นไทล์ และการเทียบโดยใช้โค้งลักษณะข้อสอบ ผลปรากฏว่า ประสิทธิภาพการเทียบคะแนนระหว่างรูปแบบสอบรวมกับรูปแบบที่ใช้ผู้สอบร่วมในแต่ละวิธีมีประสิทธิภาพไม่แตกต่างกัน ความคงที่ของการเทียบคะแนนโดยรูปแบบที่ใช้ผู้สอบร่วมได้ผลว่า วิธีการเทียบคะแนนโดยใช้วิธีอิกวิเปอร์ เซ็นไทล์มีความคงที่มากกว่าการเทียบเชิงเส้นตรง แต่มีความคงที่พอๆ กับการเทียบคะแนนโดยใช้ลักษณะโค้ง ข้อสอบความคงที่ของวิธีการเทียบคะแนนในรูปแบบที่ใช้แบบสอบรวมได้ผลเหมือนกับรูปแบบใช้แบบสอบรวม

ภาวดี ศรีสุขวัฒนานันท์ (2529 : 155-170) ได้เปรียบเทียบผลของการใช้รูปแบบการเทียบมาตราที่ แตกต่างกัน 3 รูปแบบ คือ รูปแบบอิกวิเปอร์เซ็นไทล์ รูปแบบเชิงเส้นตรง และรูปแบบอิงทฤษฎีการตอบสนอง ข้อสอบ แบบโลจิส 3 พารามิเตอร์ โดยใช้ข้อสอบร่วมภายในร้อยละ 20, 40 และ 60 ตามลำดับ ของแบบสอบที่ใช้เทียบคะแนน 35 ข้อ ศึกษากับผู้ทดสอบคัดเลือกเข้าทำงาน และนักศึกษาปริญญาตรีของสถาบันอุดมศึกษา แห่งหนึ่ง ผลปรากฏว่า การเทียบมาตรากับกลุ่มผู้สอบคัดเลือก รูปแบบที่ได้ผลดี คือ รูปแบบอิกวิเปอร์เซ็นไทล์ ส่วนกลุ่มนักศึกษาที่สอบวัดผลสัมฤทธิ์ รูปแบบที่ได้ผลดี คือ รูปแบบเชิงเส้นตรง และพบว่า ความยาวของแบบ สอบรวม มีผลต่อความแม่นยำ และความเพียงพอของการเทียบคะแนน

เรวดี อินทสระ (2530 : 64-67) ศึกษาเปรียบเทียบความคลาดเคลื่อนมาตรฐาน และความเพียงพอ ของการเทียบคะแนนตามทฤษฎีการตอบข้อสอบแบบสามพารามิเตอร์กับวิธีการวิเคราะห์องค์ประกอบ กลุ่มตัว อย่างที่ใช้ในการศึกษาเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2529 จำนวน 2,823 คน แบบทดสอบที่ใช้ เป็นแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ 2 ฉบับๆ ละ 45 ข้อ โดยมีแบบสอบภายในฉบับละ 15 ข้อ ผล การวิจัยพบว่า ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนโดยวิธีสามพารามิเตอร์น้อยกว่าวิธีการใช้ เทคนิคการวิเคราะห์องค์ประกอบ ส่วนความเพียงพอของการเทียบคะแนนโดยวิธีสามพารามิเตอร์อยู่ในระดับที่ น่าพอใจ ส่วนวิธีการวิเคราะห์องค์ประกอบอยู่ในระดับปานกลาง

นิภาพร ยอดเมือง (2539 : บทคัดย่อ) ศึกษาความคลาดเคลื่อน และความลำเอียงในการเทียบ คะแนนเชิงเส้นตรง แบบคะแนนจริงสัมพันธ์ในกลุ่มตัวอย่างขนาดเล็ก โดยศึกษากับกลุ่มตัวอย่างขนาด 30 คน 60 คน 120 คน และ 250 คนว่าแตกต่างกันหรือไม่ ผลการวิจัยพบว่า คะแนนเทียบเฉลี่ยของแต่ละกลุ่มตัวอย่าง แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ค่าความลำเอียงทางสถิติในการเทียบคะแนนก็มีความแตกต่างกันน้อย มากในแต่ละขนาดของกลุ่มตัวอย่าง และค่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนแตกต่างกันเล็กน้อย จึงสรุปได้ว่า กลุ่มตัวอย่างที่สามารถนำไปใช้ในการเทียบคะแนนได้ คือ กลุ่มตัวอย่างทั้งสี่ขนาด แต่ถ้า ต้องการให้ค่าความคลาดเคลื่อนมาตรฐาน และความลำเอียงทางสถิติในการเทียบคะแนนมีน้อย ก็สมควรใช้กลุ่ม ตัวอย่างขนาด 250 คน

พรโสภา พินิจผล (2539 : บทคัดย่อ) เปรียบเทียบผลการเทียบคะแนนในแบบทดสอบวิชาการ พยาบาลสูติศาสตร์ตามทฤษฎีการตอบข้อสอบที่ใช้วิธีการวิเคราะห์ 2 วิธี คือ วิธีการเทียบคะแนนจริงโดยการใช้แบบทดสอบรวม และวิธีการเทียบคะแนนดิบโดยการใช้แบบทดสอบรวม ผลการศึกษาพบว่า คะแนนที่ได้จากการเทียบตามวิธีการเทียบคะแนนจริงกับวิธีการเทียบคะแนนดิบ คะแนนความคลาดเคลื่อนที่ได้จากการเทียบคะแนน และความเพียงพอของการเทียบคะแนนแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

สุนา โสติณณันต์ (2539 : บทคัดย่อ) เปรียบเทียบวิธีการเทียบคะแนนแบบเส้นตรงที่ได้จากคะแนนมาคำนวณด้วยสูตร 4 วิธี โดยมีข้อตกลงว่า แบบทดสอบที่นำมาเทียบคะแนนนั้น ต้องเป็นไปตามแบบจำลองของคะแนนจริงสัมพันธ์ และมีความเชื่อมั่นเท่าเทียมกัน กลุ่มตัวอย่างที่ใช้เป็นนักศึกษาพยาบาลชั้นปีที่ 1 และ 2 พบว่า คะแนนแปลงที่ได้จากวิธีการคำนวณด้วยสูตร 4 วิธี ให้ผลแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

จากการรวบรวมงานวิจัยในประเทศพบว่า รูปแบบของการเทียบคะแนนในสถานการณ์ที่ต่างกันย่อมให้ผลไม่เหมือนกัน เช่น ในการสอบวัดผลสัมฤทธิ์รูปแบบเชิงเส้นตรงได้ผลดีและมีความเหมาะสมมากที่สุด แต่ถ้านำมาใช้ในการสอบคัดเลือกบุคคลเข้าทำงานรูปแบบอิกคิวเปอร์เซ็นต์ให้ผลที่ดีกว่ารูปแบบอื่น แบบสอบรวมที่มีความยากมากจะให้ค่าความคลาดเคลื่อนน้อยกว่า และขนาดของกลุ่มตัวอย่างขนาดเล็กที่เหมาะสมจะใช้ในการเทียบคะแนนเชิงเส้นตรงคือ 250 คน จะให้ค่าความคลาดเคลื่อนมาตรฐาน และความลำเอียงทางสถิติที่น้อยมาก

และจากงานวิจัยที่รวบรวมทั้งต่างประเทศและในประเทศนี้ แสดงให้เห็นว่า ในการศึกษาเกี่ยวกับการเทียบคะแนนระหว่างแบบทดสอบเป็นไปได้หลายวิธี และไม่มีวิธีการเทียบคะแนนวิธีใดเพียงวิธีเดียวที่ให้ผลดีที่สุดในทุกสถานการณ์ ดังนั้นในทางปฏิบัติจึงจำเป็นต้องศึกษาข้อกำหนดต่าง ๆ ของวิธีการเทียบคะแนนแต่ละวิธี เพื่อเป็นแนวทางในการตัดสินใจเลือกวิธีการเทียบคะแนนที่เหมาะสมในแต่ละสถานการณ์การเทียบ เช่น เมื่อแบบทดสอบต่างชุดที่นำมาเทียบคะแนนมีระดับความยากคล้ายคลึงกัน และกลุ่มตัวอย่างเป็นกลุ่มเท่าเทียมกัน ควรใช้วิธีการเทียบคะแนนเชิงเส้นตรง ซึ่งสะดวกในการแปลงคะแนนและมีความเป็นปรนัยมากกว่า ถ้าแบบทดสอบมีความยากต่างกัน กลุ่มตัวอย่างเป็นกลุ่มที่เท่าเทียมกันแล้ว วิธีการเทียบแบบอิกคิวเปอร์เซ็นต์ให้ผลเหมาะสมกว่า แต่ถ้าแบบทดสอบมีความยากต่างกันและกลุ่มตัวอย่างไม่ได้เป็นกลุ่มที่เท่าเทียมกันควรใช้วิธีการแบบอิงทฤษฎีการตอบข้อสอบชนิดสามพารามิเตอร์ จะให้ผลที่น่าพอใจกว่าวิธีอื่น และเมื่อใดที่ขนาดของกลุ่มตัวอย่างค่อนข้างน้อย และแบบทดสอบวัดในสิ่งเดียวกันแล้ว วิธีเชิงเส้นตรงและวิธีการอิงทฤษฎีการตอบข้อสอบชนิดหนึ่งพารามิเตอร์ ให้ผลที่น่าพอใจกว่าวิธีอิกคิวเปอร์เซ็นต์ให้ผลและวิธีอิงทฤษฎีการตอบข้อสอบชนิดสามพารามิเตอร์

โดยสรุปจากงานวิจัยข้างต้นแล้ว พบว่าส่วนใหญ่จะเป็นการศึกษาเชิงทฤษฎีมากกว่าที่จะมาคำนึงถึงการปฏิบัติในด้านการศึกษา เนื่องจากมีวิธีการวิเคราะห์ที่ค่อนข้างยาก ซับซ้อน และใช้เวลามาก ทำให้มีการนำไปใช้ประโยชน์ในการเรียนการสอนตามสภาพความเป็นจริงน้อยมาก ดังนั้นผู้วิจัยจึงมีความสนใจที่จะศึกษาวิธีการเทียบคะแนนเชิงเส้นตรง เนื่องจากเป็นวิธีที่ง่าย มีความสะดวกในการวิเคราะห์ และใช้กับกลุ่มตัวอย่างขนาดเล็กได้ เพื่อเป็นแนวทางให้ครูผู้สอนนำไปเป็นส่วนหนึ่งของการวัดและประเมินผลการเรียนต่อไป

สมมติฐานในการศึกษาค้นคว้า

1. คะแนนที่เทียบได้ของการเทียบคะแนนในแนวนอน โดยวิธีเชิงเส้นตรงเมื่อมีวิธีการแปลงคะแนนที่แตกต่างกัน 3 วิธี มีค่าแตกต่างกัน
2. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในแนวนอน โดยวิธีเชิงเส้นตรง เมื่อมีวิธีการแปลงคะแนนที่แตกต่างกัน 3 วิธี มีค่าแตกต่างกัน
3. ความเพียงพอของการเทียบคะแนนในแนวนอน โดยวิธีเชิงเส้นตรง เมื่อมีวิธีการแปลงคะแนนที่แตกต่างกัน 3 วิธี มีค่าแตกต่างกัน

บทที่ 3

วิธีการดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี ปีการศึกษา 2542 จำนวน 5,290 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี ปีการศึกษา 2542 จำนวน โรงเรียน 28 โรงเรียน จำนวน 820 คน โดยมีวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีขั้นตอนการดำเนินงานดังนี้

ขั้นที่ 1 ในการออกแบบการเทียบคะแนนต้องใช้กลุ่มตัวอย่างในการเทียบคะแนน 2 กลุ่ม และแต่ละกลุ่มควรมีขนาดไม่ต่ำกว่า 400 คน (Brennan & Kolen. 1987 : 285) ผู้วิจัยจึงใช้กลุ่มตัวอย่างทั้งสิ้น 820 คน และกลุ่มทวนผลมีขนาดไม่น้อยกว่า 10 % ของกลุ่มเทียบคะแนน ผู้วิจัยจึงใช้กลุ่มทวนผลอีก 125 คน

ขั้นที่ 2 แบ่งโรงเรียนออกเป็น 3 ขนาด คือ ขนาดใหญ่ ขนาดกลาง และขนาดเล็ก จากการสำรวจมีโรงเรียนขนาดใหญ่ 45 โรงเรียน โรงเรียนขนาดกลาง 98 โรงเรียน และโรงเรียนขนาดเล็ก 142 โรงเรียน

ขั้นที่ 3 สุ่มโรงเรียนตามขนาดของโรงเรียน โดยวิธีการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยมีขนาดโรงเรียนเป็นชั้น (Strata) และโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit) สุ่มมาประมาณ 10 เปอร์เซนต์ของแต่ละขนาด ใช้เป็นกลุ่มเทียบคะแนน และกลุ่มทวนผลอีก 125 คน ดังแสดงรายละเอียดใน

ตาราง 6

ตาราง 6 จำนวนนักเรียนในกลุ่มตัวอย่างที่ใช้ในการเทียบคะแนน

กลุ่มเทียบคะแนน				กลุ่มทานผล	
กลุ่ม 1		กลุ่ม 2			
โรงเรียน	จำนวน	โรงเรียน	จำนวน	โรงเรียน	จำนวน
<u>ขนาดใหญ่</u>		<u>ขนาดใหญ่</u>			
1.ชุมชนบ้านทุ่งโพธิ์	63	1.เมืองปราจีนบุรี	103	1.อนุบาลปราจีนบุรี	125
2.ชุมชนวัดเนินดินแดง	80	2.บ้านวังดินสอ	47		
<u>ขนาดกลาง</u>		<u>ขนาดกลาง</u>			
1.วัดโคกป่าแพ่ง	31	1.วัดราษฎร์รังษี	28		
2. บ้านโคกกระท้อน	29	2. บ้านหนองจิก	29		
3. วัดเกาะมะไฟ	28	3. บ้านเขานันทา	30		
4. บ้านโนนสะอาด	30	4. บ้านหนองสัง	29		
5. ชุมชนบ้านทุ่งแฝก	28	5. วัดศรีสวัสดิ์	29		
<u>ขนาดเล็ก</u>		<u>ขนาดเล็ก</u>			
1. บ้านหนองหอย	16	1. บ้านหนองไผ่ล้อม	17		
2. วัดไชยมงคล	17	2. บ้านบางรุ่งโรจน์	16		
3. บ้านเขาถ้ำ	20	3. วัดศรีมงคล	16		
4. บ้านหนองชะอม	18	4. บ้านหัวเขา	17		
5. บ้านคลองเจ้าแรง	16	5. นิคมพัฒนา	16		
6. บ้านโคกสว่าง	17	6. บ้านโนนสูง	17		
7. บ้านปากคลอง	18	7. บ้านบุพราหมณ์อรุณ	16		
กระดาน		อนุสรณ์			
รวม	410	รวม	410	รวม	125

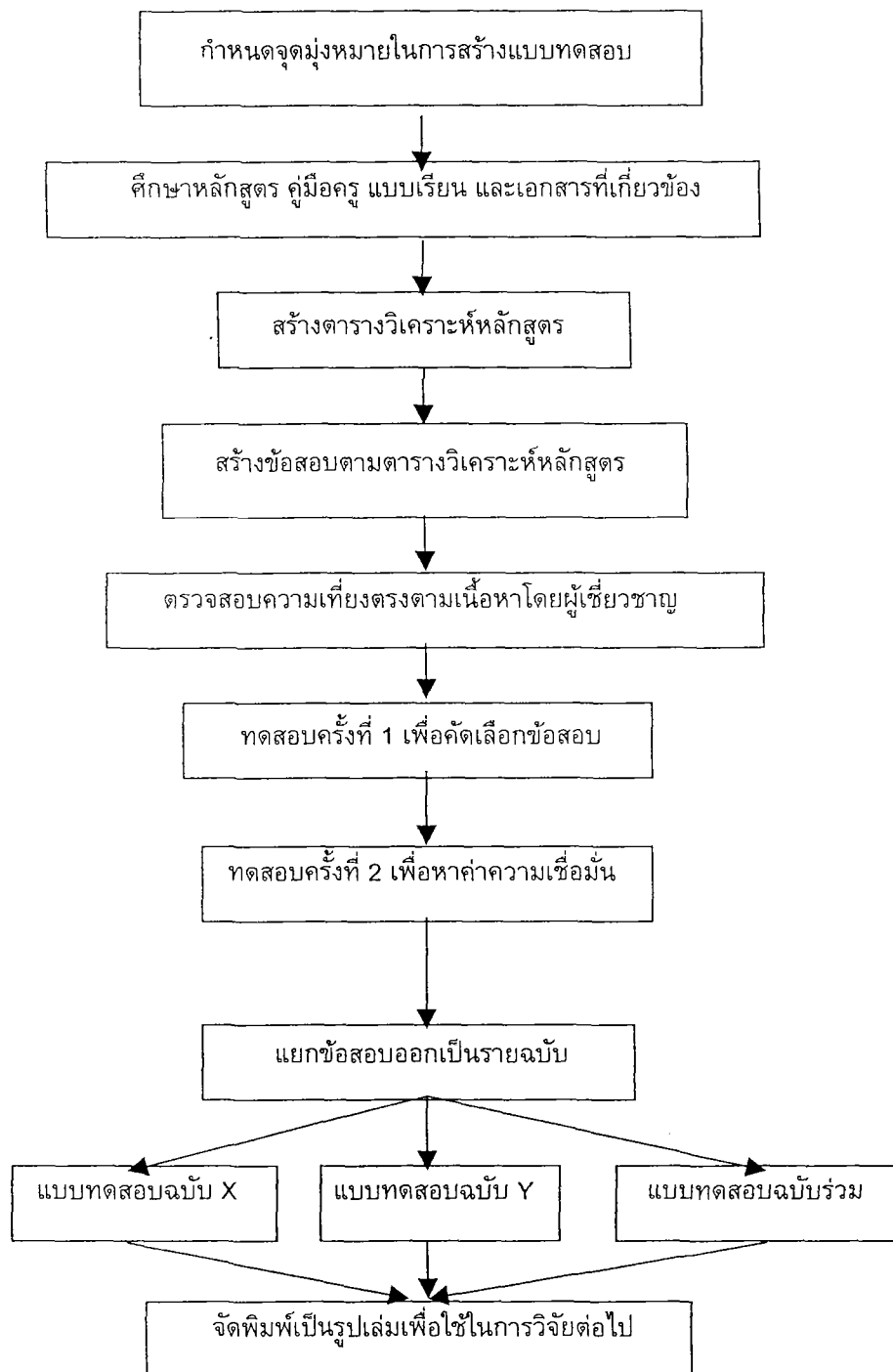
เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่อง สมการและการแก้สมการ เป็นแบบทดสอบชนิดเลือกตอบ 5 ตัวเลือก แบ่งออกเป็น 3 ฉบับ ดังนี้

1. แบบทดสอบวิชาคณิตศาสตร์ ฉบับ X จำนวน 20 ข้อ
2. แบบทดสอบวิชาคณิตศาสตร์ ฉบับ Y จำนวน 20 ข้อ
3. แบบทดสอบรวมวิชาคณิตศาสตร์ จำนวน 10 ข้อ

วิธีการดำเนินการสร้างแบบทดสอบ

การดำเนินการสร้างแบบทดสอบ ผู้วิจัยได้ดำเนินการสร้างแบบทดสอบตามขั้นตอนดังนี้



ภาพประกอบ 1 ขั้นตอนการสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

รายละเอียดในการดำเนินการสร้างแบบทดสอบ มีดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ

✓ เพื่อสร้างแบบทดสอบแบบเลือกตอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษา

ศึกษาปีที่ 6 เรื่อง สมการและการแก้สมการ และนำไปใช้ในการวิจัย

2. ศึกษาหลักสูตร คู่มือครู แบบเรียนและเอกสารที่เกี่ยวข้องในเนื้อหาวิชาคณิตศาสตร์ ชั้นประถมศึกษา

ศึกษาปีที่ 6 เรื่องสมการและการแก้สมการ

3. สร้างตารางวิเคราะห์หลักสูตร โดยกำหนดเนื้อหาและพฤติกรรมที่จะนำมาสร้างแบบทดสอบและพฤติกรรมการเรียนรู้ที่ต้องการวัด

4. สร้างข้อสอบปรนัยชนิดเลือกตอบ 5 ตัวเลือก จำนวน 100 ข้อ ตามเนื้อหาและพฤติกรรมที่กำหนดไว้ในตารางวิเคราะห์หลักสูตร

5. นำแบบทดสอบที่สร้างขึ้นทั้งหมด ไปให้ผู้เชี่ยวชาญซึ่งมีความรู้ในเนื้อหาและการเขียนข้อสอบวิชาคณิตศาสตร์ โดยมีประสบการณ์มาแล้วไม่ต่ำกว่า 5 ปี จำนวน 5 ท่าน ตรวจสอบความเที่ยงตรงเชิงเนื้อหา (Content Validity) ของแบบทดสอบเป็นรายข้อ โดยพิจารณาความสอดคล้องระหว่างข้อสอบแต่ละข้อกับเนื้อหาข้อสอบที่ถือว่าวัดได้ตรงตามเนื้อหา จะต้องได้รับความเห็นชอบจากผู้เชี่ยวชาญ จำนวน 3 ใน 5 ท่าน จึงจะถือว่าข้อสอบนั้นใช้ได้ และหลังจากให้ผู้เชี่ยวชาญพิจารณาความสอดคล้องระหว่างข้อสอบแต่ละข้อกับเนื้อหาแล้ว นำคะแนนไปหาดัชนีความสอดคล้อง ซึ่งถ้ามีค่ามากกว่า 0.5 ถือว่าข้อสอบข้อนั้นๆ วัดได้ตรงตามเนื้อหาจริง ผลปรากฏว่า ได้ข้อสอบที่มีค่าดัชนีความสอดคล้องตั้งแต่ 0.5 – 1.0 จำนวน 100 ข้อ

6. นำแบบทดสอบที่ผ่านการตรวจสอบจากผู้เชี่ยวชาญแล้ว ไปทดลองสอบครั้งที่ 1 กับนักเรียนชั้นประถมศึกษาปีที่ 6 ที่ไม่ใช่กลุ่มตัวอย่าง ได้แก่ โรงเรียนบรรหารวิทยา จำนวน 54 คน โรงเรียนอนุบาลกบินทร์บุรี จำนวน 52 คน โรงเรียนบ้านโคกสัน จำนวน 39 คน โรงเรียนวัดโคกอุดม จำนวน 33 คน และโรงเรียนวัดประสาทรังษยภูมิ จำนวน 22 คน รวมจำนวน 200 คน แล้วนำผลการทดสอบมาตรวจให้คะแนน โดยใช้หลักว่าตอบถูกให้ 1 คะแนน และข้อที่ตอบผิดหรือไม่ตอบหรือตอบมากกว่า 1 แห่ง ให้ 0 คะแนน แล้วนำผลจากการตรวจให้คะแนนไปวิเคราะห์หาค่าความยากและค่าอำนาจจำแนกของข้อสอบเป็นรายข้อและรายตัวเลือกด้วยการใช้โปรแกรมวิเคราะห์ข้อสอบ ผลปรากฏว่า ได้ข้อสอบที่มีค่าความยากตั้งแต่ 0.12 – 0.86 และมีค่าอำนาจจำแนกตั้งแต่ 0.01 – 0.61 จึงคัดเลือกข้อสอบที่มีค่าความยากตั้งแต่ 0.2 – 0.8 และมีค่าอำนาจจำแนกมากกว่าหรือเท่ากับ 0.20 ได้ข้อสอบที่อยู่ในเกณฑ์จำนวน 85 ข้อ และไม่อยู่ในเกณฑ์ 15 ข้อ ผู้วิจัยจึงได้คัดเลือกข้อสอบเลือกไว้จำนวน 50 ข้อ โดยเลือกข้อสอบที่มีค่าความยาก และค่าอำนาจจำแนกที่ใกล้เคียงกันเป็นคู่ๆ ตามเนื้อหาและพฤติกรรมเดียวกัน

7. นำแบบทดสอบที่คัดเลือกแล้ว จำนวน 50 ข้อ ไปทดสอบครั้งที่ 2 กับนักเรียนชั้นประถมศึกษาปีที่ 6 ที่ไม่ใช่กลุ่มตัวอย่าง ได้แก่ โรงเรียนบ้านคลองสิบสอง จำนวน 56 คน โรงเรียนวัดโคกป่าแพ่ง จำนวน 31 คน โรงเรียนบ้านทุ่งประสาธร์ จำนวน 32 คน และโรงเรียนบ้านหอย จำนวน 31 คน รวมจำนวน 150 คน เพื่อหาความเชื่อมั่น โดยใช้สูตร KR-20 ของคูเดอร์-ริชาร์ดสัน มีความเชื่อมั่นที่ 0.940

8. นำข้อสอบแต่ละข้อมาจัดเป็นคู่ตามเนื้อหาและพฤติกรรมเดียวกัน โดยพิจารณาจากค่าความยากและค่าอำนาจจำแนกที่ใกล้เคียงกัน แยกเป็นแบบทดสอบฉบับ X จำนวน 20 ข้อ แบบทดสอบฉบับ Y จำนวน 20 ข้อและแบบทดสอบรวม จำนวน 10 ข้อ สำหรับแบบทดสอบรวมนั้น ใช้ข้อสอบที่เหลือจากแบบทดสอบฉบับ X และแบบทดสอบฉบับ Y แล้ว

9. จัดพิมพ์เป็นรูปเล่มเพื่อใช้ในการวิจัยต่อไป

วิธีการดำเนินการเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในครั้งนี้ ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1. วางแผนการดำเนินการสอบ โดยติดต่อขอความร่วมมือในการนำแบบทดสอบไปสอบกับนักเรียนชั้นประถมศึกษาปีที่ 6 ที่เป็นกลุ่มตัวอย่าง กำหนดวัน เวลาและสถานที่สอบ
2. นำแบบทดสอบทั้ง 3 ฉบับ ไปสอบกับกลุ่มตัวอย่างจำนวน 820 คน โดยแบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม ใช้การสุ่มอย่างง่าย (Simple Random Sampling) มีโรงเรียนเป็นหน่วยการสุ่ม ซึ่งมีรายละเอียดในการทดสอบดังนี้

ตาราง 7 รายละเอียดในการจัดลำดับการสอบเพื่อนำไปเทียบคะแนน

ลำดับที่ของการสอบ	กลุ่มตัวอย่างที่ 1	กลุ่มตัวอย่างที่ 2	กลุ่มทานผล
1	แบบทดสอบฉบับ X	แบบทดสอบฉบับ Y	แบบทดสอบฉบับ X
2	แบบทดสอบฉบับ Y	แบบทดสอบฉบับ X	แบบทดสอบฉบับ Y
3	แบบทดสอบรวม		แบบทดสอบรวม

3. นำผลการสอบที่ได้มาตรวจให้คะแนน โดยใช้หลักว่า ตอบถูก ให้ 1 คะแนน หากตอบผิดหรือไม่ตอบหรือตอบมากกว่า 1 แห่ง ให้ 0 คะแนน

4. นำผลคะแนนที่ได้จากการตรวจมาเทียบคะแนนดังนี้

4.1 วิธีแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ (Random Group-One Test Administered to Each Group) ใช้ผลการสอบครั้งที่ 1 คือ คะแนนจากการสอบแบบทดสอบฉบับ X ของกลุ่มตัวอย่างที่ 1 และคะแนนจากการสอบแบบทดสอบฉบับ Y ของกลุ่มตัวอย่างที่ 2

4.2 วิธีแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบทั้งสองฉบับที่เรียงอันดับการสอบต่างกัน (Random Group-Both Forms Administered to Each Group, Counterbalanced) ใช้ผลการสอบครั้งที่ 1 และครั้งที่ 2 คือ คะแนนจากการสอบแบบทดสอบที่มีการจัดเรียงอันดับการสอบต่างกันโดยใช้คะแนนจากแบบทดสอบฉบับ X และแบบทดสอบฉบับ Y ของกลุ่มตัวอย่างที่ 1 และคะแนนจากการสอบแบบทดสอบฉบับ Y และแบบทดสอบฉบับ X ของกลุ่มตัวอย่างที่ 2

4.3 วิธีแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ และมีแบบทดสอบรวมที่ต้องสอบทั้งสองกลุ่ม (Random Groups-One Test Administered to Each Group, Common Equating Test Administered to Both Groups) ใช้ผลการสอบครั้งที่ 1 และครั้งที่ 3 คือ คะแนนจากการสอบแบบทดสอบฉบับ X และแบบทดสอบรวมของกลุ่มตัวอย่างที่ 1 และคะแนนจากการสอบแบบทดสอบฉบับ Y และแบบทดสอบรวม ของกลุ่มตัวอย่างที่ 2

5. นำผลที่ได้จากการเทียบคะแนนมาคำนวณหาค่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน

6. นำผลที่ได้จากการเทียบคะแนนมาคำนวณหาค่าความเพี้ยนพองของการเทียบคะแนนจากกลุ่มทาน

ผล

ขั้นตอนการวิเคราะห์ข้อมูล

1. การวิเคราะห์หาข้อมูลในเครื่องมือ (Try out) ได้ดำเนินการดังนี้

1.1 หาค่าความยาก (P) โดยใช้สูตรอย่างง่าย และค่าอำนาจจำแนก โดยใช้ค่าสหสัมพันธ์แบบพอยท์-ไบซีเรียล โดยใช้โปรแกรม Item Analysis

1.2 หาค่าความเชื่อมั่น โดยใช้สูตร KR-20 ของคูเดอร์-ริชาร์ดสัน โดยใช้โปรแกรม Item Analysis

2. การวิเคราะห์ข้อมูลการวิจัย ได้ดำเนินการดังนี้

2.1 หาค่าสถิติพื้นฐานของแต่ละกลุ่ม คือ ค่าเฉลี่ย และค่าเบี่ยงเบนมาตรฐาน

2.2 หาคะแนนที่เทียบได้ในแต่ละวิธีการเทียบคะแนน

2.3 หาค่าความคลาดเคลื่อนมาตรฐานของคะแนนการเทียบ โดยใช้สูตรพาร์แชล, ฮอจตัน และครอมเลย์ (Parshall, Houghton and Kromley. 1995)

2.4 หาค่าความเพี้ยนพอย โดยใช้สูตรของปีเตอร์เซนและคนอื่นๆ (Petersen and others. 1989 : 93-94)

2.5 ทดสอบความแตกต่างของสมมติฐาน โดยใช้การวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (Repeated Measures ANOVA) แล้วเปรียบเทียบแตกต่างรายคู่โดยวิธีของเชฟเฟ (Scheffe) (ล้วน สายยศ และอังคณา สายยศ. 2540 : 321)

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. สถิติที่ใช้หาคุณภาพของเครื่องมือ

1.1 หาค่าความยากโดยใช้สูตรอย่างง่าย (ล้วน สายยศ และอังคณา สายยศ. 2538 : 210) โดยใช้โปรแกรม Item Analysis

$$P = \frac{R}{N}$$

เมื่อ	P	คือ	ค่าความยากของคำถามแต่ละข้อ
	R	คือ	จำนวนผู้ตอบถูกในแต่ละข้อ
	N	คือ	จำนวนผู้เข้าสอบทั้งหมด

1.2 หาค่าอำนาจจำแนก โดยใช้ค่าสหสัมพันธ์แบบพอยท์-ไบซีเรียล (ล้วน สายยศ และอังคณา สายยศ. 2538 : 212) โดยใช้โปรแกรม Item Analysis

$$r_{p.bis} = \frac{\bar{X}_p - \bar{X}_f}{S_t} \cdot \sqrt{pq}$$

เมื่อ	$r_{p.bis}$	คือ	ค่าอำนาจจำแนก
	$\bar{X}_p$	คือ	คะแนนเฉลี่ยของนักเรียนที่ทำข้อนั้นถูก
	$\bar{X}_f$	คือ	คะแนนเฉลี่ยของนักเรียนที่ทำข้อนั้นผิด

S_i	คือ	คะแนนเบี่ยงเบนมาตรฐานของแบบทดสอบฉบับนั้น
p	คือ	สัดส่วนของคนที่ทำข้อนั้นได้
q	คือ	สัดส่วนของคนที่ทำข้อนั้นไม่ได้ หรือ $1 - p$

1.3 หาค่าความเชื่อมั่นโดยใช้สูตร KR-20 ของคูเดอร์-ริชาร์ดสัน (ล้วน สายยศ และอังคณา สายยศ. 2538 : 187) โดยใช้โปรแกรม Item Analysis

$$r_{tt} = \frac{K}{K-1} \left\{ 1 - \frac{\sum pq}{S_i^2} \right\}$$

เมื่อ	r_{tt}	คือ	ค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ
	K	คือ	จำนวนข้อสอบ
	p	คือ	สัดส่วนของคนที่ทำถูกในแต่ละข้อ
	q	คือ	$1 - p$
	S_i^2	คือ	ความแปรปรวนของคะแนนที่สอบได้

2. สถิติที่ใช้ในการวิจัย

2.1 สมการเส้นตรงที่ใช้ในการเทียบคะแนนมี 3 วิธี ดังนี้

1) วิธีเทียบคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ (Random Group-One Test Administered to Each Group) (วรเนตร พิชิตเกริกพล. 2536 ; อ้างอิงมาจาก Angoff. 1971 : 569 – 586)

$$\begin{aligned}
 Y &= AX + B \\
 \frac{Y - M_{y2}}{S_{y2}} &= \frac{X - M_{x1}}{S_{x1}} \\
 Y &= \frac{S_{y2}}{S_{x1}} X + M_{y2} - \frac{S_{y2}}{S_{x1}} M_{x1} \\
 A &= \frac{S_{y2}}{S_{x1}} \\
 B &= M_{y2} - A M_{x1}
 \end{aligned}$$

โดยที่	Y	คือ	คะแนนสอบที่เทียบได้จากแบบทดสอบฉบับ X
	M_{y2}	คือ	ค่าเฉลี่ยของคะแนนแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2
	S_{y2}	คือ	ค่าเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2

X	คือ	คะแนนสอบที่ได้จากแบบทดสอบฉบับ X
M _{x1}	คือ	ค่าเฉลี่ยของคะแนนแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1
S _{x1}	คือ	ค่าเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1

2) วิธีเทียบคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบทั้งสองฉบับที่เรียงอันดับการสอบต่างกัน (Random Group-Both Forms Administered to Each Group, Counterbalanced) (วรเนตร พิชิตเกริกพล. 2536 ; อ้างอิงมาจาก Angoff. 1971 : 569 – 586)

$$Y = AX + B$$

$$\frac{Y - M_y}{S_y} = \frac{X - M_x}{S_x}$$

$$M_x = \frac{1}{2}(M_{x1} + M_{x2} - K_x)$$

$$M_y = \frac{1}{2}(M_{y1} + M_{y2} - K_y)$$

$$S_x^2 = \frac{1}{2}(S_{x1}^2 + S_{x2}^2)$$

$$S_y^2 = \frac{1}{2}(S_{y1}^2 + S_{y2}^2)$$

โดยที่

$$A = \sqrt{\frac{S_{y1}^2 + S_{y2}^2}{S_{x1}^2 + S_{x2}^2}}$$

$$B = \frac{1}{2}(M_{y1} + M_{y2}) - \frac{A}{2}(M_{x1} + M_{x2})$$

เมื่อ	Y	คือ	คะแนนสอบที่เทียบได้จากแบบทดสอบฉบับ X
	M _{y2}	คือ	ค่าเฉลี่ยของคะแนนแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2
	S _{y2}	คือ	ค่าเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2
	X	คือ	คะแนนสอบที่ได้จากแบบทดสอบฉบับ X
	M _{x1}	คือ	ค่าเฉลี่ยของคะแนนแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1
	S _{x1}	คือ	ค่าเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1
	K _x	คือ	คะแนนที่แปลงแล้วของแบบทดสอบฉบับ X
	K _y	คือ	คะแนนที่แปลงแล้วของแบบทดสอบฉบับ Y

3) วิธีเทียบคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ โดยเพิ่มแบบทดสอบร่วมที่ต้องสอบทั้งสองกลุ่ม (Random Groups-One Test Administered to Each Group, Common Equating Test Administered to Both Groups) (วรรณทร พิชิตเกริกพล. 2536 ; อ้างอิงมาจาก Angoff. 1971 : 569 – 586)

$$Y = AX + B$$

$$A = \frac{\sigma_y^2}{\sigma_x^2}$$

$$B = \hat{\mu}_y - A\hat{\mu}_x$$

โดยที่

$$\hat{\mu}_x = M_{x1} + b_{xv1}(\mu_v - M_{v1})$$

$$\hat{\mu}_y = M_{y2} + b_{yv2}(\mu_v - M_{v2})$$

$$\sigma_x^2 = S_{x1}^2 + b_{xv1}^2(\sigma_v^2 - S_{v1}^2)$$

$$\sigma_y^2 = S_{y2}^2 + b_{yv2}^2(\sigma_v^2 - S_{v2}^2)$$

เมื่อ

$$\mu_v = M_{vt}$$

$$\sigma_v^2 = S_{vt}^2$$

$$t = n_1 + n_2$$

เมื่อ

μ_x, σ_x^2	คือ	ค่าประมาณคะแนนเฉลี่ย และความแปรปรวนของกลุ่มตัวอย่างทั้งหมดที่ทำแบบทดสอบฉบับ X
μ_y, σ_y^2	คือ	ค่าประมาณคะแนนเฉลี่ย และความแปรปรวนของกลุ่มตัวอย่างทั้งหมดที่ทำแบบทดสอบฉบับ Y
M_{y2}	คือ	ค่าเฉลี่ยของคะแนนแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2
S_{y2}	คือ	ค่าเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2
M_{x1}	คือ	ค่าเฉลี่ยของคะแนนแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1
S_{x1}	คือ	ค่าเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1
b_{xv1}	คือ	สัมประสิทธิ์ถดถอยของคะแนนจากแบบทดสอบฉบับ X จากแบบทดสอบร่วม โดยกลุ่มที่ 1
b_{yv2}	คือ	สัมประสิทธิ์ถดถอยของคะแนนจากแบบทดสอบฉบับ Y จากแบบทดสอบร่วม โดยกลุ่มที่ 2

2.2 สูตรในการหาความคลาดเคลื่อนมาตรฐาน (Parshall, Houghton and Kromley. 1995)

$$SE(\theta_i) = \sqrt{\sum(\theta_{ij} - \theta_i')^2 / n - 1}$$

เมื่อ	$SE(\theta_i)$	คือ	ความคลาดเคลื่อนมาตรฐานของคะแนนเทียบตัวที่ i
	θ_{ij}	คือ	คะแนนเทียบคะแนนตัวที่ i ของกลุ่มตัวอย่าง j
	θ_i'	คือ	คะแนนเทียบคะแนนตัวที่ i เฉลี่ยจากกลุ่มตัวอย่าง
	n	คือ	จำนวนกลุ่มตัวอย่างที่ใช้ในการเทียบคะแนน

2.3 หาค่าความเพียงพอของการเทียบคะแนน โดยประเมินค่าดัชนีความแตกต่าง อาศัยหลักเกณฑ์ในการประเมินของปีเตอร์เซน และคนอื่นๆ (Petersen and others. 1982 : 93-94) ซึ่งกำหนดคุณภาพในการเทียบมาตรฐานตามระดับการยอมรับดังนี้

ระดับที่น่าพอใจอย่างมาก	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	$C > (.20SD_x)^2$

คำนวณได้จากสูตรดังนี้

$$C = \sum(X_i - X_j')^2 / nS_x^2$$

เมื่อ	C	คือ	ค่าดัชนีความแตกต่างระหว่างรูปแบบการเทียบคะแนน
	X_i	คือ	คะแนนจากแบบทดสอบฉบับ X ของผู้เข้าสอบคนที่ i ของกลุ่มทวนผล
	X_j	คือ	คะแนนจากแบบทดสอบฉบับ X ที่ได้จากการนำคะแนนจากแบบทดสอบฉบับ Y ไปเทียบจากตารางการเทียบคะแนนที่วิเคราะห์ได้คนที่ i ของกลุ่มทวนผล
	n	คือ	จำนวนคนของกลุ่มสอบทวนผล
	S_x^2	คือ	ค่าความแปรปรวนของการแจกแจงคะแนนจากแบบทดสอบฉบับ X ของกลุ่มทวนผล

2.4 ทดสอบความแตกต่างของสมมติฐาน โดยใช้การวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (Repeated Measures ANOVA) แล้วเปรียบเทียบแตกต่างรายคู่โดยวิธีของเชฟเฟ (Scheffe) (ล้วน สายยศ และอังคณา สายยศ. 2540 : 321)

$$S = \sqrt{(k-1)F^* (MS_{\text{within}}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

เมื่อ	S	คือ	ค่าวิกฤตของ Scheffe's S-test
	k	คือ	จำนวนกลุ่มตัวอย่าง
	F*	คือ	ค่า F ที่เปิดจากตาราง (critical value)
	MS _{within}	คือ	ค่า mean square within group ที่คำนวณได้แล้วใน การวิเคราะห์ความแปรปรวน
	n ₁ , n ₂	คือ	ขนาดของกลุ่มตัวอย่างที่ 1 และ 2

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการเสนอผลการวิเคราะห์ข้อมูล ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

X	แทน	คะแนนที่ได้จากการสอบแบบทดสอบฉบับ X
Y	แทน	คะแนนที่ได้จากการเทียบคะแนนจากแบบทดสอบฉบับ Y
M_{x1}	แทน	ค่าเฉลี่ยของคะแนนจากแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1
M_{x2}	แทน	ค่าเฉลี่ยของคะแนนจากแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 2
M_{y1}	แทน	ค่าเฉลี่ยของคะแนนจากแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 1
M_{y2}	แทน	ค่าเฉลี่ยของคะแนนจากแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2
S_{x1}	แทน	ค่าเบี่ยงเบนมาตรฐานของคะแนนจากแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1
S_{x2}	แทน	ค่าเบี่ยงเบนมาตรฐานของคะแนนจากแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 2
S_{y1}	แทน	ค่าเบี่ยงเบนมาตรฐานของคะแนนจากแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 1
S_{y2}	แทน	ค่าเบี่ยงเบนมาตรฐานของคะแนนจากแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2
b_{xv1}	แทน	สัมประสิทธิ์การถดถอยของคะแนนจากแบบทดสอบฉบับ X จากแบบทดสอบรวม จากการสอบโดยกลุ่มที่ 1
b_{yv2}	แทน	สัมประสิทธิ์การถดถอยของคะแนนจากแบบทดสอบฉบับ Y จากแบบทดสอบรวม จากการสอบโดยกลุ่มที่ 2
μ_x	แทน	ค่าประมาณค่าเฉลี่ยของคะแนนจากแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1
μ_y	แทน	ค่าประมาณค่าเฉลี่ยของคะแนนจากแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2
σ_x^2	แทน	ความแปรปรวนของคะแนนจากแบบทดสอบฉบับ X จากการสอบโดยกลุ่มที่ 1
σ_y^2	แทน	ความแปรปรวนของคะแนนจากแบบทดสอบฉบับ Y จากการสอบโดยกลุ่มที่ 2
SEE	แทน	ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน
C	แทน	ดัชนีความเพียงพอของการเทียบคะแนน
n	แทน	จำนวนกลุ่มตัวอย่างที่ใช้ในการเทียบคะแนน

การนำเสนอผลการวิเคราะห์ข้อมูล

ในการวิเคราะห์ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับดังนี้

1. ค่าสถิติพื้นฐานของแบบทดสอบ
2. คะแนนที่ได้จากการเทียบคะแนน
3. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน
4. ความเพียงพอของการเทียบคะแนน

ผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานของแบบทดสอบ

จากการนำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ เรื่อง สมการและการแก้สมการ จำนวน 2 ฉบับ คือ ฉบับ X และฉบับ Y จำนวนฉบับละ 20 ข้อ คะแนนเต็ม 20 คะแนน ไปทดสอบกับนักเรียน ชั้นประถมศึกษาปีที่ 6 จำนวน 820 คน โดยแบ่งออกเป็น 2 กลุ่มๆ ละ 410 คน (รายละเอียดโรงเรียนแต่ละกลุ่ม ดังแสดงในหน้า 31) ได้ค่าสถิติพื้นฐานของคะแนนดังแสดงในตาราง 8

ตาราง 8 ค่าสถิติพื้นฐานของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

แบบทดสอบ \ กลุ่มที่	กลุ่ม 1		กลุ่ม 2	
	$\bar{X}$	S	$\bar{X}$	S
ฉบับ X	12.30	4.45	11.04	4.52
ฉบับ Y	10.56	4.39	10.76	3.66

จากตาราง 8 แสดงให้เห็นว่า ในแบบทดสอบฉบับ X ของนักเรียนกลุ่ม 1 มีคะแนนเฉลี่ยสูงกว่าแบบทดสอบฉบับ X ของนักเรียนกลุ่ม 2 อยู่ 1.26 คะแนน และมีการกระจายคะแนนในรูปความเบี่ยงเบนมาตรฐานของนักเรียนกลุ่ม 1 และนักเรียนกลุ่ม 2 มีการกระจายคะแนนที่ใกล้เคียงกัน ส่วนในแบบทดสอบฉบับ Y ของนักเรียนกลุ่ม 1 มีคะแนนเฉลี่ยต่ำกว่าแบบทดสอบฉบับ Y ของนักเรียนกลุ่ม 2 อยู่ 0.20 คะแนน และมีการกระจายคะแนนในรูปความเบี่ยงเบนมาตรฐานของนักเรียนกลุ่ม 1 มีค่าสูงกว่าความเบี่ยงเบนมาตรฐานของนักเรียนกลุ่ม 2 แสดงว่า คะแนนจากแบบทดสอบฉบับ Y ของนักเรียนกลุ่ม 1 มีการกระจายคะแนนมากกว่าคะแนนจากแบบทดสอบฉบับ Y ของนักเรียนกลุ่ม 2

2. คะแนนที่ได้จากการเทียบ

จากรูปแบบการเทียบคะแนนทั้ง 3 วิธี สามารถนำมาสร้างตารางการเทียบคะแนนได้ดังนี้ ตาราง 9 คะแนนที่ได้จากการเทียบจากแบบทดสอบฉบับ X ไปสู่แบบทดสอบฉบับ Y ในรูปแบบการเทียบคะแนนทั้ง 3 วิธี

วิธีการเทียบคะแนน	ช่วงคะแนน	$\bar{X}$	S
วิธีที่ 1	0.39 – 18.03	10.49	4.37
วิธีที่ 2	1.96 – 18.16	11.23	4.01
วิธีที่ 3	3.33 – 21.25	13.50	4.39

จากตาราง 9 แสดงให้เห็นว่า คะแนนที่ได้จากการเทียบโดยวิธีที่ 1 มีค่าอยู่ระหว่าง 0.39 – 18.03 คะแนนที่ได้จากการเทียบโดยวิธีที่ 2 มีค่าอยู่ระหว่าง 1.96 – 18.16 และคะแนนที่ได้จากการเทียบโดยวิธีที่ 3 มีค่าอยู่ระหว่าง 3.33 – 21.25 คะแนนเฉลี่ยของการเทียบโดยวิธีที่ 3 มีค่ามากที่สุดคือ 13.50 วิธีที่ 2 มีค่าเท่ากับ 11.23 และวิธีที่ 1 มีค่าเท่ากับ 10.49 ตามลำดับ และวิธีที่ 3 มีการกระจายคะแนนที่ได้จากการเทียบมากที่สุดคือ 4.39 และวิธีที่ 2 มีการกระจายคะแนนที่ได้จากการเทียบน้อยที่สุดคือ 4.01 ผู้วิจัยจึงนำคะแนนที่ได้จากการเทียบทั้ง 3 วิธี มาทดสอบความแตกต่างของคะแนนที่เทียบได้ดังแสดงในตาราง 10

ตาราง 10 การเปรียบเทียบความแตกต่างของคะแนนที่เทียบได้จากการเทียบคะแนน 3 วิธี

แหล่ง	SS	df	MS	F
ระหว่างบุคคล	22244.25	409	54.39	20625.40**
ระหว่างเงื่อนไข	2062.54	2	1031.27	
Error	39.42	818	0.05	
รวม	24346.21	1229		

$$F_{.01 (2,818)} = 3.01$$

จากตาราง 10 แสดงให้เห็นว่า คะแนนที่ได้จากการเทียบคะแนนทั้ง 3 วิธี แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 นั่นก็คือ คะแนนที่ได้จากการเทียบคะแนนทั้ง 3 วิธี มีความแตกต่างกัน อย่างน้อย 1 คู่ จึงได้ทำการเปรียบเทียบความแตกต่างรายคู่โดยใช้วิธีของเซฟเฟ (Scheffe) ดังแสดงในตาราง 11

ตาราง 11 การเปรียบเทียบความแตกต่างรายคู่ของคะแนนที่เทียบได้ 3 วิธี

ค่าเฉลี่ย	$\bar{x}_1 = 10.49$	$\bar{x}_2 = 11.23$	$\bar{x}_3 = 13.50$
$\bar{x}_1 = 10.49$	-	0.74*	3.01*
$\bar{x}_2 = 11.23$	-	-	2.27*
$\bar{x}_3 = 13.50$			-

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 11 แสดงให้เห็นว่า คะแนนที่ได้จากการเทียบวิธีที่ 1 วิธีที่ 2 และวิธีที่ 3 มีค่าแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน

จากคะแนนที่ได้จากการเทียบ เมื่อนำมาเปรียบเทียบคะแนนที่ได้จากการสอบ แล้วนำมาคำนวณหาค่าความคลาดเคลื่อนมาตรฐานในแต่ละระดับคะแนนของแต่ละวิธี ดังต่อไปนี้

ตาราง 12 ช่วงคะแนนและความคลาดเคลื่อนมาตรฐานของคะแนนที่เทียบได้ในการเทียบคะแนน 3 วิธี

วิธีการเทียบคะแนน	ช่วงคะแนน	SE (θ_i)
วิธีที่ 1	0.39 – 18.03	0.01 – 0.50
วิธีที่ 2	1.96 – 18.16	0.01 – 0.46
วิธีที่ 3	3.33 – 21.25	0.01 – 0.50

จากตาราง 12 แสดงให้เห็นว่า ช่วงคะแนนที่เทียบได้ในการเทียบคะแนนวิธีที่ 1 มีค่าน้อยที่สุด คือ อยู่ระหว่าง 0.39 – 18.03 วิธีที่ 2 มีช่วงคะแนนอยู่ระหว่าง 1.96 – 18.16 และวิธีที่ 3 มีช่วงคะแนนอยู่ระหว่าง 3.33 – 21.25 ซึ่งมีค่ามากที่สุด ส่วนค่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในรูปแบบการเทียบคะแนนวิธีที่ 1 และวิธีที่ 3 มีค่าอยู่ระหว่าง 0.01 – 0.50 เหมือนกัน และค่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในรูปแบบการเทียบคะแนนวิธีที่ 2 มีค่าอยู่ระหว่าง 0.01 – 0.46 ซึ่งมีค่าน้อยกว่าการเทียบคะแนนในรูปแบบการเทียบในวิธีที่ 1 และวิธีที่ 3 ผู้วิจัยจึงได้นำความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน 3 วิธี มาทดสอบความแตกต่างของความคลาดเคลื่อนมาตรฐานได้ ดังแสดงในตาราง 13

ตาราง 13 การเปรียบเทียบความแตกต่างของค่าความคลาดเคลื่อนมาตรฐานที่ได้จากรูปแบบการเทียบคะแนน 3 วิธี

แหล่ง	SS	df	MS	F
ระหว่างบุคคล	30.52	409	0.07	1.50
ระหว่างเงื่อนไข	0.06	2	0.03	
Error	12.89	818	0.02	
รวม	43.47	1229		

$$F_{.05 (2, 1227)} = 2.99$$

จากตาราง 13 แสดงว่า ค่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนทั้ง 3 วิธี แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4. ความเพียงพอของการเทียบคะแนน

ในการหาความเพียงพอของการเทียบคะแนน ผู้วิจัยได้นำแบบทดสอบทั้งสามฉบับไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการหาผลการเทียบคะแนน แล้วนำมาคำนวณหาความคลาดเคลื่อนที่ได้จากการเทียบคะแนน เพื่อที่จะสามารถนำไปประเมินความเพียงพอของการเทียบคะแนน ดังแสดงในตาราง 14

ตาราง 14 ความเพียงพอของการเทียบคะแนนทั้ง 3 วิธี

คะแนน	$(X - X^*)^2$ ของวิธีที่ 1	$(X - X^*)^2$ ของวิธีที่ 2	$(X - X^*)^2$ ของวิธีที่ 3
5	6.45	0.23	0.37
6	6.20	0.32	0.64
7	5.95	0.44	0.98
8	5.71	0.56	1.39
9	5.48	0.71	1.88
10	5.24	0.86	2.43
11	5.02	1.04	3.06
12	4.80	1.23	3.76
13	4.58	1.44	4.54
14	4.37	1.66	5.38
15	4.16	1.90	6.30
16	3.96	2.16	7.29
17	3.76	2.43	8.35
18	3.57	2.72	9.49
19	3.39	3.03	10.69
20	3.20	3.35	11.97
k	16	16	16
SD	3.91	4.32	4.09
c	0.3117	0.0774	0.2805
$(.05SD)^2$	0.0382	0.0467	0.0418

จากตาราง 14 แสดงให้เห็นว่า ค่าคะแนนความคลาดเคลื่อนของการเทียบคะแนนวิธีที่ 1 มีค่าอยู่ระหว่าง 3.20 – 6.45 วิธีที่ 2 มีค่าอยู่ระหว่าง 0.23 – 3.35 และวิธีที่ 3 มีค่าอยู่ระหว่าง 0.37 – 11.97 แต่เมื่อคำนวณหาค่าความเพียงพอของการเทียบคะแนนแล้วนำมาเปรียบเทียบกับเกณฑ์พบว่า วิธีการเทียบคะแนนวิธีที่ 2 ให้ค่าความเพียงพอของการเทียบคะแนนอยู่ในระดับที่น่าพอใจเพียงวิธีเดียว ซึ่งการหาค่าความเพียงพอของการเทียบคะแนนที่ได้จากการเทียบคะแนน 3 วิธี สามารถนำเสนอเปรียบเทียบให้เห็นได้ชัดดังตาราง 15

ตาราง 15 การเปรียบเทียบความเพียงพอของการเทียบคะแนนทั้ง 3 วิธี

วิธีการเทียบคะแนน	N	$\sum(x_i - x_j)^2$	S^2	C
วิธีที่ 1	125	596.56	15.31	0.3117
วิธีที่ 2	125	180.66	18.67	0.0774
วิธีที่ 3	125	587.74	16.76	0.2805

จากตาราง 15 แสดงให้เห็นว่า ความเพียงพอของการเทียบคะแนนวิธีที่ 1 ให้ค่าสูงกว่าการเทียบคะแนนวิธีที่ 3 และการเทียบคะแนนวิธีที่ 2 ตามลำดับ แต่เมื่อเปรียบเทียบกับเกณฑ์ของความเพียงพอในการเทียบคะแนนแล้วพบว่า วิธีที่ 2 อยู่ในเกณฑ์ที่น่าพอใจเพียงวิธีเดียว

บทที่ 5

สรุปผล อภิปราย และข้อเสนอแนะ

ความมุ่งหมายของการวิจัย ✓

ในการศึกษาครั้งนี้ มีความมุ่งหมายเพื่อศึกษาเปรียบเทียบผลการเทียบคะแนนโดยวิธีเชิงเส้นตรง เมื่อใช้วิธีการแปลงคะแนนที่แตกต่างกัน ในแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ โดยมีความมุ่งหมายเฉพาะดังนี้

1. เพื่อเปรียบเทียบคะแนนที่เทียบได้ของการเทียบคะแนนในแนวนอน โดยวิธีเชิงเส้นตรง เมื่อมีการแปลงคะแนนที่แตกต่างกัน
2. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในแนวนอน โดยวิธีเชิงเส้นตรง เมื่อมีการแปลงคะแนนที่แตกต่างกัน
3. เพื่อเปรียบเทียบความเพียงพอของการเทียบคะแนนในแนวนอน โดยวิธีเชิงเส้นตรง เมื่อมีการแปลงคะแนนที่แตกต่างกัน

ขอบเขตของประชากร

กลุ่มตัวอย่าง ✓

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี จำนวน 808 คน ได้มาจากการสุ่มแบบแบ่งชั้น (Stratified random sampling) โดยมีขนาดโรงเรียนเป็นชั้น (Strata) และมีโรงเรียนเป็นหน่วยการสุ่ม (Sampling Unit)

เครื่องมือที่ใช้ในการศึกษาค้นคว้า ✓

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 เรื่อง สมการและการแก้สมการ จำนวน 50 ข้อ เป็นแบบทดสอบชนิดเลือกตอบ 5 ตัว เลือก แบ่งออกเป็น 3 ฉบับ โดยพิจารณาจากค่าความยากและค่าอำนาจจำแนกที่ใกล้เคียงกันจะได้แบบทดสอบ 3 ฉบับ ดังนี้

1. แบบทดสอบวิชาคณิตศาสตร์ ฉบับ X จำนวน 20 ข้อ
2. แบบทดสอบวิชาคณิตศาสตร์ ฉบับ Y จำนวน 20 ข้อ
3. แบบทดสอบรวมวิชาคณิตศาสตร์ จำนวน 10 ข้อ

วิธีดำเนินการเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูลในครั้งนี้ ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1. วางแผนการดำเนินการสอบ โดยติดต่อขอความร่วมมือในการนำแบบทดสอบไปสอบกับนักเรียนชั้นประถมศึกษาปีที่ 6 ที่เป็นกลุ่มตัวอย่าง กำหนดวัน เวลาและสถานที่สอบ

2. นำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนที่สร้างขึ้นจำนวน 100 ข้อ ไปทดลองสอบกับนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 250 คนที่ไม่ใช่กลุ่มตัวอย่าง
3. นำผลการสอบที่ได้มาตรวจให้คะแนน และเตรียมข้อมูลสำหรับการวิเคราะห์รายข้อ แล้วหาความเชื่อมั่น ก่อนที่จะนำไปวิเคราะห์ตามวิธีการเทียบคะแนนเชิงเส้นตรง
4. นำข้อสอบที่มีคุณภาพจากการวิเคราะห์ในข้อ 3. มาจัดเป็นแบบทดสอบจำนวน 3 ฉบับ คือ ฉบับ X และฉบับ Y ฉบับละ 20 ข้อ และแบบทดสอบรวมจำนวน 10 ข้อ
5. นำแบบทดสอบทั้ง 3 ฉบับ ไปทดสอบกับกลุ่มตัวอย่างจำนวน 820 คน โดยแบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม กลุ่มละ 410 คน โดยให้แต่ละกลุ่มทำแบบทดสอบที่มีการจัดลำดับการสอบแตกต่างกัน (รายละเอียดดังตาราง 2 หน้า 33) และนำแบบทดสอบนี้ไปสอบกับกลุ่มทนายผลอีกจำนวน 125 คน

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลดำเนินการโดยแบ่งออกเป็น 4 ตอน ดังนี้

1. ค่าสถิติพื้นฐานของแบบทดสอบ
2. คะแนนที่ได้จากการเทียบคะแนน
3. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน
4. ความเพียงพอของการเทียบคะแนน

สรุปผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานของแบบทดสอบ พบว่าแบบทดสอบฉบับ X ของนักเรียนกลุ่ม 1 มีคะแนนเฉลี่ยสูงกว่าแบบทดสอบฉบับ X ของนักเรียนกลุ่ม 2 อยู่ 1.26 คะแนน และมีการกระจายคะแนนในรูปความเบี่ยงเบนมาตรฐานของนักเรียนกลุ่ม 1 และนักเรียนกลุ่ม 2 มีการกระจายคะแนนที่ใกล้เคียงกัน ส่วนในแบบทดสอบฉบับ Y ของนักเรียนกลุ่ม 1 มีคะแนนเฉลี่ยต่ำกว่าแบบทดสอบฉบับ Y ของนักเรียนกลุ่ม 2 อยู่ 0.20 คะแนน และมีการกระจายคะแนนในรูปความเบี่ยงเบนมาตรฐานของนักเรียนกลุ่ม 1 มีค่าสูงกว่าความเบี่ยงเบนมาตรฐานของนักเรียนกลุ่ม 2 แสดงว่า คะแนนจากแบบทดสอบฉบับ Y ของนักเรียนกลุ่ม 1 มีการกระจายคะแนนมากกว่าคะแนนจากแบบทดสอบฉบับ Y ของนักเรียนกลุ่ม 2

2. คะแนนที่ได้จากการเทียบ

2.1 คะแนนที่ได้จากการเทียบคะแนนจากแบบทดสอบฉบับ X ไปสู่แบบทดสอบฉบับ Y พบว่าคะแนนที่ได้จากการเทียบโดยวิธีที่ 1 มีค่าอยู่ระหว่าง 0.39 – 18.03 คะแนนที่ได้จากการเทียบโดยวิธีที่ 2 มีค่าอยู่ระหว่าง 1.96 – 18.16 และคะแนนที่ได้จากการเทียบโดยวิธีที่ 3 มีค่าอยู่ระหว่าง 3.33 – 21.25 คะแนนเฉลี่ยของการเทียบโดยวิธีที่ 3 มีค่ามากที่สุดคือ 13.50 วิธีที่ 2 มีค่าเท่ากับ 11.23 และวิธีที่ 1 มีค่าเท่ากับ 10.49 ตามลำดับ และวิธีที่ 3 มีการกระจายคะแนนที่ได้จากการเทียบมากที่สุดคือ 4.39 และวิธีที่ 2 มีการกระจายคะแนนที่ได้จากการเทียบน้อยที่สุดคือ 4.01

2.2 การเปรียบเทียบความแตกต่างของคะแนนที่เทียบได้ พบว่า คะแนนที่ได้จากการเทียบคะแนนทั้ง 3 วิธี แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 นั่นก็คือ คะแนนที่ได้จากการเทียบคะแนนทั้ง 3 วิธี มีความแตกต่างกันอย่างน้อย 1 คู่ จึงได้ทำการเปรียบเทียบความแตกต่างรายคู่โดยใช้วิธีของเชฟเฟ

(Scheffe) ผลปรากฏว่า คะแนนที่ได้จากการเทียบวิธีที่ 1 วิธีที่ 2 และวิธีที่ 3 แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน

3.1 ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน พบว่าค่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในรูปแบบการเทียบคะแนนวิธีที่ 1 และวิธีที่ 3 มีค่าอยู่ระหว่าง 0.01 – 0.50 เหมือนกัน และค่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในรูปแบบการเทียบคะแนนวิธีที่ 2 มีค่าอยู่ระหว่าง 0.01 – 0.46 ซึ่งมีค่าน้อยกว่าการเทียบคะแนนในรูปแบบการเทียบในวิธีที่ 1 และวิธีที่ 3

3.2 การเปรียบเทียบความแตกต่างของคะแนนความคลาดเคลื่อนที่ได้จากการเทียบคะแนน พบว่าค่าความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนทั้ง 3 วิธี แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ นั่นก็คือคะแนนความคลาดเคลื่อนที่ได้จากการเทียบคะแนนทั้ง 3 วิธี ไม่แตกต่างกัน

4. ความเพียงพอของการเทียบคะแนน

4.1 ความเพียงพอของการเทียบคะแนน พบว่าค่าความคลาดเคลื่อนของวิธีที่ 1 มีค่าอยู่ระหว่าง 3.20 – 6.45 วิธีที่ 2 มีค่าอยู่ระหว่าง 0.23 – 3.35 และวิธีที่ 3 มีค่าอยู่ระหว่าง 0.37 – 11.97 จะเห็นได้ว่า วิธีที่ 2 มีค่าความคลาดเคลื่อนอยู่ในช่วงน้อยที่สุด และวิธีที่ 3 มีค่าความคลาดเคลื่อนอยู่ในช่วงกว้างมากที่สุด

4.2 การเปรียบเทียบความเพียงพอของการเทียบคะแนน พบว่าความเพียงพอของการเทียบคะแนนวิธีที่ 1 ให้ค่าสูงกว่าการเทียบคะแนนวิธีที่ 3 และการเทียบคะแนนวิธีที่ 2 ตามลำดับ แต่เมื่อเปรียบเทียบกับเกณฑ์ของความเพียงพอในการเทียบคะแนนแล้วพบว่า วิธีที่ 2 อยู่ในเกณฑ์ที่น่าพอใจเพียงวิธีเดียว

อภิปรายผล

1. การเปรียบเทียบความแตกต่างของคะแนนที่เทียบได้จากการเทียบคะแนนตามวิธีการแปลงคะแนนที่แตกต่างกัน 3 วิธี มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะว่า วิธีการจัดอันดับในการสอบแตกต่างกัน กล่าวคือ ในวิธีที่ 1 ผู้สอบในแต่ละกลุ่มจะทำแบบทดสอบเพียงฉบับเดียว ดังนั้นจึงไม่มีความเหนื่อยล้า และไม่สามารถฝึกฝนการทำข้อสอบได้ แต่จะไม่มีข้อมูลที่จะนำมาใช้ปรับความแตกต่างของกลุ่มตัวอย่าง ส่วนในวิธีที่ 2 ผู้สอบมีโอกาสดำเนินการแบบทดสอบทั้งสองฉบับ ทำให้ผู้สอบเกิดความเหนื่อยล้า และเกิดการเรียนรู้จากการทำแบบทดสอบฉบับแรก และในวิธีที่ 3 ผู้สอบมีโอกาสดำเนินการแบบทดสอบคนละฉบับ แต่มีแบบทดสอบรวมเพื่อนำมาใช้ปรับความแตกต่างของกลุ่มตัวอย่าง (Petersen and others. 1989 : 244-246) ดังนั้นจะเห็นได้ว่า ค่าสถิติที่คำนวณได้แต่วิธีของการแปลงคะแนนนั้นมีวิธีการวิเคราะห์ที่แตกต่างกันตามการจัดอันดับการสอบ และอาจเนื่องมาจากการวิเคราะห์วิธีการแปลงคะแนนแต่ละวิธีมีรูปแบบของสมการในการเทียบคะแนนที่แตกต่างกัน ทำให้ผลของการเทียบคะแนนที่ได้มีค่าแตกต่างกันตามไปด้วย

2. การเปรียบเทียบความแตกต่างของคะแนนความคลาดเคลื่อนมาตรฐานที่ได้จากการเทียบคะแนนเชิงเส้นตรงทั้ง 3 วิธี มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ เนื่องจากความคลาดเคลื่อนมาตรฐานที่ได้จากการเทียบคะแนนทั้ง 3 วิธีไม่แตกต่างกัน ทั้งนี้อาจเป็นเพราะว่า ไม่ว่าจะใช้วิธีการแปลงคะแนนในแบบใด หรือมีสมการในการเทียบคะแนนที่แตกต่างกันทั้ง 3 วิธีก็ตาม แต่วิธีการวิเคราะห์หาค่าความคลาดเคลื่อนมาตรฐานใช้วิธีการคิดวิเคราะห์เดียวกันทั้ง 3 วิธี ทำให้ค่าความคลาดเคลื่อนมาตรฐานที่ได้มีค่าแตกต่างกันไม่มากนัก หรือมีค่าใกล้เคียงกัน ซึ่งสอดคล้องกับงานวิจัยของพรโสภา พินิจผล

(2539 : 67) ที่ได้ทำการความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน 2 วิธี คือ การเทียบคะแนนจริงและการเทียบคะแนนดิบ พบว่าการเทียบคะแนนทั้งสองวิธีมีคะแนนความคลาดเคลื่อนมาตรฐานไม่แตกต่างกัน

3. การเปรียบเทียบความแตกต่างของความเพียงพอที่ได้จากการเทียบคะแนนตามวิธีการเทียบคะแนนทั้ง 3 วิธี มีค่าแตกต่างกัน ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะว่า การเทียบคะแนนทั้ง 3 วิธีนั้น มีแบบแผนการเก็บรวบรวมข้อมูลที่แตกต่างกัน ซึ่งสอดคล้องกับผลการศึกษาของโคเลน และวิทนี (Kolen & Whitney, 1982 : 291) พบว่า การศึกษาจะให้ผลของความเพียงพอได้มักต้องอาศัยปัจจัยหลายอย่างร่วมกัน ทั้งลักษณะของแบบทดสอบ แบบแผนของการเก็บรวบรวมข้อมูล และขนาดของกลุ่มตัวอย่าง ซึ่งจะเห็นได้ว่า แบบแผนการเก็บรวบรวมข้อมูลแบบที่ 2 ให้ผลความเพียงพออยู่ในเกณฑ์ที่น่าพอใจเพียงวิธีเดียว นั่นคือรูปแบบการเทียบคะแนนที่นำมาใช้นั้น มีความเหมาะสมและเพียงพอ

ข้อเสนอแนะ

จากการศึกษาค้นคว้าครั้งนี้ ทำให้ผู้วิจัยได้แนวคิดที่จะเสนอแนะดังนี้

1. ควรมีการศึกษาเปรียบเทียบผลการเทียบคะแนนเชิงเส้นตรง โดยใช้แบบทดสอบอื่นๆ เช่น แบบทดสอบวัดความถนัด แทนแบบทดสอบวัดผลสัมฤทธิ์บ้าง ว่าจะให้ผลการเทียบที่แตกต่างหรือไม่
2. ใช้แบบแผนการเก็บรวบรวมข้อมูลแบบอื่นมาเปรียบเทียบกัน เพื่อหาวิธีที่ให้ความคลาดเคลื่อนมาตรฐานที่น้อยที่สุด และมีความเพียงพออยู่ในเกณฑ์ที่น่าพอใจมากที่สุด
3. ควรมีการศึกษาเปรียบเทียบผลการเทียบคะแนนเชิงเส้นตรง โดยใช้กลุ่มตัวอย่างที่มีความสามารถแตกต่างกันบ้าง
4. ควรทำการศึกษาขนาดของกลุ่มตัวอย่างที่มีความเหมาะสมในการเทียบคะแนนเชิงเส้นตรงว่าขนาดของกลุ่มตัวอย่างที่เล็กที่สุดควรมีจำนวนเท่าใด เพื่อประหยัดงบประมาณ ค่าใช้จ่าย และเวลาในการนำผลที่ได้จากการเทียบคะแนนเชิงเส้นตรงมาใช้ได้อย่างมีประสิทธิภาพและประสิทธิผล

บรรณานุกรม

บรรณานุกรม

- ชูชีพ พงษ์สมบุญ. การเปรียบเทียบการเทียบมาตราระหว่างรูปแบบที่ใช้ผู้สอบร่วมกับรูปแบบที่ใช้แบบสอบรวม. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2528.
- ชูศักดิ์ ชัมภลจิต. Test Equating Procedures. (เอกสารประกอบการบรรยาย). ม.ป.ท.
- นิคม ดังคะพิภพ. Equating. (เอกสารประกอบการเรียน). กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, ม.ป.ป.
- นิภาพร ยอดเมือง. ความคลาดเคลื่อน และความลำเอียงทางสถิติในการเทียบคะแนนเชิงเส้นตรง แบบคะแนนจริงสัมพันธ์ในกลุ่มตัวอย่างขนาดเล็ก. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2539.
- พรโสภา พินิจผล. ผลการเทียบคะแนนแบบทดสอบวิชาการศึกษาตามทฤษฎีการตอบข้อสอบที่ใช้วิธีวิเคราะห์ต่างกัน. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2539.
- ภาวิณี ศรีสุขวัฒนานนท์. การเปรียบเทียบผลจากการใช้รูปแบบการเทียบมาตราที่ต่างกันเมื่อแบบสอบร่วมมีความยาวต่างกัน. วิทยานิพนธ์ ค.ด. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2529.
- เรวดี อินทสระ. การเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตราระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์ห้อยค์ประกอบ. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2530.
- ล้วน สายยศ และอังคณา สายยศ. เทคนิคการวิจัยทางการศึกษา. กรุงเทพฯ : สุวีริยาสาส์น, 2538.
- _____. สถิติวิทยาทางการวิจัย. กรุงเทพฯ : สุวีริยาสาส์น, 2540.
- วรเนตร พิชิตเกริกพล. การเปรียบเทียบคุณภาพของวิธีการเทียบมาตราระหว่างแบบอิกคิวเปอร์เซ็นไทล์กับแบบเชิงเส้นตรง. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2536.
- วิรัช วรรณรัตน์. "การเทียบมาตรา (Equating)," วารสารวัดผลการศึกษา. 11(29) : กันยายน – ธันวาคม, 69-73, 2530.
- สงบ ลักษณะ. "การเทียบคะแนนระหว่างแบบทดสอบ (Equating)," วารสารวัดผลการศึกษา. 4(2) : กันยายน – ธันวาคม, 21-32, 2525.
- _____. "ความสามารถเกี่ยวกับการตอบข้อสอบ," วารสารวัดผลการศึกษา. 4 : พฤษภาคม – สิงหาคม, 47-54, 2528.
- สุมนา ไสถิตผลอนันต์. การศึกษาวิธีการเทียบคะแนนแบบเส้นตรงตามแบบจำลองคะแนนจริงสัมพันธ์. วิทยานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2539.
- Angoff, W.H. "Scales, Norms and Equivalent Scores," In R.L. Thorndike (Ed.) Educational Measurement. 2nd ed. Washington, D.C. : American Council on Educational, 508-600; 1971.
- _____. "Summer and Derivation of Equating Methods Used at ETS," In Test Equating. P.55-70. New York : Academic Press, 1982.
- Brennan, R.L. & Kolen, M.J. "Some Practical Issues in Equating". Applied Psychological Measurement. 11(September) : p. 279 – 290, 1987.
- Bruan, Henry & Holland, P.W. "Observed-score Test Equating : Mathematical Analysis of Some ETS Equating Procedures," Test Equating. P.9-50. New York : Academic Press, 1982.

- Cook, L.L. & Petersen, N.S. "Problem Related to the Use of Conventional and Item Response Theory Equating Methods in Less Than Optimal Circumstances," Applied Psychology Measurement. 11(3) : 225-244, 1987.
- Cope, R.T. "How Well Do the Angoff Design V Linear Equating Methods Compare With the Tucker and Levine Methods," Applied Psychology Measurement. 11(2) : 143 – 149, 1987.
- Flanagan, T.C. "Unit, Scores and Norm," Educational Measurement. P.695-763. Washington D.C. : American Council on Education, 1951.
- Gulliksen, Harold. Theory of Mental Tests. New York : John Wiley & Sons, Inc., 1950.
- Hambleton, R.K. & H. Swaminathan. "Development in Latent Trait Theory Model Technical Issues and Applications," Review of Educational Research. P.465-510. 48, Fall, 1978.
- _____. Item Response Theory. Kluwer. Nijhoff, Boston, 312, 1985.
- Holland, Paul W. & Rubin Donald B. Test Equating. New York, Academic Press, 1982.
- Jaeger, R.M. "Some Exploratory Indices for Selection of a Test Equating Method," Journal of Educational Measurement. 18: 23-38, Spring, 1981.
- Kolen, M.J. "Comparison of Traditional and Item Response Theory Methods for Equating Tests," Journal of Educational Measurement. 18(1) : 1-11, Spring, 1981.
- Kolen, M.J. & Whitney, D.R. "Comparison of Four Procedures for Equating the Tests of General Educational Development," Journal of Educational Measurement. 19 : 279-293, Winter, 1982.
- Lord, F.M. "Equating Test Scores-A Maximum Likelihood Solution," Psychometrika. 20 : 193-200, 1955a.
- _____. "Practical Applications of Item Characteristic Curve Theory," Journal of Educational Measurement. 14 : 117-138, Summer, 1977.
- _____. Applications of Item Response Theory to Practical Testing Problem. Hillsdale, N.J. : Erlbaum, 1980.
- _____. "Standard Error of an Equating by Item Response Theory," Educational Testing Service, Princeton, New Jersey : ERIC Document Reproduction Service, November, 1981.
- _____. "The Standard Error of an Equipercentile Equating," Journal of Educational Statistics. 7(3) : 165-174, 1982.
- Lord, F.M. & Novick, M.R. Statistical Theories of Mental Test Scores. Reading, M.A. : Addison-Wesley 1968.
- Macro, Cary L. "Equating Tests in an Era of Test Disclosure," New Directions for Testing and Measurement. 8 : 105-122, September, 1981.
- Parshall, C.G. , Houghton, P.B. & Kromley, J.D. "Equating Error and Statistics Bias in Small Sample Linear Equating," Journal of Educational Measurement. 32(1) : 37-54, Spring, 1995.
- Petersen, N.S., Cook, L.L. & Stocking, M.L. "IRT Versus Conventional Equating Methods : A Comparative Study of Scale Stability," Journal of Educational Statistics. 8 : 137 – 156, 1983.
- Petersen, N.S., Kolen, M.J. & Hoover, H.D. "Scaling, Norming and Equating," Educational Measurement. P.221-262. 3rd ed. New York : Macmillan, 1983.

- Petersen, N.S., Macro, G.L. & Stewart, E.E. "A Test of the Adequacy of Linear Score Equating Methods," Test Equating. P.71-135. New York : Academic Press, 1982.
- Petersen & others. "Scaling, Norming and Equating," In Educational Measurement. Edited by R.L.Linn, p.241-260. New York : American Council on Education and Macmillan Publishing Company, 1989.
- Rock, D.A. "Equating Using the Confirmatory Factor Analysis Model," Test Equating. P.251-254, ed. By P.W. Holland and D.B. Rubin. New York, Academic Press, 1982.
- Woodruff, D. "Derivations of Observed Score Linear Equating Methods Base on Test Scores Models for the Common Item nonequivalent Populations Design," Journal of Educational Statistics. 11 (4) : 254-257, 1986.

ภาคผนวก

ภาคผนวก ก

รายชื่อผู้เชี่ยวชาญที่ตรวจสอบคุณภาพเครื่องมือ

รายชื่อผู้เชี่ยวชาญที่ตรวจสอบคุณภาพเครื่องมือ

- | | |
|---|--|
| 1. รองศาสตราจารย์ ดร.ส.วาสนา ประवालพุกษ์ | อาจารย์คณะศึกษาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ |
| 2. ผู้ช่วยศาสตราจารย์ระวีวรรณ พันธุ์พานิช | อาจารย์คณะศึกษาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ |
| 3. อาจารย์ชวลิต รวยอาจิณ | อาจารย์คณะศึกษาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ |
| 4. อาจารย์จิตติพร ทองสุข | ศึกษานิเทศก์ 7
สำนักงานการประถมศึกษาจังหวัดปราจีนบุรี |
| 5. อาจารย์สมพร นวลอุทัย | อาจารย์ 2 ระดับ 7
โรงเรียนอนุบาลประจันตคาม
จังหวัดปราจีนบุรี |

ภาคผนวก ข
ตารางวิเคราะห์หลักสูตรวิชาคณิตศาสตร์

ตารางวิเคราะห์หลักสูตรวิชาคณิตศาสตร์
เรื่อง สมการและการแก้สมการ
ชั้นประถมศึกษาปีที่ 6

เนื้อหา	พฤติกรรม (จำนวนข้อ(%))			รวม
	ความจำ	ความเข้าใจ	การนำไปใช้	
1. ความหมายของสมการ	5 (5%)	5 (5%)	0 (0%)	10 (10%)
2. สมการที่เป็นจริง	7 (7%)	8 (8%)	0 (0%)	15 (15%)
3. สมการที่มีตัวไม่ทราบค่า	8 (8%)	7 (7%)	0 (0%)	15 (15%)
4. คำตอบของสมการ	5 (5%)	5 (5%)	20 (20%)	30 (30%)
5. การแก้สมการ	5 (5%)	5 (5%)	20 (20%)	30 (30%)
รวม	30 (30%)	30 (30%)	40 (40%)	100 (100%)

ภาคผนวก ค
คุณภาพของแบบทดสอบจากการทดลองเครื่องมือ

คุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์
เรื่อง สมการและการแก้สมการ
ชั้นประถมศึกษาปีที่ 6

ข้อที่	ตัวเลือก	P	$R_{p.BIS}$	Δ	การพิจารณา
1	ก	0.045	0.230	19.82	ตัดทิ้ง
	ข	0.055	0.222	19.43	
	ค*	0.840	0.382	8.98	
	ง	0.055	0.152	19.43	
	จ	0.005	0.097	23.34	
2	ก*	0.780	0.354	9.87	คัดเลือกไว้
	ข	0.100	0.227	18.17	
	ค	0.035	0.194	20.29	
	ง	0.030	0.228	20.56	
	จ	0.055	0.018	19.43	
3	ก	0.155	0.085	17.10	คัดเลือกไว้
	ข	0.115	0.196	17.84	
	ค	0.130	0.151	17.54	
	ง*	0.530	0.396	12.66	
	จ	0.070	0.209	18.94	
4	ก	0.045	0.278	19.82	คัดเลือกไว้
	ข	0.090	0.072	18.40	
	ค	0.055	0.178	19.43	
	ง*	0.775	0.244	9.94	
	จ	0.035	0.130	20.29	
5	ก	0.080	0.169	18.66	คัดเลือกไว้
	ข	0.090	0.165	18.40	
	ค	0.055	0.197	19.43	
	ง	0.050	0.238	19.62	
	จ*	0.725	0.426	10.57	
6	ก	0.080	0.224	18.66	คัดเลือกไว้
	ข	0.075	0.239	18.81	
	ค	0.125	0.218	17.64	
	ง	0.050	0.074	19.62	
	จ*	0.670	0.451	11.20	

ข้อที่	ตัวเลือก	P	$R_{P.BIS}$	Δ	การพิจารณา
7	ก	0.100	0.124	18.17	คัดเลือกไว้
	ข*	0.585	0.431	12.10	
	ค	0.140	0.170	17.36	
	ง	0.120	0.182	17.74	
	จ	0.055	0.249	19.43	
8	ก	0.085	0.318	18.53	คัดเลือกไว้
	ข	0.080	0.320	18.66	
	ค	0.235	0.087	15.93	
	ง*	0.5255	0.315	12.71	
	จ	0.075	0.070	18.80	
9	ก	0.075	0.205	18.80	คัดเลือกไว้
	ข*	0.575	0.274	12.20	
	ค	0.035	0.259	20.29	
	ง	0.035	0.164	20.29	
	จ	0.280	0.009	15.37	
10	ก	0.045	0.168	19.82	คัดเลือกไว้
	ข	0.070	0.279	18.94	
	ค	0.180	0.223	16.70	
	ง	0.165	0.184	16.94	
	จ*	0.540	0.522	12.56	
11	ก	0.080	0.177	18.66	คัดเลือกไว้
	ข*	0.780	0.376	9.87	
	ค	0.075	0.141	18.80	
	ง	0.045	0.207	19.82	
	จ	0.020	0.198	21.25	
12	ก*	0.550	0.191	12.46	ตัดทิ้ง
	ข	0.070	0.239	18.94	
	ค	0.065	0.147	19.09	
	ง	0.045	0.153	18.92	
	จ	0.270	0.077	15.49	
13	ก	0.125	0.174	17.64	คัดเลือกไว้
	ข*	0.595	0.424	12.00	
	ค	0.145	0.142	17.54	
	ง	0.065	0.122	19.09	
	จ	0.070	0.278	18.94	

ข้อที่	ตัวเลือก	P	R _{P.BIS}	Δ	การพิจารณา
14	ก	0.265	0.044	15.55	คัดเลือกไว้
	ข	0.155	0.218	17.10	
	ค*	0.405	0.258	17.10	
	ง	0.110	0.065	17.94	
	จ	0.065	0.190	19.09	
15	ก*	0.670	0.323	11.20	คัดเลือกไว้
	ข	0.095	0.167	18.28	
	ค	0.070	0.177	18.94	
	ง	0.070	0.146	18.94	
	จ	0.095	0.071	18.28	
16	ก	0.090	0.213	18.40	คัดเลือกไว้
	ข	0.080	0.268	18.66	
	ค	0.080	0.230	18.66	
	ง*	0.655	0.461	11.36	
	จ	0.095	0.079	18.28	
17	ก*	0.360	0.393	14.47	คัดเลือกไว้
	ข	0.155	0.254	17.10	
	ค	0.155	0.192	17.10	
	ง	0.205	0.002	16.33	
	จ	0.125	0.085	17.64	
18	ก*	0.280	0.163	15.37	ตัดทิ้ง
	ข	0.170	0.100	16.86	
	ค	0.165	0.212	16.95	
	ง	0.195	0.001	16.48	
	จ	0.190	0.176	16.55	
19	ก	0.115	0.125	17.84	ตัดทิ้ง
	ข*	0.365	0.191	14.42	
	ค	0.370	0.080	14.37	
	ง	0.085	0.170	18.53	
	จ	0.065	0.175	19.09	
20	ก	0.080	0.232	18.66	คัดเลือกไว้
	ข	0.100	0.220	18.17	
	ค*	0.635	0.412	11.58	
	ง	0.100	0.219	18.17	
	จ	0.085	0.110	18.53	

ข้อที่	ตัวเลือก	P	R _{P.BIS}	Δ	การพิจารณา
21	ก	0.305	0.007	14.42	คัดเลือกไว้
	ข*	0.380	0.312	14.26	
	ค	0.120	0.141	17.74	
	ง	0.065	0.245	19.09	
	จ	0.070	0.163	18.94	
22	ก	0.080	0.272	18.66	คัดเลือกไว้
	ข	0.120	0.134	17.74	
	ค*	0.425	0.304	16.80	
	ง	0.185	0.033	16.63	
	จ	0.190	0.050	16.55	
23	ก	0.210	0.099	16.27	ตัดทิ้ง
	ข	0.095	0.224	18.28	
	ค	0.340	0.114	14.69	
	ง	0.195	0.135	16.48	
	จ*	0.160	0.067	17.02	
24	ก	0.410	0.373	13.95	ตัดทิ้ง
	ข	0.195	0.125	16.48	
	ค*	0.175	0.012	16.78	
	ง	0.135	0.189	17.45	
	จ	0.085	0.232	18.53	
25	ก	0.130	0.165	17.54	คัดเลือกไว้
	ข*	0.455	0.358	13.49	
	ค	0.165	0.172	16.94	
	ง	0.140	0.021	17.36	
	จ	0.110	0.165	17.92	
26	ก*	0.855	0.419	8.46	ตัดทิ้ง
	ข	0.060	0.284	19.21	
	ค	0.035	0.263	20.29	
	ง	0.025	0.195	20.88	
	จ	0.025	0.007	20.88	
27	ก	0.030	0.157	20.56	ตัดทิ้ง
	ข*	0.850	0.396	8.82	
	ค	0.065	0.275	19.09	
	ง	0.045	0.201	19.82	
	จ	0.010	0.053	22.34	

ข้อที่	ตัวเลือก	P	$R_{P,BIS}$	Δ	การพิจารณา
28	ก	0.060	0.345	19.26	ตัดทิ้ง
	ข	0.060	0.182	19.26	
	ค*	0.840	0.458	8.98	
	ง	0.030	0.157	20.56	
	จ	0.010	0.161	22.34	
29	ก	0.040	0.239	20.04	ตัดทิ้ง
	ข*	0.835	0.453	9.06	
	ค	0.040	0.233	20.07	
	ง	0.030	0.160	20.56	
	จ	0.055	0.213	19.43	
30	ก	0.015	0.118	21.68	คัดเลือกไว้
	ข	0.070	0.250	18.94	
	ค	0.075	0.288	18.80	
	ง	0.045	0.144	19.82	
	จ*	0.795	0.455	9.67	
31	ก*	0.800	0.488	9.59	คัดเลือกไว้
	ข	0.040	0.205	20.04	
	ค	0.070	0.274	18.94	
	ง	0.040	0.233	20.04	
	จ	0.050	0.181	19.62	
32	ก	0.045	0.212	19.82	คัดเลือกไว้
	ข	0.065	0.217	19.09	
	ค*	0.810	0.415	9.45	
	ง	0.060	0.183	19.26	
	จ	0.020	0.156	21.25	
33	ก	0.050	0.188	19.62	คัดเลือกไว้
	ข	0.055	0.236	19.43	
	ค	0.060	0.271	19.26	
	ง*	0.760	0.461	10.14	
	จ	0.075	0.120	18.80	
34	ก	0.150	0.301	17.18	คัดเลือกไว้
	ข	0.080	0.198	18.66	
	ค*	0.650	0.520	11.42	
	ง	0.075	0.128	18.80	
	จ	0.045	0.256	19.82	

ข้อที่	ตัวเลือก	P	R _{P.BIS}	Δ	การพิจารณา
35	ก	0.100	0.261	18.17	คัดเลือกไว้
	ข	0.075	0.347	18.81	
	ค	0.080	0.172	18.06	
	ง	0.060	0.175	19.26	
	จ*	0.685	0.556	11.03	
36	ก	0.125	0.334	17.64	คัดเลือกไว้
	ข	0.070	0.241	18.94	
	ค	0.045	0.205	19.82	
	ง	0.070	0.263	18.94	
	จ*	0.685	0.556	11.03	
37	ก	0.030	0.260	20.56	คัดเลือกไว้
	ข	0.135	0.294	17.45	
	ค	0.055	0.296	19.43	
	ง	0.080	0.211	18.66	
	จ*	0.700	0.588	10.86	
38	ก	0.095	0.269	18.28	คัดเลือกไว้
	ข	0.065	0.251	19.09	
	ค	0.075	0.266	18.80	
	ง	0.060	0.244	19.26	
	จ*	0.705	0.590	10.81	
39	ก	0.145	0.311	17.54	คัดเลือกไว้
	ข	0.025	0.220	20.88	
	ค	0.060	0.248	19.26	
	ง	0.040	0.179	20.04	
	จ*	0.730	0.536	10.51	
40	ก	0.065	0.259	19.09	คัดเลือกไว้
	ข	0.060	0.268	19.26	
	ค	0.065	0.238	19.09	
	ง	0.100	0.238	18.17	
	จ*	0.710	0.568	10.75	
41	ก	0.090	0.356	18.40	คัดเลือกไว้
	ข	0.070	0.217	18.94	
	ค*	0.620	0.365	11.74	
	ง	0.145	0.029	17.54	
	จ	0.075	0.038	18.80	

ข้อที่	ตัวเลือก	P	R _{P.BIS}	Δ	การพิจารณา
42	ก	0.065	0.228	19.09	คัดเลือกไว้
	ข*	0.630	0.506	11.63	
	ค	0.100	0.258	18.17	
	ง	0.115	0.250	17.84	
	จ	0.090	0.109	18.40	
43	ก	0.115	0.337	17.84	คัดเลือกไว้
	ข	0.055	0.207	19.43	
	ค	0.055	0.213	19.43	
	ง	0.105	0.080	18.05	
	จ*	0.670	0.485	11.20	
44	ก	0.085	0.252	18.53	คัดเลือกไว้
	ข	0.155	0.257	17.84	
	ค*	0.655	0.566	11.36	
	ง	0.085	0.155	18.53	
	จ	0.060	0.310	19.26	
45	ก	0.070	0.197	18.94	คัดเลือกไว้
	ข*	0.780	0.494	9.87	
	ค	0.065	0.262	19.09	
	ง	0.045	0.166	19.82	
	จ	0.040	0.283	20.04	
46	ก	0.080	0.195	18.66	คัดเลือกไว้
	ข	0.125	0.228	17.64	
	ค*	0.470	0.464	13.34	
	ง	0.060	0.193	19.26	
	จ	0.265	0.130	15.55	
47	ก	0.220	0.165	16.13	คัดเลือกไว้
	ข	0.130	0.386	17.54	
	ค*	0.465	0.424	13.39	
	ง	0.145	0.011	17.54	
	จ	0.040	0.088	20.04	
48	ก	0.130	0.098	17.54	คัดเลือกไว้
	ข	0.070	0.197	18.94	
	ค	0.085	0.238	18.53	
	ง*	0.435	0.367	13.69	
	จ	0.280	0.072	15.37	

ข้อที่	ตัวเลือก	P	R _{P.BIS}	Δ	การพิจารณา
49	ก	0.205	0.104	16.33	คัดเลือกไว้
	ข	0.080	0.196	18.66	
	ค	0.095	0.184	18.28	
	ง*	0.460	0.421	13.44	
	จ	0.160	0.165	17.02	
50	ก	0.155	0.149	17.10	คัดเลือกไว้
	ข*	0.525	0.420	12.71	
	ค	0.085	0.224	18.53	
	ง	0.125	0.079	17.94	
	จ	0.110	0.208	17.64	
51	ก	0.085	0.236	18.53	ตัดทิ้ง
	ข	0.360	0.120	14.47	
	ค	0.350	0.192	14.58	
	ง*	0.185	0.305	16.63	
	จ	0.020	0.131	21.25	
52	ก	0.050	0.189	19.62	คัดเลือกไว้
	ข	0.105	0.207	18.05	
	ค*	0.590	0.537	12.05	
	ง	0.165	0.284	16.94	
	จ	0.090	0.190	18.40	
53	ก	0.100	0.265	18.17	คัดเลือกไว้
	ข*	0.505	0.536	12.96	
	ค	0.175	0.129	16.78	
	ง	0.130	0.220	17.54	
	จ	0.090	0.229	18.40	
54	ก	0.085	0.167	18.53	คัดเลือกไว้
	ข*	0.425	0.345	13.80	
	ค	0.280	0.013	15.37	
	ง	0.095	0.176	18.28	
	จ	0.115	0.208	17.84	
55	ก	0.115	0.164	17.84	ตัดทิ้ง
	ข	0.095	0.171	18.28	
	ค*	0.115	0.207	17.84	
	ง	0.505	0.473	12.96	
	จ	0.170	0.180	16.86	

ข้อที่	ตัวเลือก	P	R _{p.BIS}	Δ	การพิจารณา
56	ก	0.180	0.087	16.70	คัดเลือกไว้
	ข	0.170	0.228	16.86	
	ค	0.090	0.150	18.40	
	ง*	0.505	0.392	12.96	
	จ	0.055	0.150	19.43	
57	ก	0.055	0.188	19.43	คัดเลือกไว้
	ข*	0.555	0.534	12.41	
	ค	0.085	0.200	18.53	
	ง	0.200	0.303	16.41	
	จ	0.105	0.148	18.05	
58	ก	0.060	0.156	19.26	คัดเลือกไว้
	ข	0.290	0.024	15.25	
	ค	0.105	0.287	18.05	
	ง*	0.475	0.293	13.29	
	จ	0.070	0.126	18.94	
59	ก	0.145	0.077	17.54	คัดเลือกไว้
	ข	0.145	0.267	17.54	
	ค	0.110	0.273	17.94	
	ง*	0.545	0.475	12.51	
	จ	0.055	0.132	19.43	
60	ก*	0.650	0.594	11.42	คัดเลือกไว้
	ข	0.100	0.287	18.17	
	ค	0.100	0.213	18.17	
	ง	0.110	0.314	17.94	
	จ	0.040	0.178	20.04	
61	ก*	0.505	0.194	12.96	คัดเลือกไว้
	ข	0.150	0.113	17.18	
	ค	0.185	0.359	16.63	
	ง	0.055	0.206	19.43	
	จ	0.105	0.017	18.05	
62	ก*	0.580	0.583	12.15	คัดเลือกไว้
	ข	0.115	0.236	17.84	
	ค	0.070	0.259	18.94	
	ง	0.190	0.283	16.55	
	จ	0.045	0.152	19.82	

ข้อที่	ตัวเลือก	P	$R_{P,BIS}$	Δ	การพิจารณา
63	ก	0.110	0.157	17.94	คัดเลือกไว้
	ข*	0.580	0.570	12.15	
	ค	0.065	0.252	19.09	
	ง	0.065	0.214	19.09	
	จ	0.180	0.306	16.70	
64	ก	0.155	0.036	17.10	คัดเลือกไว้
	ข	0.055	0.228	19.43	
	ค	0.090	0.188	18.40	
	ง	0.175	0.248	16.78	
	จ*	0.525	0.375	12.71	
65	ก	0.080	0.173	18.66	คัดเลือกไว้
	ข*	0.530	0.437	12.66	
	ค	0.210	0.037	16.27	
	ง	0.130	0.378	17.54	
	จ	0.050	0.138	19.62	
66	ก	0.140	0.212	17.36	ตัดทิ้ง
	ข*	0.295	0.017	15.19	
	ค	0.320	0.251	14.91	
	ง	0.115	0.080	17.84	
	จ	0.130	0.078	17.54	
67	ก	0.160	0.071	17.02	คัดเลือกไว้
	ข	0.050	0.207	19.62	
	ค	0.095	0.266	18.28	
	ง*	0.630	0.333	11.63	
	จ	0.065	0.047	19.09	
68	ก	0.135	0.236	17.45	คัดเลือกไว้
	ข	0.145	0.241	17.54	
	ค	0.105	0.220	18.05	
	ง*	0.565	0.561	12.31	
	จ	0.050	0.207	19.62	
69	ก	0.060	0.194	19.26	คัดเลือกไว้
	ข*	0.680	0.488	11.09	
	ค	0.115	0.187	17.84	
	ง	0.095	0.309	18.28	
	จ	0.050	0.145	19.62	

ข้อที่	ตัวเลือก	P	R _{P.BIS}	Δ	การพิจารณา
70	ก	0.110	0.326	17.94	คัดเลือกไว้
	ข	0.145	0.265	17.54	
	ค*	0.610	0.506	11.84	
	ง	0.060	0.019	19.26	
	จ	0.075	0.220	18.80	
71	ก	0.090	0.074	18.40	คัดเลือกไว้
	ข	0.265	0.242	15.55	
	ค*	0.415	0.357	13.96	
	ง	0.155	0.142	17.10	
	จ	0.075	0.014	18.80	
72	ก	0.205	0.000	16.33	คัดเลือกไว้
	ข*	0.365	0.281	14.42	
	ค	0.110	0.171	17.94	
	ง	0.125	0.053	17.64	
	จ	0.195	0.163	16.48	
73	ก	0.120	0.097	17.74	คัดเลือกไว้
	ข	0.205	0.242	16.33	
	ค*	0.380	0.383	14.26	
	ง	0.200	0.195	16.41	
	จ	0.095	0.071	18.28	
74	ก	0.065	0.199	19.09	คัดเลือกไว้
	ข*	0.465	0.398	13.39	
	ค	0.250	0.142	15.71	
	ง	0.135	0.116	17.45	
	จ	0.085	0.174	18.53	
75	ก	0.195	0.172	16.48	ตัดทิ้ง
	ข*	0.390	0.366	14.16	
	ค	0.165	0.146	16.94	
	ง	0.175	0.100	16.78	
	จ	0.075	0.068	18.80	
76	ก*	0.375	0.256	14.31	คัดเลือกไว้
	ข	0.225	0.044	16.06	
	ค	0.175	0.135	16.78	
	ง	0.125	0.060	17.64	
	จ	0.100	0.114	18.17	

ข้อที่	ตัวเลือก	P	R _{P.BIS}	Δ	การพิจารณา
77	ก	0.120	0.162	17.74	คัดเลือกไว้
	ข	0.265	0.032	15.55	
	ค	0.220	0.195	16.13	
	ง*	0.310	0.363	15.02	
	จ	0.085	0.072	18.53	
78	ก	0.155	0.149	17.10	คัดเลือกไว้
	ข	0.135	0.109	17.45	
	ค*	0.430	0.277	13.75	
	ง	0.125	0.046	17.64	
	จ	0.155	0.084	17.10	
79	ก	0.125	0.092	17.64	คัดเลือกไว้
	ข	0.205	0.089	16.33	
	ค*	0.410	0.269	13.95	
	ง	0.135	0.141	17.45	
	จ	0.125	0.053	17.64	
80	ก	0.185	0.164	16.63	คัดเลือกไว้
	ข	0.245	0.073	15.80	
	ค	0.175	0.125	16.78	
	ง*	0.315	0.343	14.97	
	จ	0.080	0.061	18.66	
81	ก	0.165	0.160	16.94	คัดเลือกไว้
	ข	0.185	0.187	16.63	
	ค*	0.370	0.365	14.37	
	ง	0.165	0.028	16.94	
	จ	0.115	0.107	17.84	
82	ก	0.145	0.018	17.54	คัดเลือกไว้
	ข	0.180	0.228	16.70	
	ค	0.235	0.268	15.93	
	ง	0.165	0.025	16.94	
	จ*	0.275	0.444	15.43	
83	ก*	0.295	0.398	15.19	คัดเลือกไว้
	ข	0.230	0.218	15.99	
	ค	0.155	0.132	17.10	
	ง	0.165	0.033	16.94	
	จ	0.155	0.082	17.10	

ข้อที่	ตัวเลือก	P	R _{P.BIS}	Δ	การพิจารณา
84	ก	0.210	0.114	16.27	คัดเลือกไว้
	ข	0.155	0.080	17.10	
	ค*	0.310	0.381	15.02	
	ง	0.225	0.167	16.06	
	จ	0.100	0.103	18.17	
85	ก	0.340	0.156	14.69	คัดเลือกไว้
	ข	0.160	0.193	17.02	
	ค	0.125	0.087	17.64	
	ง*	0.225	0.479	16.06	
	จ	0.150	0.074	17.18	
86	ก	0.145	0.043	17.54	คัดเลือกไว้
	ข	0.155	0.131	17.10	
	ค	0.240	0.118	15.86	
	ง	0.160	0.079	17.02	
	จ*	0.300	0.309	15.14	
87	ก	0.130	0.048	17.54	คัดเลือกไว้
	ข*	0.365	0.315	14.42	
	ค	0.190	0.213	16.55	
	ง	0.135	0.140	17.45	
	จ	0.180	0.011	16.70	
88	ก	0.160	0.191	17.02	คัดเลือกไว้
	ข	0.265	0.022	15.55	
	ค	0.155	0.023	17.10	
	ง*	0.310	0.327	15.02	
	จ	0.110	0.202	17.94	
89	ก*	0.145	0.135	17.54	ตัดทิ้ง
	ข	0.230	0.070	15.99	
	ค	0.380	0.177	14.26	
	ง	0.145	0.027	17.54	
	จ	0.100	0.060	18.17	
90	ก	0.160	0.119	17.02	คัดเลือกไว้
	ข	0.420	0.416	13.85	
	ค	0.145	0.231	17.54	
	ง	0.125	0.125	17.64	
	จ*	0.150	0.109	17.18	

ข้อที่	ตัวเลือก	P	R _{P.BIS}	Δ	การพิจารณา
91	ก*	0.445	0.424	13.59	คัดเลือกไว้
	ข	0.190	0.113	16.55	
	ค	0.195	0.254	16.48	
	ง	0.115	0.131	17.84	
	จ	0.055	0.106	19.43	
92	ก	0.170	0.045	16.86	คัดเลือกไว้
	ข*	0.410	0.281	13.95	
	ค	0.140	0.225	17.36	
	ง	0.180	0.060	16.70	
	จ	0.100	0.067	18.17	
93	ก	0.160	0.227	17.02	คัดเลือกไว้
	ข	0.240	0.089	15.86	
	ค*	0.375	0.332	14.31	
	ง	0.100	0.134	18.71	
	จ	0.125	0.001	17.64	
94	ก	0.095	0.194	18.28	คัดเลือกไว้
	ข*	0.500	0.461	13.00	
	ค	0.215	0.145	16.20	
	ง	0.085	0.093	18.53	
	จ	0.105	0.288	18.05	
95	ก*	0.380	0.315	14.26	คัดเลือกไว้
	ข	0.195	0.065	16.48	
	ค	0.175	0.201	16.76	
	ง	0.170	0.067	16.86	
	จ	0.080	0.094	18.66	
96	ก	0.235	0.007	15.93	คัดเลือกไว้
	ข	0.170	0.094	16.86	
	ค	0.310	0.071	15.02	
	ง	0.100	0.077	18.17	
	จ*	0.185	0.059	16.63	
97	ก	0.295	0.098	15.19	คัดเลือกไว้
	ข	0.215	0.146	16.20	
	ค	0.175	0.177	16.78	
	ง	0.105	0.233	18.05	
	จ*	0.210	0.303	16.27	

ข้อที่	ตัวเลือก	P	$R_{P.BIS}$	Δ	การพิจารณา
98	ก	0.170	0.155	16.86	คัดเลือกไว้
	ข*	0.415	0.237	13.90	
	ค	0.185	0.038	16.63	
	ง	0.167	0.036	17.10	
	จ	0.111	0.121	18.86	
99	ก	0.165	0.030	16.94	ตัดทิ้ง
	ข	0.240	0.098	15.86	
	ค	0.205	0.204	16.37	
	ง	0.200	0.067	16.41	
	จ*	0.190	0.219	16.55	
100	ก	0.225	0.203	16.06	คัดเลือกไว้
	ข*	0.430	0.394	13.71	
	ค	0.130	0.043	17.54	
	ง	0.120	0.061	17.74	
	จ	0.095	0.258	18.28	

ภาคผนวก ง

แบบทดสอบที่ใช้ในการเทียบคะแนน 3 ฉบับ

ฉบับที่ 1 จำนวน 20 ข้อ

ฉบับที่ 2 จำนวน 20 ข้อ

ฉบับที่ 3 (ฉบับรวม) จำนวน 10 ข้อ

แบบทดสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ฉบับที่ 1
เรื่อง สมการและการแก้สมการ

1. ข้อใดเป็นสมการ

- ก. $\frac{4}{5} - \frac{1}{5} = \frac{3}{5}$
 ข. $8 - 4 \neq 4$
 ค. $\frac{3}{4} \times 12 > 5$
 ง. $13 \times 4 < 134$
 จ. $10 \times \frac{1}{3} > 30$

2. ข้อใดไม่ใช่สมการ

- ก. $14 + 8 = 22$
 ข. $30 \times 10 = 300$
 ค. $12 \div 6 = 2 \times 1$
 ง. $20 + 17 = 17 + 20$
 จ. $25 + 12 \neq 25 \times 12$

3. ข้อใดเป็นสมการที่เป็นจริง

- ก. $23 + 12 = \frac{121}{11}$
 ข. $234 - 34 = 100$
 ค. $14 \times 5 = 5 + 14$
 ง. $\frac{23}{7} \times 21 = 23 \times \frac{7}{21}$
 จ. $7 + 12.75 = 19.75$

4. ข้อใดเป็นสมการที่เป็นจริง

- ก. $360 \div 6 = 61$
 ข. $13 + 110 = 113$
 ค. $300 - 20 = 200$
 ง. $8 \times 500 = 4000$
 จ. $25 \times 10 = 2500$

5. ข้อใดเป็นสมการที่เป็นเท็จ

- ก. $10 \times (5 \times 4) = (10 \times 5) \times 4$
 ข. $10 + (5 + 4) = (10 + 5) + 4$
 ค. $10 - (5 - 4) = (10 - 5) - 4$
 ง. $10 + (5 + 4) = (5 + 10) + 4$
 จ. $10 \times (5 \times 4) = (5 \times 4) \times 10$

6. สมการในข้อใดมีตัวไม่ทราบค่า

- ก. $32 \times 11 = \square$
 ข. $35 \times 5 = 175$
 ค. $102 \div 2 = 51$
 ง. $50 + 91 = 141$
 จ. $5000 \div 1000 = 5$

7. สมการในข้อใดไม่มีตัวไม่ทราบค่า

- ก. $785 \div 5 = 5$
 ข. $152 \div 25 = 6$
 ค. $517 \times 6 = 1034$
 ง. $5 + 300 = 2400$
 จ. $1234 + 4321 = 5555$

8. สมการในข้อใดไม่มีตัวไม่ทราบค่า

- ก. $18 \div 5 = 9$
 ข. $40 - \square = 10$
 ค. $12 \times 6 = 36$
 ง. $32 + 12 = 5$
 จ. $200 \div 4 = 50$

9. จำนวนในข้อใดเป็นคำตอบของสมการ

$$320 = 5 + 230$$

- ก. 70
 ข. 80
 ค. 90
 ง. 100
 จ. 110

10. จำนวนในข้อใดเป็นคำตอบของสมการ

$$62 \times 6 = 434$$

- ก. 3
 ข. 4
 ค. 5
 ง. 6
 จ. 7

11. สมการในข้อใดที่แทนค่าด้วย 9 แล้วทำให้สมการนั้นเป็นจริง

- ก. $18 = x \times 9$
- ข. $99 = x + 100$
- ค. $92 = 828 \div x$
- ง. $109 = 100 - x$
- จ. $108 = 100 + x$

12. พ่อใช้เงิน จ บาท ซื้อกางเกง 10 ตัว ราคาตัวละ 45 บาท พ่อใช้เงินไปเท่าใด

- ก. 35 บาท
- ข. 55 บาท
- ค. 440 บาท
- ง. 450 บาท
- จ. 460 บาท

13. วันชัยมีเงิน 325 บาท แม่ให้อีก ง บาท วันชัยมีเงิน 650 บาท แม่ให้เงินวันชัยกี่บาท

- ก. 325 บาท
- ข. 335 บาท
- ค. 965 บาท
- ง. 975 บาท
- จ. 985 บาท

14. สมศรีมีเงาะ ค กอง กองละ 7 ผล เมื่อนับแล้วมีเงาะทั้งหมด 84 ผล สมศรีมีเงาะกี่กอง

- ก. 11 กอง
- ข. 12 กอง
- ค. 77 กอง
- ง. 91 กอง
- จ. 588 กอง

15. จะหาคำตอบของสมการ $x - 10 = 20$ ได้โดยวิธีใด

- ก. นำ 10 มาลบทั้งสองข้างของสมการ
- ข. นำ 10 มาบวกทั้งสองข้างของสมการ
- ค. นำ 10 มาหารทั้งสองข้างของสมการ
- ง. นำ 20 มาบวกทั้งสองข้างของสมการ
- จ. นำ 20 มาลบทั้งสองข้างของสมการ

16. จะแก้สมการ $124 + x = 248$ ได้ตามวิธีการในข้อใด

- ก. $124 + x - x = 248$
- ข. $124 + x + 124 = 248$
- ค. $124 + x - 124 = 248 - 124$
- ง. $124 \div 124 + x = 248 - 124$
- จ. $124 + x \div 124 = 248 \div 124$

17. $x - 9 = 56$ มีความหมายตรงกับข้อใด

- ก. อาริมีเงิน x บาท ได้อีก 9 บาท เป็นเงิน 56 บาท
- ข. อำไพมีส้มอยู่ 56 ผล ให้เพื่อนไป 9 ผล เหลือ x ผล
- ค. มีนักเรียน x คน ยืนเป็นแถวๆ ละ 9 คนจะยืนได้ทั้งหมด 56 แถว
- ง. ร้านค้ามียางลบ x อัน ขายไป 9 อัน เหลือยางลบทั้งหมด 56 อัน
- จ. มานะมีเงิน x บาท มาลัยมีเงิน 9 บาท รวมเป็นเงิน 56 บาท

18. $58 + 46 = k$ มีความหมายตรงกับข้อใด

- ก. ซื้อแก้วมา 58 ใบ ทำแตกไป 46 ใบ เหลือแก้ว k ใบ
- ข. พ่อซื้อทุเรียน 58 ผล แม่ซื้ออีก 46 ผล รวมเป็นทุเรียน k ผล
- ค. แม่ค้าขายนม 58 กล่อง ราคากล่องละ 46 บาท ได้เงินทั้งสิ้น k บาท
- ง. แก้วมีเงิน 58 บาท แม่มีเงิน 46 บาท แก้วมีเงินมากกว่าแม่อยู่เท่าไร
- จ. นักเรียนชาย 58 คน นักเรียนหญิง 46 คน มีนักเรียนชายมากกว่านักเรียนหญิงกี่คน

19. "นุชซื้อส้มมา 9 กิโลกรัม กิโลกรัมละ x บาท จ่ายเงินไป 63 บาท ส้มราคา กิโลกรัมละกี่บาท" เขียนเป็นสมการได้อย่างไร

- ก. $9 + x = 63$
- ข. $9 - x = 63$
- ค. $9 \times x = 63$
- ง. $9 \div x = 63$
- จ. $9 = 63 \div x$

20. “จงซื้อน้ำมันมา 5 โหล ราคาโหลละ ม บาท เป็นเงิน 1,050 บาท นมราคาโหลละเท่าใด” เขียนเป็นสมการได้อย่างไร

ก. $1050 \times 5 = \text{ม}$

ข. $5 \times \text{ม} = 1050$

ค. $\text{ม} \div 5 = 1050$

ง. $\text{ม} \div 1050 = 5$

จ. $1050 \times \text{ม} = 5$

แบบทดสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ฉบับที่ 2

เรื่อง สมการและการแก้สมการ

1. ข้อใดเป็นสมการ

- ก. $6 \times 7 < 10$
- ข. $7 + 8 > 15$
- ค. $14 - 5 \neq 8$
- ง. $13 + 12 = 25$
- จ. $13 + 12 \neq 15$

2. ข้อใดไม่ใช่สมการ

- ก. $7 \div \frac{21}{3} = 1$
- ข. $23 - 18 = \frac{25}{5}$
- ค. $\frac{3}{7} + \frac{4}{7} \neq \frac{7}{3} + \frac{7}{4}$
- ง. $\frac{25}{5} \times 3 = \frac{25 \times 3}{5}$
- จ. $25 \div 5 = \frac{25}{5}$

3. ข้อใดเป็นสมการที่เป็นจริง

- ก. $25 + 27 = 52$
- ข. $34 + 18 = 42$
- ค. $200 - 40 = 60$
- ง. $49 \div 7 = 7 \div 49$
- จ. $3 \times (4 + 5) = (3 \times 4) + 5$

4. ข้อใดเป็นสมการที่เป็นจริง

- ก. $11 + (2 \times 1) = (11 \times 2) + (11 \times 1)$
- ข. $11 \times (2 + 1) = (11 \times 2) + (11 \times 1)$
- ค. $11 \times (2 \times 1) = (11 + 2) + (11 + 1)$
- ง. $11 \times (2 \times 1) = (11 \times 2) \times (11 \times 1)$
- จ. $11 \times (2 + 1) = (11 + 2) \times (11 + 1)$

5. ข้อใดเป็นสมการที่เป็นเท็จ

- ก. $\frac{1}{5} \times \frac{2}{5} = \frac{2}{5} \times \frac{1}{5}$
- ข. $\frac{6}{9} \times \frac{9}{6} = \frac{9}{6} \times \frac{9}{6}$
- ค. $\frac{2}{5} \times \frac{7}{10} = \frac{7}{10} \times \frac{2}{5}$
- ง. $\frac{3}{10} \times \frac{7}{10} = \frac{7}{10} \times \frac{3}{10}$
- จ. $\frac{10}{20} \times \frac{5}{10} = \frac{5}{10} \times \frac{10}{20}$

6. สมการในข้อใดมีตัวไม่ทราบค่า

- ก. $500 \div 20 = 25$
- ข. $22 + 78 = 100$
- ค. $8 + 234 = 242$
- ง. $39 \times 100 = 3900$
- จ. $B + 2007 = 3007$

7. สมการในข้อใดไม่มีตัวไม่ทราบค่า

- ก. $526 = 50 + n$
- ข. $g \div 20 = 100$
- ค. $190 = x \div 15$
- ง. $240 + j = 2400$
- จ. $5000 \div 1000 = 5$

8. สมการในข้อใดไม่มีตัวไม่ทราบค่า

- ก. $n \div 8 = 7$
- ข. $72 \times m = 8$
- ค. $k - 17 = 17$
- ง. $25 + j = 42$
- จ. $99 - 18 = 81$

9. จำนวนในข้อใดเป็นคำตอบของสมการ

$$414 \div n = 69$$

- ก. 4
- ข. 6
- ค. 7
- ง. 8
- จ. 9

10. จำนวนในข้อใดเป็นคำตอบของสมการ

$$325 = x + 237$$

- ก. 12
- ข. 78
- ค. 88
- ง. 188
- จ. 562

11. สมการในข้อใดที่แทนค่าด้วย 4 แล้วทำให้สมการนั้นเป็นจริง

- ก. $9 - n = 13$
- ข. $9 + n = 18$
- ค. $n + 16 = 64$
- ง. $n \times 16 = 64$
- จ. $8 \div n = 32$

12. นายลอยทำถนน 10 ตารางเมตร ราคาตาราง

เมตรละ n บาท นายลอยได้เงินค่าจ้าง 6300 บาท
ราคาทำถนนตารางเมตรละเท่าใด

- ก. 630 บาท
- ข. 6,290 บาท
- ค. 6,310 บาท
- ง. 6,350 บาท
- จ. 63,000 บาท

13. พระเลี้ยงเปิด ช ตัว กรมปศุสัตว์แจกมาอีก 101 ตัว
รวมเป็น 412 ตัวเดิมพระเลี้ยงเปิดกี่ตัว

- ก. 301 ตัว
- ข. 311 ตัว
- ค. 321 ตัว
- ง. 503 ตัว
- จ. 513 ตัว

14. มีน้ำมันในถัง n ลิตร ใช้วันละ 21 ลิตร ใช้ได้ 6 วัน
เดิมมีน้ำมันในถังกี่ลิตร

- ก. 3.5 ลิตร
- ข. 15 ลิตร
- ค. 27 ลิตร
- ง. 126 ลิตร
- จ. 136 ลิตร

15. จะหาคำตอบของสมการ $n + 100 = 102$ ได้โดยวิธีใด

- ก. นำ 100 ไปลบทั้งสองข้างของสมการ
- ข. นำ 100 ไปบวกทั้งสองข้างของสมการ
- ค. นำ 100 ไปหารทั้งสองข้างของสมการ
- ง. นำ 102 ไปลบทั้งสองข้างของสมการ
- จ. นำ 102 ไปบวกทั้งสองข้างของสมการ

16. จะแก้สมการ $n + 105 = 200$ ได้ตามวิธีการในข้อใด

- ก. $n + 105 - 105 = 200$
- ข. $n + 105 - 105 = 200 + 105$
- ค. $n + 105 - 105 = 200 - 105$
- ง. $n + 105 + 105 = 200 + 105$
- จ. $n + 105 + 105 = 200 - 105$

17. $n + 5 = 18$ มีความหมายตรงกับข้อใด

- ก. สุนัขมีผลไม้ n ผล แบ่งให้เด็ก 5 คน ได้รับคนละ 18 ผล
- ข. บานเย็นมีเงิน n บาท ซื้อสมุดไป 5 บาท เหลือเงิน 18 บาท
- ค. แก้วน้ำ n ใบ ราคาใบละ 5 บาท คิดเป็นเงิน 18 บาท
- ง. วิชัยมีขนม n ชิ้น แบ่งให้เพื่อนไป 5 ชิ้น เหลือขนม 18 ชิ้น
- จ. วิชามีเงิน n บาท แม่ให้มาอีก 5 บาท รวมเป็นเงิน 18 บาท

18. $20 - n = 12$ มีความหมายตรงกับข้อใด

- ก. มีฝรั่ง 20 ผล นำไป n ผล เหลือ 12 ผล
- ข. พ่อมีเงิน 20 บาท แม่มีเงิน 12 บาท รวมเป็นเงิน n บาท
- ค. น้องมีเงิน 20 บาท พี่ชายให้อีก 12 บาท รวมเป็นเงิน n บาท
- ง. ป้ากการาคาด้ามละ 12 บาท ซื้อ n ด้าม คิดเป็นเงิน n บาท
- จ. ก้อยมีกล้วย 12 หวี ซื้อมาอีก n หวี รวมมีกล้วยทั้งหมด 20 หวี

19. "แม่มีเงินอยู่ n บาท แจกให้ลูก 3 คนๆ ละ 15 บาท เท่าๆ กัน แม่มีเงินกี่บาท"

เขียนเป็นสมการได้อย่างไร

- ก. $n \div 3 = 15$
- ข. $n \times 3 = 15$
- ค. $n - 3 = 15$
- ง. $15 \times 3 = n$
- จ. $15 + 3 = n$

20. "จัดลูกเสือเป็นแถวได้ n แถว แถวละ 17 คน จากลูกเสือทั้งหมด 204 คน จัดลูกเสือได้กี่แถว"

เขียนเป็นสมการได้อย่างไร

- ก. $n = 204 - 17$
- ข. $n \times 17 = 204$
- ค. $n + 17 = 204$
- ง. $n - 17 = 204$
- จ. $n = 204 \times 17$

แบบทดสอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ฉบับที่ 3

เรื่อง สมการและการแก้สมการ

1. ข้อใดเป็นสมการ

ก. $\frac{4}{6} + \frac{2}{6} > \frac{5}{6}$

ข. $\frac{5}{6} - \frac{4}{5} < \frac{1}{2}$

ค. $\frac{2}{5} + \frac{1}{5} \neq \frac{4}{5}$

ง. $\frac{3}{5} - \frac{1}{10} = \frac{1}{2}$

จ. $\frac{3}{4} + \frac{2}{5} > \frac{8}{20}$

2. ข้อใดเป็นสมการที่เป็นจริง

ก. $29 + 93 = 123$

ข. $5 \times 600 = 3000$

ค. $35 \div 7 = 7 \div 35$

ง. $75 - 25 = 70 + 15$

จ. $25 \times 25 = 25 + 25$

3. สมการในข้อใดไม่มีตัวไม่ทราบค่า

ก. $(33 \div 11) - 3 = 0$

ข. $(6 \times 21) \times 3 = 6 \times (21 \times 3)$

ค. $(58 - 3) + 5 = (58 + 5) - 3$

ง. $(9 + 3) + 11 = 9 + (\square + 11)$

จ. $9 \times (3 + 2) = (9 \times 3) + (9 \times 2)$

4. สมการในข้อใดไม่มีตัวไม่ทราบค่า

ก. $\frac{1}{5} + \frac{3}{5} = \text{ม}$

ข. $39 \times \text{ป} = 9$

ค. $62 \div 2 = 31$

ง. $\text{พ} - 45 = 50$

จ. $72 + \text{ก} = 86$

5. จำนวนในข้อใดเป็นคำตอบของสมการ

$31 + \square = 45$

ก. 4

ข. 14

ค. 24

ง. 74

จ. 76

6. สมการในข้อใดที่แทนค่าด้วย 7 แล้วทำให้สมการนั้นเป็นจริง

ก. $\text{ง} - 4 = 2$

ข. $\text{ง} + 7 = 15$

ค. $\text{ง} - 14 = 7$

ง. $42 + \text{ง} = 48$

จ. $35 - \text{ง} = 28$

7. ผ้าชิ้นหนึ่งยาว 4.50 เมตร ตัดไปใช้ ก เมตร เหลือผ้า 1.70 เมตร ตัดผ้าไปใช้กี่เมตร

ก. 2.01 เมตร

ข. 2.21 เมตร

ค. 2.80 เมตร

ง. 3.20 เมตร

จ. 6.20 เมตร

8. จะหาคำตอบของสมการ $\text{ป} \times 30 = 60$ ได้โดยวิธีใด

ก. นำ 30 มาลบทั้งสองข้างของสมการ

ข. นำ 30 มาบวกทั้งสองข้างของสมการ

ค. นำ 30 มาหารทั้งสองข้างของสมการ

ง. นำ 60 มาลบทั้งสองข้างของสมการ

จ. นำ 60 มาหารทั้งสองข้างของสมการ

9. $\text{จ} \div 5 = 50$ มีความหมายตรงกับข้อใด

ก. มีเงิน จ บาท ซื้อสมุดไป 5 บาท เหลือเงิน 50 บาท

ข. ป้อมมีเงิน จ บาท ให้เด็ก 5 คน เด็กได้รับคนละ 50 บาท

ค. มานี้มีแตงโม จ ผล ขายไป 5 ผล เหลือแตงโม 50 ผล

ง. นักเรียน 50 คน ยืนแถวหน้า จ คน ที่เหลือยืนแถวหลัง 5 คน

จ. วันจันทร์ขายของได้ จ บาท วันอังคารขายได้ 5 บาท รวมเป็น 50 บาท

10. “ฉันอายุ 15 ปี พี่อายุ k ปี ฉันกับพี่มีอายุรวมกันได้ 35 ปี พี่ฉันมีอายุกี่ปี”
เขียนเป็นสมการได้อย่างไร

ก. $15 = 35 + k$

ข. $15 + k = 35$

ค. $15 - k = 35$

ง. $25 - k = 15 + k$

จ. $25 + k = 15 + k$

ภาคผนวก จ
คะแนนที่ได้จากการเทียบคะแนน 3 วิธี

ตาราง 16 คะแนนที่ได้จากการเทียบจากแบบทดสอบฉบับ X ไปสู่แบบทดสอบฉบับ Y ในรูปแบบการเทียบ
คะแนนทั้ง 3 วิธี

คะแนน (X)	คะแนนที่เทียบได้ วิธีที่ 1	คะแนนที่เทียบได้ วิธีที่ 2	คะแนนที่เทียบได้ วิธีที่ 3
1	-0.59	1.06	2.34
2	0.39	1.96	3.33
3	1.37	2.86	4.32
4	2.35	3.76	5.31
5	3.33	4.66	6.30
6	4.31	5.56	7.29
7	5.29	6.46	8.28
8	6.27	7.36	9.27
9	7.25	8.26	10.26
10	8.23	9.16	11.25
11	9.21	10.06	12.24
12	10.19	10.96	13.23
13	11.17	11.86	14.22
14	12.15	12.76	15.21
15	13.13	13.66	16.20
16	14.11	14.56	17.19
17	15.09	15.46	18.18
18	16.07	16.36	19.17
19	17.05	17.26	20.16
20	18.03	18.16	21.25

ภาคผนวก จ

ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน 3 วิธี

ตาราง 17 ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในแต่ละระดับคะแนนในรูปแบบการเทียบทั้ง 3 วิธี

คะแนน (X)	วิธีที่ 1		วิธีที่ 2		วิธีที่ 3	
	คะแนน (Y')	SE (θ_i)	คะแนน (Y')	SE (θ_i)	คะแนน (Y')	SE (θ_i)
1	-0.59	0.55	1.06	0.50	2.34	0.55
2	0.39	0.50	1.96	0.46	3.33	0.50
3	1.37	0.45	2.86	0.41	4.32	0.45
4	2.36	0.40	3.76	0.37	5.31	0.40
5	3.33	0.35	4.66	0.32	6.30	0.36
6	4.31	0.31	5.56	0.28	7.29	0.31
7	5.29	0.26	6.46	0.24	8.28	0.26
8	6.27	0.21	7.36	0.19	9.27	0.21
9	7.25	0.16	8.26	0.15	10.26	0.16
10	8.23	0.11	9.16	0.10	11.25	0.11
11	9.21	0.06	10.06	0.06	12.24	0.06
12	10.19	0.01	10.96	0.01	13.23	0.01
13	11.17	0.03	11.86	0.03	14.22	0.04
14	12.15	0.01	12.76	0.08	15.21	0.08
15	13.13	0.08	13.66	0.12	16.20	0.13
16	14.11	0.18	14.56	0.16	17.19	0.18
17	15.09	0.23	15.46	0.21	18.18	0.23
18	16.07	0.28	16.36	0.25	19.17	0.28
19	17.05	0.32	17.26	0.30	20.16	0.33
20	18.03	0.37	18.16	0.34	21.15	0.38

ภาคผนวก ข
ความเพียงพอของการเทียบคะแนน 3 วิธี

ตาราง 18 ความเพียงพอของการเทียบคะแนนทั้ง 3 วิธี ที่ระดับคะแนนเดียวกัน

คะแนน (X)	$(X_i - X_j)^2$ วิธีที่ 1	$(X_i - X_j)^2$ วิธีที่ 2	$(X_i - X_j)^2$ วิธีที่ 3
5	6.45	0.23	0.37
6	6.20	0.32	0.64
7	5.95	0.44	0.98
8	5.71	0.56	1.39
9	5.48	0.71	1.88
10	5.24	0.86	2.43
11	5.02	1.04	3.06
12	4.80	1.23	3.76
13	4.58	1.44	4.54
14	4.37	1.66	5.38
15	4.16	1.90	6.30
16	3.96	2.16	7.29
17	3.76	2.43	8.35
18	3.57	2.72	9.49
19	3.39	3.03	11.97
20	3.20	3.35	11.97
k	16	16	16
SD	3.91	4.32	4.09
c	0.3117	0.0774	0.2805
$(.05SD)^2$	0.0382	0.0467	0.0418

ประวัติย่อผู้วิจัย

ชื่อ ชื่อสกุล	นางสาวทิพาภรณ์ ทุงฤทธิ์
วันเดือนปีเกิด	2 ตุลาคม 2515
สถานที่เกิด	อำเภอเมือง จังหวัดปราจีนบุรี
สถานที่อยู่ปัจจุบัน	55 หมู่ 3 ตำบลเนินหอม อำเภอเมือง จังหวัดปราจีนบุรี 25230
ตำแหน่งหน้าที่การงานในปัจจุบัน	อาจารย์ 1 ระดับ 4
สถานที่ทำงานปัจจุบัน	โรงเรียนวัดดงบัง (ดงบังบำรุงวิทย์) อำเภอประจันตคาม จังหวัดปราจีนบุรี
ประวัติการศึกษา	พ.ศ.2533 มัธยมศึกษาตอนปลาย โรงเรียนปราจีนกัลยาณี จังหวัดปราจีนบุรี พ.ศ.2537 การศึกษามัธยมศึกษา (การประถมศึกษา) มหาวิทยาลัยบูรพา จังหวัดชลบุรี พ.ศ.2544 การศึกษามหาบัณฑิต (วัดผลการศึกษา) มหาวิทยาลัยศรีนครินทรวิ โรฒ

การเปรียบเทียบผลการเทียบคะแนนโดยวิธีเชิงเส้นตรง เมื่อใช้วิธีการแปลงคะแนนที่แตกต่างกัน
ในแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์

บทคัดย่อ
ของ
ทิพาภรณ์ ทุงฤทธิ์

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
มีนาคม 2544

การศึกษาค้นคว้าครั้งนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบคะแนนที่เทียบได้ ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน และความเพียงพอของการเทียบคะแนนในแนวนอน โดยใช้วิธีเชิงเส้นตรง เมื่อใช้การเทียบคะแนนที่ใช้วิธีการแปลงคะแนนที่แตกต่างกัน 3 วิธี ดังนี้

1. การแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ
2. การแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบทั้งสองฉบับที่เรียงอันดับการสอบต่างกัน
3. การแปลงคะแนนจากการสุ่มกลุ่มนักเรียนให้สอบโดยใช้แบบทดสอบกลุ่มละฉบับ และมีแบบทดสอบรวมที่ต้องสอบทั้งสองกลุ่ม

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลเป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ เรื่อง สมการและการแก้สมการ จำนวน 3 ฉบับ ประกอบด้วย ฉบับ X ฉบับ Y และฉบับรวม โดยมีกลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2542 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษาจังหวัดปราจีนบุรี จำนวน 820 คน ซึ่งได้มาจากการสุ่มแบบแบ่งชั้น

ผลการวิจัยสรุปได้ดังนี้

1. คะแนนที่ได้จากการเทียบคะแนน 3 วิธี แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
2. ความคลาดเคลื่อนมาตรฐานของการเทียบคะแนน 3 วิธี แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ
3. ความเพียงพอของการเทียบคะแนน 3 วิธี แตกต่างกัน

A COMPARISON OF DIFFERENT LINEAR EQUATING PROCEDURES ON
THE MATHEMATICS ACHIEVEMENT TEST SCORES.

AN ABSTRACT
BY
TIPAPORN TOONGRIT

Presented in partial fulfillment of the requirements
for the Master of Education degree in Educational Measurement
at Srinakarinwirot University
March 2001

The purpose of this study were to compare the equating scores, the standard error of equating and the discrepancy of equating with three different linear equating procedures; 1) random group-one test administered to each group, 2) random group-both forms administered to each group, counterbalanced and 3) random groups-one administered to each group, common equating test administered to both groups. The instruments were 3 subtests that composed of test X, test Y and common equating of mathematics achievement test. The samples were 820 students of prathom 6 in the academic year 1999 from primary schools in Prachinburi and selected by stratified random sampling.

The result of this study showed that

- 1) The equating scores from three linear equating procedures were different with significant at .01 level.
- 2) The standard errors of equating scores from three linear equating procedures were not statistical significant different.
- 3) The discrepancy of equating from three linear equating procedures was different and the second procedures had satisfied of discrepancy in the equating scores.